

# Statistische Inferenz für Performancemaße

Carsten Körner

Universität zu Köln 2013

Referent: Univ.-Prof. Dr. Friedrich Schmid  
Korreferent: Univ.-Prof. Dr. Karl Mosler  
Tag der Promotion: 20. Juni 2013

# Statistische Inferenz für Performancemaße

Inauguraldissertation  
zur  
Erlangung des Doktorgrades  
der  
Wirtschafts- und Sozialwissenschaftlichen Fakultät  
der  
Universität zu Köln

2013

vorgelegt  
von

Diplom-Kaufmann  
Carsten Körner

aus

Huchem-Stammeln



# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>1</b>
<b>2</b>	<b>Renditen</b>	<b>5</b>
2.1	Renditebegriffe . . . . .	5
2.2	Renditemodelle . . . . .	7
2.2.1	Unabhängig identisch verteilte Renditen . . . . .	7
2.2.2	Intertemporal abhängige Renditen . . . . .	10
2.2.2.1	Autokorrelation . . . . .	10
2.2.2.2	Abhängige Varianzen . . . . .	12
2.2.3	Multivariate Renditemodelle . . . . .	14
2.2.3.1	Multivariate unabhängig und identisch verteilte Zufallsvariablen . . . . .	15
2.2.3.2	MGARCH . . . . .	17
2.3	Empirische Renditen . . . . .	22
2.3.1	Aktien und andere börsengehandelte Wertpapiere . . . . .	22
2.3.1.1	Stilisierte Fakten . . . . .	22
2.3.1.2	Renten und Derivate . . . . .	25
2.3.2	Investmentfonds . . . . .	26
2.3.3	Hedgefonds . . . . .	33
<b>3</b>	<b>Performancemessung</b>	<b>43</b>
3.1	Rendite, Risiko und Nutzen . . . . .	43
3.2	Performancemaße . . . . .	48
3.2.1	Überblick und Klassifikation . . . . .	48
3.2.2	Die klassischen Performancemaße . . . . .	49
3.2.2.1	Das Sharpe-Ratio . . . . .	49
3.2.2.2	Das Treynor-Ratio . . . . .	50
3.2.2.3	Jensens Alpha . . . . .	52
3.2.3	Neuere Performancemaße . . . . .	53
3.2.3.1	Risk-Adjusted Performance nach Modigliani / Modigliani . . . . .	53
3.2.3.2	Performancemaße auf Basis partieller Momente . . . . .	54
3.2.3.3	Performancemaße auf Basis des Value-at-Risk . . . . .	57
3.3	Punktschätzung von Performancemaßen . . . . .	59

---

3.4	Risikoloser Zins, Marktportfolio und Zeitskala . . . . .	62
3.4.1	Der risikolose Zins . . . . .	62
3.4.2	Das Marktportfolio . . . . .	64
3.4.3	Die Zeitskala . . . . .	66
3.5	Studien zur Performancemessung . . . . .	67
<b>4</b>	<b>Statistische Inferenz</b>	<b>69</b>
4.1	Die Delta-Methode . . . . .	70
4.1.1	Ansatz . . . . .	70
4.1.2	Varianzen für einzelne Maße . . . . .	71
4.1.2.1	Sharpe-Ratio . . . . .	71
4.1.2.2	Treynor-Ratio . . . . .	77
4.1.2.3	Jensens Alpha . . . . .	80
4.1.2.4	Modigliani-Modigliani . . . . .	83
4.1.2.5	Kappa . . . . .	88
4.1.2.6	Excess Return on Value-at-Risk . . . . .	92
4.1.3	Konfidenzintervalle auf Basis der Varianzformeln . . . . .	98
4.2	Die Bootstrap-Methode . . . . .	98
4.2.1	Der nichtparametrische Bootstrap . . . . .	99
4.2.2	Bootstrap-Konfidenzintervalle . . . . .	101
4.2.2.1	Bootstrap-Konfidenzintervalle über Standardfehler . . . . .	101
4.2.2.2	Bootstrap-Konfidenzintervalle über Perzentile . . . . .	103
4.2.3	Bootstrapping von Performancemaßen . . . . .	105
4.3	Weitere Ansätze . . . . .	106
<b>5</b>	<b>Simulationen</b>	<b>107</b>
5.1	Methodik . . . . .	107
5.1.1	Datenerzeugende Prozesse . . . . .	107
5.1.2	Auswertung . . . . .	109
5.2	Konvergenz der Schätzer . . . . .	114
5.3	Ergebnisse für einzelne Maße . . . . .	120
5.3.1	Sharpe-Ratio . . . . .	120
5.3.2	Treynor-Ratio . . . . .	134
5.3.3	Jensens Alpha . . . . .	142
5.3.4	Modigliani-Modigliani . . . . .	148
5.3.5	Kappa . . . . .	156
5.3.6	Expected Return on Value-at-Risk . . . . .	162
5.4	Konfidenzintervalle . . . . .	167
<b>6</b>	<b>Schlussfolgerungen</b>	<b>181</b>
6.1	Zusammenfassung . . . . .	181
6.2	Kann man Performancemaße aussagekräftig schätzen? . . . . .	184

---

<b>A Anhang</b>	<b>191</b>
A.1 Umformungen höherer Momente . . . . .	191
A.2 Varianzformeln für Performancemaße . . . . .	193
A.2.1 Sharpe-Ratio . . . . .	193
A.2.2 Treynor-Ratio . . . . .	193
A.2.3 Jensens Alpha . . . . .	195
A.2.4 Modigliani-Modigliani . . . . .	197
A.2.5 Kappa . . . . .	198
A.2.6 ERVaR . . . . .	198
A.3 Weitere, ergänzende Berechnungen . . . . .	199
A.3.1 Überleitung von $\sigma_{\hat{\kappa}_{1,z}}^2$ zu $\sigma_{\hat{\Omega}_z}^2$ . . . . .	199
A.3.2 Erläuterung von Formel 5.1, S. 113 . . . . .	200
<b>Literaturverzeichnis</b>	<b>201</b>





# Tabellenverzeichnis

2.1	Kennzahlen der Monatsrenditen von 2 378 amerikanischen Investmentfonds 2001-2011 . . . . .	30
2.2	Vergleich der Kennzahlen für Investmentfondsmonatsrenditen . . . . .	32
2.3	Kennzahlen für Hedgefondsmonatsrenditen nach Getmansky et al. (2004) . . . . .	39
3.1	Punktschätzer für klassische Performancemaße . . . . .	60
3.2	Korrelation der Schätzer ausgewählter Performancemaße nach Eling und Schuhmacher (2007) . . . . .	68
5.1	Parameter bei der Simulation von Zeitreihen . . . . .	109
5.2	Übersicht über die in den folgenden Abschnitten verwendeten Simulationsstatistiken . . . . .	114
5.3	Simulationsparameter in Abschnitt 5.2 . . . . .	115
5.4	Simulationsergebnisse für die benchmarkfreien Maße bei i.i.d.-normalverteilten Daten und Zeitreihenlänge $n = 500$ . . . . .	116
5.5	Simulationsergebnisse für die Maße mit Benchmark bei i.i.d.-normalverteilten Daten und Zeitreihenlänge $n = 500$ . . . . .	118
5.6	Simulationsergebnisse für Differenzen von Performancemaßen bei i.i.d.-normalverteilten Daten und Zeitreihenlänge $n = 500$ . . . . .	119
5.7	Simulationsergebnisse bei i.i.d.- $t_5$ -verteilten Daten und Zeitreihenlänge $n = 500$ . . . . .	121
5.8	Ergebnisse für das Sharpe-Ratio bei i.i.d.-normalverteilten Daten und Zeitreihenlänge $n = 60$ , Ausgangsparameter . . . . .	122
5.9	Ergebnisse für das Sharpe-Ratio bei nicht normalverteilten Daten und Zeitreihenlänge $n = 60$ , Ausgangsparameter . . . . .	124
5.10	Ergebnisse für die Differenz der Sharpe-Ratios zweier Portfolios $X$ und $Y$ mit Zeitreihenlänge $n = 60$ , Ausgangsparameter . . . . .	130
5.11	Ergebnisse für das Treynor-Ratio bei Zeitreihenlänge $n = 60$ , Ausgangsparameter . . . . .	135
5.12	Ergebnisse für die Differenz der Treynor-Ratios zweier Portfolios bei Zeitreihenlänge $n = 60$ , Ausgangsparameter . . . . .	141
5.13	Ergebnisse für Jensens Alpha bei Zeitreihenlänge $n = 60$ , Ausgangsparameter . . . . .	143

5.14	Ergebnisse für die Differenz der Jensens Alphas zweier Portfolios bei Zeitreihenlänge $n = 60$ , Ausgangsparameter . . . . .	146
5.15	Ergebnisse für die Originalversion von $MM$ bei Zeitreihenlänge $n = 60$ , Ausgangsparameter . . . . .	148
5.16	Ergebnisse für die alternative Version von $MM$ bei Zeitreihenlänge $n = 60$ , Ausgangsparameter . . . . .	151
5.17	Ergebnisse für die Differenz der $MM$ zweier Portfolios bei Zeitreihenlänge $n = 60$ , Ausgangsparameter . . . . .	154
5.18	Ergebnisse für $Kappa(\alpha, z = 0)$ bei Zeitreihenlänge $n = 60$ , Ausgangsparameter . . . . .	157
5.19	Ergebnisse für die Differenz von $Kappa(\alpha, z = 0)$ zweier Portfolios bei Zeitreihenlänge $n = 60$ , Ausgangsparameter . . . . .	162
5.20	Ergebnisse für $ERVaR(\alpha = 0, 05)$ bei Zeitreihenlänge $n = 60$ , Ausgangsparameter . . . . .	164
5.21	Annahmebereiche von $H_0$ bei zweiseitigem Gauß-Anteilswerttest	168
5.22	Überdeckungshäufigkeiten verschiedener Konfidenzintervalle für $Sh_X$ und $\Delta Sh_{XY}$ . . . . .	170
5.23	Überdeckungshäufigkeiten verschiedener Konfidenzintervalle für $Tr_X$ und $\Delta Tr_{XY}$ . . . . .	171
5.24	Überdeckungshäufigkeiten verschiedener Konfidenzintervalle für $Tr_X$ bei $\rho_{M,X} = 0, 25$ . . . . .	172
5.25	Überdeckungshäufigkeiten verschiedener Konfidenzintervalle für $\alpha_X$ und $\Delta \alpha_{XY}$ . . . . .	173
5.26	Überdeckungshäufigkeiten verschiedener Konfidenzintervalle für Schätzer des Performancemaßes von Modigliani-Modigliani . . . . .	174
5.27	Überdeckungshäufigkeiten verschiedener Konfidenzintervalle für $\kappa_X(\alpha = 1, 2, 3)$ . . . . .	176
5.28	Überdeckungshäufigkeiten verschiedener Konfidenzintervalle für $ERVaR_X(\alpha = 0, 05; 0, 02)$ . . . . .	176
5.29	Vergleich der Schätzmethode für Konfidenzintervalle . . . . .	177
5.30	Überdeckungshäufigkeiten für Konfidenzintervalle des Sharpe-Ratios bei Variation von $n$ und $B$ . . . . .	180
6.1	Übersicht über vorhandene Varianzformeln für Performancemaßschätzer . . . . .	182
6.2	Wahrscheinlichkeiten für falsches Vorzeichen von $\hat{\theta}$ . . . . .	186
6.3	Asymptotische Varianzen von Performancemaßen bei verschiedenen erwarteten Renditen und Stichprobenlängen . . . . .	187
6.4	Minimale erwartete Portfoliorendite pro Monat bzw. Jahr für eine mit mind. 84% Wahrscheinlichkeit korrekte Schätzung des Vorzeichens von $\bar{X}$ . . . . .	189

# Abbildungsverzeichnis

2.1	Histogramm der Hedgefondsrenditen 1994 - 2005 aus CISDM . . .	38
4.1	Nichtparametrischer Dichteschätzer . . . . .	96
4.2	Abhängigkeit des Parameters $k$ von $n \cdot \alpha$ . . . . .	97
5.1	Grafische Erläuterung zu $\psi_{f_1, f_2}$ . . . . .	112
5.2	Grafiken zur Verteilung von $\widehat{Sh}$ für i.i.d.-normalverteilte Daten mit $Sh = 0,02$ und Zeitreihenlänge $n = 60$ . . . . .	123
5.3	Grafiken zur Verteilung von $\widehat{Sh}$ für i.i.d.- $t_{10}$ -verteilte Daten mit $Sh = 0,02$ und Zeitreihenlänge $n = 60$ . . . . .	125
5.4	Grafiken zur Verteilung von $\widehat{Sh}$ für GARCH- und AR-Zeitreihen mit $Sh = 0,02$ und Zeitreihenlänge $n = 60$ . . . . .	127
5.5	Simulationsstatistiken des Sharpe-Ratios bei Variation von $\mu_X$ für i.i.d.-Daten (links) und AR(1)-Daten (rechts) . . . . .	129
5.6	Grafiken zur Verteilung von $\widehat{\Delta}_{Sh}$ , i.i.d.- $t_{10}$ -verteilte Renditen, Zeitreihenlänge $n = 60$ . . . . .	132
5.7	Grafiken zu $\widehat{\Delta}_{Sh}$ bei Parametervariation . . . . .	133
5.8	Grafiken zur Verteilung von $\widehat{Tr}$ für i.i.d.- $t_{10}$ -verteilte Daten mit $Tr = 0,004$ und Zeitreihenlänge $n = 60$ . . . . .	136
5.9	Grafiken zu $\widehat{Tr}_X$ bei Parametervariation . . . . .	138
5.10	Grafiken zu $\widehat{Tr}_X$ bei Variation des Korrelationskoeffizienten und $n = 1250$ . . . . .	140
5.11	Grafiken zur Verteilung von $\widehat{\Delta}_{Tr}$ , i.i.d.- $t_{10}$ -verteilte Renditen, Zeitreihenlänge $n = 60$ . . . . .	142
5.12	Grafiken zur Verteilung von Jensens $\widehat{\alpha}$ für i.i.d.- $t_{10}$ -verteilte Daten mit $\alpha = 0,0015$ und Zeitreihenlänge $n = 60$ . . . . .	144
5.13	Grafiken zu $\widehat{\alpha}_X$ bei Parametervariation . . . . .	145
5.14	Grafiken zur Verteilung von $\widehat{\Delta}_{\alpha}$ , i.i.d.- $t_{10}$ -verteilte Renditen, Zeitreihenlänge $n = 60$ . . . . .	147
5.15	Grafiken zur Verteilung von $\widehat{MM}_{ori}$ für i.i.d.- $t_{10}$ -verteilte Daten mit Zeitreihenlänge $n = 60$ . . . . .	149
5.16	Grafiken zu $\widehat{MM}_{ori}(X)$ bei Parametervariation . . . . .	150

5.17	Grafiken zur Verteilung von $\widehat{MM}_{alt}$ für i.i.d.- $t_{10}$ -verteilte Daten mit Zeitreihenlänge $n = 60$ . . . . .	152
5.18	Grafiken zu $\widehat{MM}_{alt}(X)$ bei Parametervariation . . . . .	153
5.19	Streudiagramme zu Modigliani-Modigliani mit 95%- und 99%-NV-Zentralbereichen; i.i.d.- $t_{10}$ -verteilte Renditen, Zeitreihenlänge $n = 60$ , je 2000 Zeitreihen. . . . .	154
5.20	Grafiken zur Verteilung von $\widehat{\Delta}_{MM_{X,Y}}$ für i.i.d.- $t_{10}$ -verteilte Daten mit Zeitreihenlänge $n = 60$ . . . . .	155
5.21	Grafiken für $\text{Kappa}_X(\alpha)$ bei verschiedenen Werten für $\alpha$ mit i.i.d.- $t_{10}$ -verteilten Daten und Zeitreihenlänge $n = 60$ : Histogramme (20 000 Werte, links); Punktdiagramme $\widehat{\kappa}_X(\alpha, z = 0)$ gegen $\widehat{\sigma}_{\kappa_X(\alpha, z=0)}$ (2 000 Werte, rechts). . . . .	159
5.22	Grafiken zu $\widehat{\kappa}_X(\alpha)$ bei Variation von $\mu_X$ . . . . .	161
5.23	Grafiken für $\Delta_{\kappa_X}(\alpha)$ bei verschiedenen Werten für $\alpha$ mit i.i.d.- $t_{10}$ -verteilten Daten und Zeitreihenlänge $n = 60$ . . . . .	163
5.24	Grafiken zur Verteilung von $\widehat{ERV}aR_X$ für i.i.d.- $t_{10}$ -verteilte Daten mit Zeitreihenlänge $n = 60$ . . . . .	165
5.25	Grafiken zu $\widehat{ERV}aR_X(\alpha)$ bei Variation von $\mu_X$ . . . . .	166

# Kapitel 1

## Einleitung

Die Performancemessung für Wertpapiere und andere Vermögenswerte gehört schon lange zu den Teilgebieten der Finanzmarktforschung. Als sich die Portfoliotheorie in den 1950er und 1960er Jahren etablierte, wuchs das Bewusstsein dafür, dass man Vermögenswerte nicht allein nach ihrer erwarteten Rendite beurteilen sollte, sondern auch das Risiko, das ein Anleger durch Investition in bestimmte Anlageobjekte eingeht, geeignet einzubeziehen ist. Vor bald fünfzig Jahren veröffentlichten Jack L. Treynor (1965), William F. Sharpe (1966) und Michael C. Jensen (1968) Vorschläge für Performancemaße, d.h. Kennzahlen zur risikoadjustierten Messung und Bewertung der Wertentwicklung von Kapitalanlagen: Treynor-Ratio, Sharpe-Ratio und Jensens Alpha. Hat man Kenntnisse über bestimmte Renditeeigenschaften verschiedener Portfolios oder trifft Annahmen darüber, kann man durch die Anwendung von Performancemaßen Urteile über die Vorteilhaftigkeit der Portfolios, ihre Rangfolge untereinander und die optimale Auswahl aus ihnen fällen. Will man solche Annahmen fundieren oder die vergangene Performance bestimmter Portfolios bzw. Portfoliomanager bewerten, muss man die Performancemaße aus beobachteten Finanzmarktdaten schätzen.

Nach einiger Zeit begann sich die Wissenschaft für die statistischen Eigenschaften dieser Schätzer zu interessieren. Miller und Gehr (1978) zeigten, dass der Schätzer des Sharpe-Ratios verzerrt ist und ermittelten dessen Erwartungswert für unabhängig und identisch normalverteilte Renditen. In einem grundlegenden Artikel leiteten Jobson und Korkie (1981) die asymptotischen Varianzen der Schätzer für Sharpe-Ratio und Treynor-Ratio her. Vinod und Morey (2000) wendeten Bootstrap-Ansätze zur Schätzung von Konfidenzintervallen für das Treynor-Ratio an. Seit diesen Arbeiten wurden einige Verfeinerungen, Erweiterungen und Korrekturen veröffentlicht, aber bisher fehlt eine zusammenhängende Darstellung über die Berechnung von Varianzen und Konfidenzintervallen für eine größere Zahl von Performancemaßen.

Diese Arbeit bietet Formeln für asymptotische Varianzen der Schätzer von Sharpe-Ratio, Treynor-Ratio, Jensens Alpha, Risk-Adjusted Performance nach Modigliani/Modigliani (MM), Omega, Sortino-Ratio und Excess Return on Value-

at-Risk für unabhängig und identisch verteilte Renditen sowie für die Differenzen dieser Maße beim Vergleich zweier Portfolios.<sup>1</sup> Auf dieser Basis sowie mittels verschiedener Bootstrap-Varianten werden Konfidenzintervalle konstruiert. Alle Ansätze werden ausführlich mittels Monte-Carlo-Simulation getestet und miteinander verglichen, wobei auch die Verteilung der Schätzer unter verschiedenen Renditemodellen dargestellt wird. Die Untersuchung dieser statistischen Eigenschaften der Performancemaßschätzer wird ergänzt um eine ökonomische Betrachtung: Wie genau können Performancemaße geschätzt werden? Inwieweit kann man mittels Schätzung von Performancemaßen gut und weniger gut gemanagte Portfolios unterscheiden?

Diese Arbeit ist wie folgt gegliedert: Kapitel 2 befasst sich mit der Berechnung von Renditen, verschiedenen Ansätzen zur Modellierung von Renditen und Renditezeitreihen sowie einer Sichtung der Literatur zu den Eigenschaften der Renditen von Aktien, Publikumsfonds und Hedgefonds. In Kapitel 3 werden die in dieser Arbeit behandelten Performancemaße und deren Schätzer vorgestellt. Kapitel 4 ist der theoretische Hauptteil dieser Arbeit und beinhaltet die Herleitung der asymptotischen Varianzen der o. g. Performancemaße mittels der Delta-Methode sowie eine Einführung in die Schätzung von Konfidenzintervallen per nichtparametrischem Bootstrap. Kapitel 5 enthält für jedes einzelne Performancemaß Simulationen auf Basis verschiedener Verteilungsmodelle, um die Genauigkeit der Varianzschätzungen, die Anpassung der Verteilung der Schätzer an die Normalverteilung und die Einhaltung des Konfidenzniveaus durch verschiedene Konfidenzintervall-Schätzverfahren zu prüfen. Das Schlusskapitel 6 fasst die Ergebnisse zusammen und diskutiert deren Konsequenzen für die ökonomische Praxis. Im Anhang A.2 findet der Leser eine Auflistung aller in Kapitel 4 hergeleiteten Formeln.

Diese Arbeit greift wie jede wissenschaftliche Veröffentlichung auf Vorarbeiten anderer Wissenschaftler zurück, vor denen man einige gesondert aufgreifen sollte, um die originalen Leistungen dieser Arbeit von den übernommenen abzugrenzen: Die Idee, asymptotische Varianzen von Performancemaß-Schätzern mittels der Delta-Methode zu ermitteln, stammt, wie oben erwähnt, von J. D. Jobson und Bob M. Korkie. Ihr Ansatz für Sharpe-Ratio und Treynor-Ratio wird in Kapitel 4 nachvollzogen. Für die asymptotischen Varianzen von Jensens Alpha, die Risk-Adjusted Performance (MM) und Omega konnte ich auf teilweise recht weitreichende Vorarbeiten meines Doktorvaters Prof. Dr. Friedrich Schmid und Herrn PD Dr. Rafael Schmidt zurückgreifen.<sup>2</sup>

Im wesentlichen Anteil oder vollständig selbstständig erarbeitete neue Leistungen in dieser Dissertation sind:

- Ausarbeitung und Anwendung der Varianzansätze für Jensens Alpha, MM und Kappa<sup>3</sup> (Kap. 4),

---

<sup>1</sup>Für den Excess Return on Value-at-Risk wird keine Varianzformel für den Schätzer der Differenz dieses Performancemaßes für zwei Portfolios gezeigt.

<sup>2</sup>Siehe hierzu z. B. Schmid und Schmidt (2008)

<sup>3</sup>Kappa ist ein verallgemeinertes Performancemaß, zu dessen Spezialfällen Omega und Sortino-Ratio gehören.

- Konzeption, Ausarbeitung und Anwendung der Varianzformel für den Excess Return on Value-at-Risk (Kap. 4),
- Überprüfung der Varianzformeln aller genannter Performancemaße per Simulation (Kap. 5),
- Vergleich von Konfidenzintervallen auf Basis der hergeleiteten Varianzformeln mit verschiedenen Bootstrap-Konfidenzintervallen für alle genannten Performancemaße (Kap. 5).

Eher hilfswise entstanden, aber für den einen oder anderen Leser vielleicht nicht uninteressant, sind ein einfaches multivariates Renditemodell mit GARCH-Varianzen (Abschnitt 2.2.3.2, S. 20 f.) sowie ein Maß für die Abweichung einer empirischen Verteilung von einer Normalverteilung (Abschnitt 5.1.2, S. 111 ff.).





# Kapitel 2

## Renditen

Inhalte dieses Kapitels sind der Begriff der Rendite, verschiedene Modelle zur Darstellung von Finanzmarktrenditen und die konkrete Eignung bestimmter Verteilungsmodelle für die in der Performancemessung relevanten Renditedaten. Häufigster Gegenstand von Performancemessungen sind von Managern zusammengestellte Anlageportfolios, insbesondere Investmentfonds und Hedgefonds. Die Eigenschaften dieser Fondstypen und ihrer Renditen werden im letzten Teil des Kapitels behandelt.

### 2.1 Renditebegriffe

Als Rendite einer Anlage wird gemeinhin der Erfolg  $W_t - W_0 = \Delta W$  einer Anlage in einer Anlageperiode  $\Delta t = t - t_0$ , geteilt durch das in  $t_0$  eingesetzte Kapital  $W_0$  verstanden.

$$R_{D;0,t} = \frac{\Delta W}{W_0} = \frac{W_0 + \Delta W}{W_0} - 1 = \frac{W_t}{W_0} - 1$$

Die Renditedefinition  $R_D$  wird im Allgemeinen 'diskrete Rendite' genannt. Eine andere, häufiger angewendete Renditedefinition ist die 'stetige Rendite'

$$R_{S;0,t} = \ln\left(\frac{W_t}{W_0}\right) = \ln(W_t) - \ln(W_0) = \ln(R_D + 1).$$

Stetige Renditen haben den Vorteil, dass sie im Gegensatz zu diskreten Renditen additiv aggregierbar sind, während diskrete Renditen multiplikativ aggregiert werden.

$$R_{D;0,2} = \frac{W_2}{W_0} - 1 = \frac{W_2}{W_1} \frac{W_1}{W_0} - 1 = (R_{D;1,2} + 1)(R_{D;0,1} + 1) - 1$$

$$R_{S;0,2} = \ln\left(\frac{W_2}{W_0}\right) = \ln\left(\frac{W_2}{W_1} \frac{W_1}{W_0}\right) = \ln\left(\frac{W_2}{W_1}\right) + \ln\left(\frac{W_1}{W_0}\right) = R_{S;1,2} + R_{S;0,1}$$

Durch die lineare Aggregierbarkeit stetiger Renditen können sämtliche üblichen Rechenoperationen (arithmetischer Mittelwert, Varianzen, Regressionen) mit den bekannten Formeln durchgeführt werden. Unter anderem aus diesem Grund<sup>1</sup> wird in der gesamten Arbeit der stetige Renditebegriff verwendet. Zur Vereinfachung der Notation sei  $R_t$ , wenn nicht anders angegeben, gleichbedeutend mit  $R_{S;t-1,t}$ .

In dieser Arbeit werden Renditezeitreihen betrachtet, also durch einen Zeitindex  $t \in \mathbb{Z}$  geordnete Renditen. Im Allgemeinen betrachten wir Zeitreihen, die einen Zeitraum von  $t = 0$  bis  $T$  umfassen, also mit der Rendite  $R_1 = R_{0,1}$  beginnen, mit  $R_T = R_{T-1,T}$  enden und den Zeitraum von 0 bis  $T$  lückenlos abdecken.

In der Arbeit werden nur zeitdiskrete Modelle verwendet, also Modelle, die nur zu festen, höchstens abzählbar unendlich vielen Zeitpunkten Werte aufweisen. Außerdem wird stets von äquidistanten Zeitreihen ausgegangen: Der Abstand zweier aufeinanderfolgender Zeitpunkte sei innerhalb einer Zeitreihe immer konstant.<sup>2</sup> Die kürzeste modellierte Zeitspanne ist der Abstand zwischen zwei Börsentagen (Tagesrenditen).

Man kann Renditen als Ergebnisse von Zufallsversuchen oder -prozessen auffassen und beobachtete Renditen daher als Realisierungen von Zufallsvariablen.  $R_X$  bezeichnet in diesem Sinne die zufällige Rendite eines Portfolios  $X$ , eine Realisierung der Zufallsvariable wird mit  $r_X$  bezeichnet. Die Zufallsvariable  $R_X$  hat die Momente Erwartungswert  $\mu_{R_X}$  und Varianz  $\sigma_{R_X}^2$ , sofern einige recht milde Anforderungen an die Verteilung von  $R_X$  erfüllt sind.

In der Performancemessung wird oft mit *Überrenditen* gerechnet – Rohrenditen, von denen der Ertrag einer sicheren Geldanlage abgezogen wurde. Überrenditen darstellende Zufallsvariablen werden in dieser Arbeit zur Vereinfachung der Notation mit einfachen Großbuchstaben  $X, Y$  usw. bezeichnet.

In den folgenden Abschnitten stelle ich einige Modellierungsansätze für Finanzmarktrenditen vor und werde dabei ggf. kurz auf deren Motivation eingehen. Ein Vergleich mit empirischen Erkenntnissen über das Verhalten von Finanzmarktrenditen erfolgt zusammengefasst in einem anschließenden Abschnitt.

---

<sup>1</sup>Weiterhin haben diskrete Renditen den Nachteil, dass ihr Wertebereich bei Vermögensbestandteilen, deren Wert nicht negativ werden kann, auf das Intervall  $[-1; \infty)$  beschränkt ist. Dies gilt für die meisten börsengehandelten Vermögenswerte (Aktien, Anleihen, Fonds, Long-Optionen). Stetige Renditen dagegen haben den Wertebereich  $(-\infty; \infty)$ . Verschiedene mathematische Modelle (z.B. normalverteilte Renditen) gehen von unbeschränkten Wertebereichen aus. Einen Nachteil haben stetige Renditen bei der Abbildung wertlos gewordener Vermögensgegenstände (z.B. Aktien von nach Insolvenz liquidierten Unternehmen oder verfallene Optionen), bei denen  $R_{S;t,\text{Endzeitpunkt}} = -\infty$ . Solche Vermögenswerte werden in dieser Arbeit nicht betrachtet und sind in der Praxis auch selten Gegenstand von Performancemessungen.

<sup>2</sup>Praktisch ist das in vielen Fällen nicht ganz erfüllt. Benutzt man Monatsrenditen, so haben die einzelnen Monate unterschiedliche Längen zwischen 28 und 31 Tagen. Bei Tagesrenditen, berechnet etwa auf Basis der Börsenschlusskurse zweier aufeinanderfolgender Börsentage, nimmt man mit der Annahme der Äquidistanz an, dass sich Renditen zwischen zwei Werktagen nicht von denjenigen über das Wochenende von Freitag auf Montag oder über Feiertage hinweg unterscheiden. Die Feiertage können sich zudem von Börsenplatz zu Börsenplatz unterscheiden.

## 2.2 Renditemodelle

Bei stochastischen Modellen für Renditezeitreihen sind zwei Aspekte zu modellieren: einerseits das Verteilungsmodell der Rendite zu einem gegebenen Zeitpunkt, andererseits die zeitliche Abhängigkeitsstruktur zwischen den Renditen. Wir beginnen mit unabhängig und identisch verteilten Renditen und stellen einige Verteilungsmodelle vor. Anschließend gehen wir auf zwei Arten der Abhängigkeit zeitlich aufeinanderfolgender Renditen ein. Der dritte Unterabschnitt beschäftigt sich mit Verallgemeinerungen für die simultane Modellierung mehrerer, ggf. voneinander abhängiger Zeitreihen.

### 2.2.1 Unabhängig identisch verteilte Renditen

In diesem ersten Unterabschnitt machen wir die Annahme, die Komponenten einer (univariaten) Renditezeitreihe seien voneinander global unabhängig und die Renditeverteilung der einzelnen Komponenten sei identisch. Entsprechend sind Erwartungswert, Varianz und alle anderen Eigenschaften jeder einzelnen Zeitreihenkomponente gleich. Diese Annahme, abgekürzt *i.i.d.* (für *independent and identically distributed*), ermöglicht eine direkte vollständige Modellierung der Verteilung der Rendite einer Periode, ohne dass Einflüsse von Renditen anderer Perioden berücksichtigt werden müssen.

Die Annahme **unabhängig identisch normalverteilter** Renditen stellt die einfachste hier verwendete Verteilungsannahme für Portfoliorenditen dar. Die Normalverteilung hat eine Fülle von Eigenschaften, die ihre Verwendung auch in der Finanzmarktstatistik sehr angenehm machen:

- Die Normalverteilung ist vollständig durch ihre zwei Parameter  $\mu$  und  $\sigma$  definiert, die gleichzeitig Erwartungswert und Standardabweichung der Verteilung sind.
- Linearkombinationen von Zufallsvariablen, die gemeinsam multivariat normalverteilt sind, sind wiederum normalverteilt. Da unabhängige univariat normalverteilte Zufallsvariablen multivariat normalverteilt sind (mit einer Diagonalmatrix als Kovarianzmatrix), gilt dies auch für Summen unabhängiger univariat normalverteilter Zufallsvariablen:  
Ist  $Z = \sum_{i=1}^n a_i X_i$  mit  $X_i \sim N(\mu_i, \sigma_i^2)$  und seien alle  $X_i$  global unabhängig, dann ist

$$Z \sim N\left(\sum_{i=1}^n a_i \mu_i, \sum_{i=1}^n a_i^2 \sigma_i^2\right)$$

(*Reproduktionseigenschaft der Normalverteilung*). Somit verbleibt man bei der Aggregation von Renditen innerhalb derselben Lage-Skalen-Familie.

- Nach dem *Zentralen Grenzwertsatz* konvergieren auch Summen von  $n$  unabhängig identisch nicht normalverteilter Zufallsvariablen für  $n \rightarrow \infty$  gegen eine Normalverteilung, solange die ersten beiden Momente der Ausgangsverteilung endlich sind.

Seien  $X_i$ ,  $i = 1, \dots, n$  unabhängig und identisch verteilte Zufallsvariablen mit Erwartungswert  $\mu_X$  und Varianz  $\sigma_X < \infty$ . Für  $n \rightarrow \infty$  gelten dann folgende Konvergenzen in Verteilung:

$$\sum_i X_i \sim N(n\mu_X, n\sigma_X^2) \quad \text{bzw.} \quad \bar{X} = \frac{1}{n} \sum_i X_i \sim N\left(\mu_X, \frac{\sigma_X^2}{n}\right)$$

Auch nicht normalverteilte Renditen einer Zeitreihe können, wenn sie auf ein größeres Zeitintervall aggregiert werden, in ausreichender Näherung einer Normalverteilung entsprechen.

- Die Normalverteilung ist sehr gut erforscht und in allen gängigen Softwarepaketen ohne Abweichungen in der Definition mit vielen Zusatzfunktionen und Testverfahren implementiert. Viele vor- und nachgelagerte Forschungsansätze verwenden die Normalverteilung oder eine ihrer Verallgemeinerungen.

Aus empirischer Sicht wird eingewendet, dass die Normalverteilung für die Verteilung von Wertpapierrenditen - wie für viele andere reale Vorgänge auch - nur eine unzureichende Näherung darstelle. Insbesondere modelliere die Normalverteilung bei gegebener Varianz der Verteilung deutlich zu kleine Wahrscheinlichkeiten für extreme Ereignisse. Gemessen wird dieser Sachverhalt mittels der Kurtosis einer Verteilung, die für Renditeverteilungen in der Regel höher als drei ist, dem Wert der Kurtosis für die Normalverteilung. Zudem werde empirisch eine mehr oder weniger stark ausgeprägte Linksschiefe von Renditeverteilungen festgestellt, während die Normalverteilung symmetrisch um dem Erwartungswert  $\mu$  ist. Inwieweit diese Sachverhalte auch für im Rahmen der Performance-messung verwendete Renditedaten zutreffen, wird im Abschnitt 2.3 (Empirische Renditen) untersucht.

Eine Verallgemeinerung, die bei einer weiterhin symmetrischen Verteilung eine höhere Kurtosis erlaubt, ist die **Studentsche t-Verteilung** mit  $\nu \in \mathbb{N}$  Freiheitsgraden. Sie ist die Verteilung des Quotients einer standardnormalverteilten Zufallsvariablen und des Mittelwertes  $\nu$  quadrierter standardnormalverteilter Zufallsvariablen, wobei alle Zufallsvariablen voneinander unabhängig sind. Für  $\nu \rightarrow \infty$  konvergiert die  $t$ -Verteilung gegen die Standardnormalverteilung; mit sinkendem  $\nu$  liegt ein immer größeres Gewicht der Verteilung auf den Flanken der Verteilung. Es sind jeweils nur die  $(\nu-1)$ -ten Momente einer  $t$ -Verteilung definiert. Die Kurtosis der  $t$ -Verteilung beträgt  $3+6/(\nu-4)$ , wobei die Kurtosis als viertes Moment nur für  $t$ -Verteilungen mit  $\nu \geq 5$  definiert ist. Die  $t$ -Verteilung ist symmetrisch, so dass die Schiefe null ist. Bei der Diskussion über die  $t$ -Verteilung wird meistens die jeweilige Standard- $t$ -Verteilung mit  $\nu$  Freiheitsgraden behandelt. Ihr Erwartungswert ist 0 und ihre Varianz beträgt  $\sigma^2 = \nu/(\nu-2)$ . Eine  $t$ -verteilte Zufallsvariable  $X$  mit Erwartungswert  $\mu_X$  und Varianz  $\sigma_X^2$  erhält man über die Lage-Skalen-Beziehung  $X = \mu_X + \sigma_X \sqrt{(\nu-2)/\nu} \cdot t_\nu$ , wobei  $t_\nu$

eine Standard- $t$ -verteilte Zufallsvariable ist.<sup>3</sup>

Ein weiteres einfaches Modell zur Abbildung von Wertpapierrenditen ist die Mischung normalverteilter Zufallsvariablen. Eine Mischungsverteilung kann man sich als zweistufige Verteilung vorstellen, bei der über eine Gewichtungsverteilung  $G$  die Verteilungsfunktion  $F(x|G)$  zufällig ausgewählt wird, aus der anschließend  $X$  gezogen wird. Bei einer diskreten Mischung aus  $K \in \mathbb{N}$  verschiedenen Verteilungen ist  $G$  eine diskrete Verteilung, die  $K$  Werte annehmen kann, jeweils mit Wahrscheinlichkeit  $\pi_k$ ,  $k = 1, \dots, K$ . Wird ein bestimmtes  $k$  gezogen, folgt  $X$  der Verteilung  $F_k(x)$ . Möglicherweise hängen die  $F_k$  von einem Parametervektor  $\theta$  ab. Die Dichtefunktion einer diskreten Mischungsverteilung kann man schreiben als<sup>4</sup>

$$f(x; \theta, \pi) = \sum_{k=1}^K \pi_k f_k(x; \theta) \quad \text{mit} \quad \sum_{k=1}^K \pi_k = 1.$$

Ist  $G$  eine stetige Verteilung, so wird über den Mischungsparameter  $\pi$  integriert<sup>5</sup>:

$$f(x; \theta) = \int g(\pi) h(x; \pi, \theta) d\pi \quad \text{mit} \quad \int g(\pi) d\pi = 1$$

Mischt man  $K$  unterschiedliche  $N(\mu_0, \sigma_k^2)$ -verteilte Zufallsvariablen mit Gewichten  $\pi_k > 0$ , so dass  $\pi_1 + \dots + \pi_K = 1$ , dann hat die resultierende Mischungsverteilung eine höhere Kurtosis als die Normalverteilung. Lässt man zudem unterschiedliche Mittelwerte  $\mu_k$  zu, sind auch schiefe und platykurtische Mischungsverteilungen möglich.

Ein Problem dieser Ansätze ist, dass sie die Reproduktionseigenschaft der Normalverteilung nicht teilen: Summen unabhängiger identisch nicht-normalverteilter Zufallsvariablen liegen im Allgemeinen nicht in der Lage-Skalen-Familie der Ausgangsverteilung. Hat eine Verteilung die Eigenschaft, dass Summen unabhängig (identisch) verteilter Zufallsvariablen eine Verteilung aus derselben Lage-Skalen-Familie wie die ursprüngliche Verteilung haben, so handelt es sich um eine (faltungs-)stabile Verteilung.

Eine Verteilungsklasse, bei der die Stabilitätseigenschaft erhalten bleibt, ist die Klasse der (Alpha-)stabilen Verteilungen oder Stable Paretian Distributions. Sie können durch Variation ihres Parameters  $\alpha$  ( $0 < \alpha \leq 2$ ) deutlich mehr Wahrscheinlichkeitsmasse auf die Flanken der Verteilung legen. Für  $\alpha = 2$  erhält man die Normalverteilung. Zusätzlich ist durch Variation eines zweiten Parameters  $\beta$  die Schiefe der Verteilung innerhalb gewisser Grenzen variierbar. Diese Verteilungen haben jedoch einige deutliche Nachteile: Erstens hat die Dichtefunktion bis auf wenige Ausnahmen keine geschlossene Form.<sup>6</sup> Die Verteilung

<sup>3</sup>Über eine Generalisierung der  $t$ -Verteilung zur nicht-zentralen  $t$ -Verteilung lässt sich zusätzlich noch eine Schiefe ungleich null modellieren.

<sup>4</sup>analog zu McLachlan und Basford (1988, S. 9)

<sup>5</sup>analog zu Everitt (1982, S. 560)

<sup>6</sup>Neben der Normalverteilung für  $\alpha = 2$  erhält man im Fall ( $\alpha = 1, \beta = 0$ ) die Cauchy-Verteilung und im Fall ( $\alpha = 0.5, \beta = 1$ ) die Lévy-Verteilung.

kann lediglich über ihre charakteristische Funktion  $E(e^{iXt})$  beschrieben werden. Diese Beschreibung ist zwar vollständig und eindeutig, für Anwendungen muss die Dichtefunktion jedoch jeweils numerisch bestimmt werden. Ebenso wird kritisiert, dass für viele Anwendungen die Wahrscheinlichkeitsmasse auf den Flanken von nicht-normalen stabilen Verteilungen wiederum zu hoch sei.<sup>7</sup> Schwerer wiegt im Rahmen der Performanceanalyse der Nachteil, dass für stabile Verteilungen in den meisten Fällen nur wenige bis gar keine Momente existieren. Nur im Fall  $\alpha = 2$  (Normalverteilung) existieren alle Momente und sind endlich. Für  $1 < \alpha < 2$  existiert nur der Erwartungswert; für  $\alpha \leq 1$  existieren gar keine Momente. Da die Performancemessung in den meisten Fällen über Momente von Verteilungen (bzw. Funktionen dieser Momente) erfolgt, sind Modelle über nicht-normale stabile Verteilungen nur für sehr wenige Performancemaße geeignet; für die wichtigste Anwendung in dieser Arbeit, die Ermittlung der Varianzen von Performancemaß-Schätzern, werden stets mindestens die ersten beiden Momente einer Verteilung benötigt. Deshalb werden im Folgenden keine auf stabilen Verteilungen beruhenden Modelle verwendet. Kurze Einführungen in univariate Renditeverteilungsmodelle findet man z.B. in Rachev et al. (2005) und Schmid und Trede (2006), ausführlicher in Rachev und Mittnik (2000).

## 2.2.2 Intertemporal abhängige Renditen

Im Folgenden wird die Annahme der intertemporalen Unabhängigkeit der Renditen aufgegeben.

### 2.2.2.1 Autokorrelation

Die Annahme unabhängig und identisch verteilter Renditen impliziert die Unkorreliertheit aller Renditen eines Wertpapiers, das heißt, dass alle Korrelationskoeffizienten  $\rho_{R_t, R_{t+s}}$  für alle  $s \in \{\mathbb{Z} \setminus 0\}$  gleich null sind. Gibt man diese Annahme auf, kann intertemporale Korrelation von Renditen z. B. mittels eines ARMA(p,q)-Modells

$$R_t = \mu + \alpha_1 R_{t-1} + \dots + \alpha_p R_{t-p} + \varepsilon_t + \beta_1 \varepsilon_{t-1} + \dots + \beta_q \varepsilon_{t-q}$$

modelliert werden, wobei die  $\varepsilon_i$  unabhängig und identisch verteilte Störterme oder Schocks mit  $E(\varepsilon) = 0$  sind. Die Rendite in  $t$  hängt also von den Störtermen aus den zurückliegenden Perioden ab. Über den Parameter  $\mu$  wird der unbedingte Erwartungswert der Renditen gesteuert: Der Erwartungswert eines

<sup>7</sup>Ein Problem liegt hier darin, dass sich zwar die Nullhypothese schwacher Flanken bei moderaten Stichprobengrößen mit relativ hoher Entscheidungskraft statistisch testen lässt, die Normalverteilungshypothese etwa über den Jarque-Bera-Test. Umgekehrt ist dies wesentlich schwerer. Geht man von einem stabilen Verteilungsmodell aus, werden für den Flankenparameter  $\alpha$  bei normalverteilten Stichproben ( $\hat{=}$  stabil verteilt mit  $\alpha = 2$ ) häufig Werte deutlich unter 2 geschätzt. Auch die Unterscheidung stabiler Verteilungen von nicht stabilen Verteilungen mit endlicher Varianz (z.B. Laplace-Verteilung) ist bei moderaten Stichprobengrößen ( $n < 1000$ ) schwierig, siehe z.B. DuMouchel (1983).

stationären ARMA-Prozesses beträgt

$$E(R_t) = \frac{\mu}{1 - \sum_{k=1}^p \alpha_k}.$$

Die Varianzen  $\sigma_{R_t}^2$  und Kovarianzen  $\sigma_{R_t, R_{t-s}}$  der Renditen können mittels der Yule-Walker-Gleichungen<sup>8</sup> ermittelt werden. Für einen ARMA(1,0)-Prozess, kurz: AR(1)-Prozess, erhält man die Formeln

$$E(R_t) = \frac{\mu}{1 - \alpha_1}, \quad \sigma_{R_t}^2 = \frac{\sigma_\varepsilon^2}{1 - \alpha_1^2} \quad \text{und} \quad \sigma_{R_t, R_{t-s}} = \frac{\alpha_1^{|t-s|} \sigma_\varepsilon^2}{1 - \alpha_1^2}.$$

Es sei an dieser Stelle kurz auf die Bedingungen eingegangen, unter denen ein ARMA-Prozess (schwach) stationär ist, d.h. dass der unbedingte Mittelwert  $\mu_{R_t}$  der Zeitreihe für alle  $t$  konstant ist und die Varianz sowie die Kovarianzen  $\sigma_{R_s, R_t}$  endlich sind und nur vom Abstand  $t - s$  abhängen. Die Eigenschaft der Stationarität sichert auch, dass der Prozess nicht irreversibel abdriftet („ex- oder implodiert“).

Die Stationarität eines ARMA(p,q)-Modells hängt nur von den autoregressiven AR-Termen  $\alpha_1, \dots, \alpha_p$  ab. Aus dem AR-Teil des Prozesses

$$R_t = \mu + \alpha_1 R_{t-1} + \dots + \alpha_p R_{t-p}$$

lässt sich das sogenannte „charakteristische Polynom“

$$P(z) = 1 - \alpha_1 z - \dots - \alpha_p z^p$$

bilden. Der ARMA-Prozess ist stationär, wenn alle (möglicherweise auch komplexen) Nullstellen von  $P(z)$  einen Betrag größer 1 haben. Für den häufig verwendeten AR(1)-Prozess (bzw. den ARMA(1,q)-Prozess) bedeutet das, dass Stationarität garantiert ist, solange  $|\alpha_1| < 1$  ist.<sup>9</sup> Solange wir uns auf Portfoliorenditen beschränken (und keine Portfoliowerte oder Wertpapierkurse einbeziehen), haben wir es bis auf einige Ausnahmen<sup>10</sup> mit erkennbar stationären Zeitreihen zu tun. Statistisch getestet wird Stationarität über Einheitswurzeltests.<sup>11</sup>

Sind die Störterme  $\varepsilon_t$  eines stationären ARMA-Prozesses normalverteilt, so sind auch die Realisationen  $R_t$  des Prozesses normalverteilt, da sie Linearkombinationen der unabhängig und identisch normalverteilten Störterme  $\varepsilon_t, \varepsilon_{t-1}, \dots$  und der Konstanten  $\mu$  sind.

In ökonomischer Hinsicht muss die Bedingung der intertemporalen Unkorreliertheit der Renditen auf vollkommenen Kapitalmärkten erfüllt sein, da sich ansonsten Arbitragemöglichkeiten, etwa über Termingeschäfte, bieten würden.

<sup>8</sup>siehe z.B. Cryer und Chan (2008, S. 76)

<sup>9</sup>Zur Stationarität von ARMA-Prozessen siehe ausführlicher z.B. Greene (2008, S. 718 ff.)

<sup>10</sup>Z. B. Geldmarktanlagen, deren Renditen der Höhe des jeweiligen kurzfristigen Zinssatzes folgen oder Wertpapieren von Unternehmen in Ausnahmesituationen (z.B. Insolvenz), bei denen kein liquider Markt mehr existiert.

<sup>11</sup>siehe z.B. Greene (2008, S. 744 - 756), Cryer und Chan (2008, S. 128 ff.)

Durch Ausnutzung dieser Arbitragemöglichkeiten und dessen Auswirkung auf Angebot und Nachfrage würden die Wertpapierkurse so angepasst, dass die Autokorrelation wieder verschwindet. An sich sollte man bei Wertpapierrenditen also keine Autokorrelation beobachten können. Sie findet sich aber mitunter bei nicht börsengehandelten Vermögenswerten, zum Beispiel bei veröffentlichten (Hedge-)Fondsrenditen, sowie in Hochfrequenzdaten, also Renditen über sehr kurze Zeiträume (Minuten, Sekunden(-bruchteile) oder die Zeitspanne zwischen zwei Kursfeststellungen).

### 2.2.2.2 Abhängige Varianzen

Es ist möglich, dass zwar die Renditen selber zeitlich unkorreliert sind, deren Potenzen jedoch nicht. Ein gut modellierbarer Fall in dieser Gruppe betrifft Abhängigkeiten der quadrierten Renditen entsprechend der Beobachtung, dass die Schwankung von Finanzmarktzeitreihen (Volatilität) auf mittelfristige Sicht nicht konstant ist, d.h. dass es durchaus länger andauernde Perioden mit höherer bzw. niedrigerer Volatilität gibt. Der wichtigste (diskrete) Ansatz auf diesem Gebiet ist das GARCH-Modell (Generalized Autoregressive Conditional Heteroscedasticity) von Bollerslev, eine Verallgemeinerung des von Engle<sup>12</sup> entwickelten ARCH-Modells. Die Definition eines GARCH(p,q)-Prozesses  $\epsilon_t$  lautet nach Bollerslev (1986, dortige Notation):

$$\epsilon_t | \psi_{t-1} \sim N(0, h_t),$$

mit

$$\begin{aligned} \psi_{t-1} &\hat{=} \text{Menge aller verfügbaren Informationen bis zum Zeitpunkt } t-1 \\ h_t &= \alpha_0 + \sum_{i=1}^q \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^p \beta_j h_{t-j}; \end{aligned}$$

sowie  $p \in \mathbb{N}_0$ ,  $q \in \mathbb{N}$ ,  $\alpha_0 > 0$ ,  $\alpha_i \geq 0$  ( $i = 1, \dots, q$ ),  $\beta_j \geq 0$  ( $j = 1, \dots, p$ ).<sup>13</sup> In der Praxis<sup>14</sup> – und auch in dieser Arbeit – beschränkt man sich häufig auf den GARCH(1,1)-Fall

$$h_t = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 h_{t-1}.$$

In diesem Fall ist die Varianz in  $t$  bedingt auf die Varianz in  $t-1$  sowie den konkreten Wert des Störterms in  $t-1$ . Dabei wirkt das Quadrat des Störterms des Prozesses in  $t-1$  zunächst über den MA-Term in der Varianzformel direkt auf die bedingte Varianz der Folgeperiode  $h_t$  ein und wird danach über den AR-Term (mit exponentiell abnehmender Wirkung) in die weiteren Folgeperioden übertragen. Damit wird der empirisch gut beobachtbare Fall modelliert, dass

<sup>12</sup>Engle (1982)

<sup>13</sup>Die Störterme müssen nicht notwendigerweise normalverteilt sein. Man kann die Annahme auf  $\epsilon_t | \psi_{t-1} \sim h_t \cdot \text{White Noise Typ 1}(\mu = 0, \sigma = 1)$  verallgemeinern. Die folgenden Ausführungen beziehen sich lediglich auf GARCH-Modelle mit normalverteilten Störtermen.

<sup>14</sup>vgl. Mills und Markellos (2008, S. 177)



ein starker Schock durch eine den Markt erreichende und den Kurs stark beeinflussende Nachricht auch zu höheren Kursschwankungen in den Folgeperioden führt. Erfolgt kein weiterer Schock, beruhigt sich der Markt allmählich wieder. Eine ökonomische Erklärung hierfür ist, dass vor Eintreffen des Schocks ein gewisses Marktgleichgewicht mit einigermaßen stabilen Preiserwartungen hinsichtlich des betroffenen Vermögenswertes besteht. Durch den Schock wird den Marktteilnehmern klar, dass sich der Wert des Vermögenswertes nunmehr geändert hat. Es dauert aber einige Zeit, bis sich am Markt ein impliziter Konsens herausgebildet hat, wie hoch der neue Wert des Vermögenswertes ist.

Ein GARCH(1,1)-Prozess lässt sich als ARMA(1,1)-Prozess in den quadrierten Renditen auffassen.<sup>15</sup> Für eine Stationarität des Prozesses muss die Bedingung

$$\sum_{i=1}^q \alpha_i + \sum_{j=1}^p \beta_j < 1$$

erfüllt sein. In diesem Fall beträgt die unbedingte Varianz des Prozesses

$$\sigma_{\epsilon_t}^2 = \frac{\alpha_0}{1 - \sum_{i=1}^q \alpha_i + \sum_{j=1}^p \beta_j}.$$

Bei einem GARCH(1,1)-Prozess reduzieren sich die Formeln auf  $\alpha_1 + \beta_1 < 1$  und  $\sigma_{\epsilon_t}^2 = \frac{\alpha_0}{1 - \alpha_1 - \beta_1}$ .

Es ist möglich, ARMA- und GARCH-Prozesse miteinander zu verbinden, etwa indem ein ARMA-Modell keine unabhängig und identisch verteilten Störterme hat, sondern Störterme, die einem GARCH-Prozess folgen<sup>16</sup>. Im weiteren Verlauf wird lediglich ein Modell „GARCH(1,1) + Konstante“

$$R_t = \gamma + h_t \epsilon_t \quad \text{mit} \quad h_t = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 h_{t-1} \quad \text{und} \quad \epsilon \sim N(0, \sigma^2)$$

verwendet, wobei die Konstante  $\gamma$  die angenommene langfristige (Über-)Rendite eines Portfolios darstellt.

GARCH- bzw. ARMA-GARCH-Modelle führen zu unbedingten Verteilungen, die als Mischungen von Normalverteilungen die gegenüber der Normalverteilung höhere Kurtosis empirischer Renditeverteilungen nachbilden. Eine praktische Schwierigkeit ist, dass sich die unbedingten Verteilungen nicht in geschlossener Form ausdrücken lassen. Viele Eigenschaften, z.B. Quantile, können nur numerisch bzw. simulativ ermittelt werden. Andere Größen wie die (unbedingte) Kurtosis sind nur für bestimmte Parameterkonstellationen definiert. Die unbedingte Kurtosis für einen GARCH(1,1)-Prozess mit normalverteilten Störtermen beträgt z.B.:<sup>17</sup>

$$\gamma_{2,u} = 3 \cdot \frac{1 - \beta_1^2 - 2\alpha_1\beta_1 - \alpha_1^2}{1 - \beta_1^2 - 2\alpha_1\beta_1 - 3\alpha_1^2},$$

<sup>15</sup>Schmid und Trede (2006, S. 179)

<sup>16</sup>siehe hierzu z.B. Gouriéroux und Jasiak (2001, S. 128 f.)

<sup>17</sup>Bollerslev (1986), zitiert nach Jondeau et al. (2007, S. 146 f.)

was nur definiert ist, wenn der Nenner positiv, also  $\beta_1^2 + 2\alpha_1\beta_1 + 3\alpha_1^2 < 1$  erfüllt ist. Wenn beispielsweise  $\alpha_1 \geq \sqrt{1/2} \approx 0,7071$  ist, kann das vierte Moment nicht existieren.<sup>18</sup>

Die häufig beobachtete Schiefe von Renditeverteilungen lässt sich mittels „gewöhnlichem“ GARCH nicht darstellen. GARCH unterstellt, dass positive und negative Schocks dieselbe (positive) Wirkung auf die bedingte Varianz der Folgeperioden haben, während in der Praxis vor allem nach negativen Schocks stark erhöhte Volatilitäten zu beobachten sind.<sup>19</sup> Für Diskussion und empirische Untersuchung dieses „Leverage-Effekt“ genannten Phänomens siehe Christie (1982) und French et al. (1987). Unter den vielen weiteren Fortentwicklungen des GARCH-Modells<sup>20</sup> können u.a. das Exponential GARCH (EGARCH) von Nelson (1991), das Threshold GARCH (TGARCH) von Zakoian (1994) und GJR-GARCH von Glosten et al. (1993) asymmetrische unbedingte Renditeverteilungen modellieren.

Insgesamt dominiert nach wie vor das GARCH(1,1)-Modell, auch aus praktischen Gründen: Das GARCH(1,1)-Modell ist ein Standardmodell mit relativ wenigen und gut interpretierbaren Parametern, das häufig keine schlechtere Anpassung an die Daten liefert als GARCH-Modelle höherer Ordnung. Vielfach erweist sich beim Test mehrerer Modelle keines hinsichtlich Anpassung und Prognose als eindeutig überlegen.<sup>21</sup> Wer höhere Lag-Längen nutzen will, muss eine Wahl treffen und begründen, wie hoch  $p$  und  $q$  sein sollen. Klassische Informationskriterien wie AIC und BIC sind für heteroskedastische Daten häufig wenig geeignet.<sup>22</sup> In einer Vergleichsstudie von Hansen und Lunde (2005) mit ARCH, GARCH und 14 Modifikationen von GARCH – insgesamt 330 Modellvarianten – konnten sie für Wechselkursdaten kein Modell finden, das GARCH(1,1) eindeutig übertraf. In einem Aktiendatensatz erwies sich GARCH(1,1) lediglich Modellen, die den Leverage-Effekt explizit abbildeten, unterlegen.

### 2.2.3 Multivariate Renditemodelle

Viele Performancemaße werden aus der Renditezeitreihe eines einzigen Portfolios berechnet – desjenigen, dessen Wertentwicklung quantifiziert werden soll. Andere Performancemaße setzen die Renditezeitreihe eines Portfolios zu einer anderen Renditezeitreihe (Benchmark) in Beziehung, häufig einem näher zu definierenden Marktportfolio. Da in die Berechnung derartiger Performancemaße

$$^{18} \beta_1^2 + 2\alpha_1\beta_1 + 3\alpha_1^2 < 1 \Leftrightarrow \underbrace{(\beta_1 + \alpha_1)^2}_{\geq 0} + 2\alpha_1^2 < 1 \Rightarrow 2\alpha_1^2 < 1 \Leftrightarrow |\alpha_1| < \sqrt{1/2}$$

<sup>19</sup>Auch eine umgekehrte Wirkungsbeziehung ist hier erklärbar: Steigt durch einen Schock die Unsicherheit über die Wertentwicklung einer Anlage, so nimmt nach dem  $\mu$ - $\sigma$ -Prinzip der Nutzen der Anlage für einen Investor ab. Um die Anlage ceteris paribus wieder attraktiv zu machen, müsste deren erwartete Rendite ansteigen, was kurzfristig nur über einen sinkenden Kurs der Anlage möglich ist. Dies würde die Wirkung von positiven Schocks dämpfen und jene von negativen Schocks verstärken.

<sup>20</sup>Einen Überblick über univariate GARCH-Modelle liefern Jondeau et al. (2007, Kap. 4) und Teräsvirta (2009)

<sup>21</sup>siehe z.B. Brailsford und Faff (1996, S. 436), Brooks und Burke (2003)

<sup>22</sup>siehe Brooks und Burke (2003)

bzw. deren Varianzen regelmäßig gemischte Momente wie die Kovarianz eingehen, können die Zeitreihen nicht separat modelliert werden, sondern müssen einem multivariaten Modell entstammen. Dies gilt auch dann, wenn Performancemessungen zweier Portfolios miteinander verglichen werden. Es werden also bivariate Modelle gebraucht; werden zwei Portfolios mittels eines Performancemaßes mit Benchmark – die Benchmark ist für beide Portfolios notwendigerweise dieselbe – verglichen, ist ein trivariates Modell erforderlich. Es folgt nun eine Übersicht über die in dieser Arbeit verwendeten multivariaten Renditemodelle sowie einige weitere Varianten.

### 2.2.3.1 Multivariate unabhängig und identisch verteilte Zufallsvariablen

Das Ausgangsmodell für multivariate Renditeverteilungen ist die **multivariate Normalverteilung** mit Dichtefunktion für  $d$  Dimensionen

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{(2\pi)^{\frac{d}{2}} |\boldsymbol{\Sigma}|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu})\right),$$

wobei  $\boldsymbol{\mu} \in \mathbb{R}^{1 \times d}$  der Erwartungswertvektor und  $\boldsymbol{\Sigma} \in \mathbb{R}^{d \times d}$  die Kovarianzmatrix der Verteilung ist. Die Kovarianzmatrix kann eine beliebige symmetrische, positiv definite Matrix sein.

Alle Randverteilungen einer multivariaten Normalverteilung sind ebenfalls Normalverteilungen. Auch alle auf  $j \in \{1, \dots, d-1\}$  Komponenten bedingten Verteilungen sind  $(d-j)$ -dimensionale Normalverteilungen. Die multivariate Normalverteilung hat eine Reihe von Eigenschaften, die über ihre Verwendung als Renditemodell hinaus für Anwendungen in den folgenden Kapiteln wichtig sind:

- Die multivariate Normalverteilung ist durch ihre Parameter  $\boldsymbol{\mu}$  und  $\boldsymbol{\Sigma}$ , die gleichzeitig auch Erwartungswertvektor und Kovarianzmatrix der Komponenten sind, vollständig definiert.
- Ist die Kovarianzmatrix eine Diagonalmatrix, so sind alle Komponenten stochastisch global unabhängig.
- Die Reproduktionseigenschaft der Normalverteilung bleibt erhalten: Jede lineare Funktion eines Vektors  $\mathbf{X}^{d \times 1}$  gemeinsam multivariat normalverteilter Zufallsvariablen ist ebenfalls multivariat normalverteilt<sup>23</sup>: Wenn  $\mathbf{X} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$  sowie  $\mathbf{b}^{d \times 1}$  ein reellwertiger Vektor und  $\mathbf{A}$  eine reellwertige  $d$ -spaltige Matrix sind, dann gilt  $\mathbf{A}\mathbf{X} + \mathbf{b} \sim N(\mathbf{A}\boldsymbol{\mu} + \mathbf{b}, \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}')$ .
- Der aus dem univariaten Bereich bekannte Zentrale Grenzwertsatz gilt auch im Multivariaten<sup>24</sup>:  
Seien die Zufallsvektoren  $\mathbf{Y}_1^{1 \times d}, \dots, \mathbf{Y}_n^{1 \times d}$  unabhängig und identisch verteilt mit Erwartungswertvektor  $\boldsymbol{\mu}^{1 \times d}$  und Kovarianzmatrix  $\boldsymbol{\Sigma}^{d \times d}$ . Dann konvergiert die Folge  $\sqrt{n}(\bar{\mathbf{Y}} - \boldsymbol{\mu})$  für  $n \rightarrow \infty$  in Verteilung gegen  $N(\mathbf{0}, \boldsymbol{\Sigma})$ .

<sup>23</sup>Vgl. Greene (2008, S. 1015)

<sup>24</sup>Beweis siehe z.B. Anderson (1984, S. 81)

Wie im univariaten Fall kann die multivariate Normalverteilung die Schiefe und Kurtosis empirischer Renditeverteilungen nicht abbilden. Hinzu kommt, dass die Abhängigkeitsstruktur zwischen den univariaten Verteilungen, die die multivariate Normalverteilung impliziert, sich von empirisch beobachteten Abhängigkeitsstrukturen unterscheidet. So wird oft beobachtet, dass die Abhängigkeit zwischen den Wertpapierrenditen eines Marktes dann besonders hoch ist, wenn die Rendite des Gesamtmarktes stark in eine Richtung ausschlägt; hierbei spricht man von *tail dependence*. Häufig unterscheidet man zwischen dem Verhalten bei Extremwerten an der oberen und der unteren Flanke der Renditeverteilung (*upper tail* und *lower tail*). In der multivariaten Normalverteilung ist die Korrelation zwischen den Renditen konstant.

Diese die Abhängigkeitsstruktur betreffenden Phänomene kann man mittels Copulas modellieren. Eine Copula stellt im Falle von  $d$  gleichzeitig betrachteten eindimensionalen Renditeverteilungen als Funktion  $C(u_1, \dots, u_d)$  im  $d$ -dimensionalen Einheitswürfel  $[0; 1]^d$  die Verbindung der Randverteilungen  $F_{X_1}(x_1), \dots, F_{X_d}(x_d)$  mit der gemeinsamen Verteilung  $F_{X_1, \dots, X_d}(x_1, \dots, x_d)$  her:<sup>25</sup>

$$F_{X_1, \dots, X_d}(x_1, \dots, x_d) = C(F_{X_1}(x_1), \dots, F_{X_d}(x_d))$$

Die Copula  $C$  ist also die gemeinsame Verteilung der Verteilungsfunktionen der einzelnen Komponenten.

In dieser Arbeit wird auf Copulas und ihre Theorie nicht weiter eingegangen. Eine praxisorientierte Einführung geben Frees und Valdez (1998); eine Einführung auf mathematischerem Niveau liefern Nelsen (2006) und Joe (1997), die Anwendung im Bereich Finanzmarkttheorie stellen Cherubini et al. (2004) vor. Für Copula-Anwendungen im MGARCH-Kontext siehe Manner (2010), eine kurze Übersicht zu Literaturquellen findet man in Patton (2009, S. 772f.).

Die multivariate Erweiterung der  $t$ -Verteilung ist die **multivariate  $t$ -Verteilung**. Sie wird spezifiziert über die Anzahl ihrer Freiheitsgrade und die Matrix  $\Sigma$ . Die multivariate  $t$ -Verteilung hat als Randverteilungen  $t$ -Verteilungen, die jeweils dieselbe Anzahl Freiheitsgrade  $\nu$  aufweisen müssen. Hier ist also – wie im univariaten Fall – die Modellierung einer bestimmten Kurtosis möglich, während die Schiefe aufgrund der weiterhin symmetrischen Randverteilungen gleich null bleibt. Dass die Anzahl Freiheitsgrade für alle Dimensionen gleich ist, stellt eine Einschränkung dar, da es z.B. bei der Modellierung einer Wertpapier- und einer Marktzeitreihe angemessener erscheinen könnte, unterschiedliche Freiheitsgrade anzunehmen, etwa weil ein einzelnes Wertpapier ausreißeranfälliger sein könnte als der entsprechende Marktindex.

Über die als Parameter vorgegebene Matrix  $\Sigma$  kann jede mögliche tatsächliche Kovarianzmatrix – über die Beziehung  $\text{Cov}(\mathbf{X}) = \frac{\nu}{\nu-2}\Sigma$  – direkt modelliert werden, also auch vollständige Korrelation oder völlige Unkorreliertheit. Allerdings sind die einzelnen Dimensionen der Verteilungen auch bei Unkorreliertheit nicht unabhängig.<sup>26</sup> Durch beidseitige Multiplikation von  $\Sigma$  mit einem Vektor und

<sup>25</sup>Zusammen mit der Feststellung, dass für jede multivariate Verteilungsfunktion eine Copula existiert, heißt dieser Satz *Sklar's Theorem*, nach Sklar (1959).

<sup>26</sup>vgl. McNeil et al. (2005, S. 191)

durch Addition eines Erwartungswertvektors ist die multivariate t-Verteilung beliebig skalier- und verschiebbar.<sup>27</sup> Für  $\nu \rightarrow \infty$  konvergiert die multivariate t-Verteilung gegen die multivariate Normalverteilung.

### 2.2.3.2 MGARCH

Während die Erweiterung der Normal- und t-Verteilung in den multivariaten Bereich eindeutig ist, existieren verschiedene, teilweise sehr komplexe Ansätze für multivariates GARCH, von denen sich keiner gegenüber den anderen eindeutig durchgesetzt hat. Im Folgenden werden einige etablierte Modelle in Reihenfolge absteigender Komplexität kurz vorgestellt. Die Darstellung beschränkt sich auf multivariate Versionen von GARCH(1,1)-Modellen.<sup>28</sup>

Der älteste Ansatz, das *Vech-MGARCH*, wurde in Bollerslev et al. (1988) vorgestellt. Nach diesem Modell hängen Varianzen und Kovarianzen der modellierten Zeitreihen linear von allen Varianzen und Störtermen der Vorperiode(n) ab. Zur besseren Handhabbarkeit werden die entsprechenden Matrizen mittels des *vech*-Operators (Halbvektorisierungs-Operator) in Vektoren umgewandelt.<sup>29</sup> Bei einem bivariaten Vech-GARCH(1,1) hängt die Kovarianzmatrix  $\Sigma_t$  des Wertevektors  $(r_{1,t}, r_{2,t})'$  wie folgt von den Werten der Vorperiode ab:<sup>30</sup>

$$\begin{aligned} \text{vech}(\Sigma_t) &= \text{vech} \begin{pmatrix} \sigma_{1,t}^2 & \sigma_{12,t} \\ \sigma_{12,t} & \sigma_{2,t}^2 \end{pmatrix} = \begin{pmatrix} \sigma_{1,t}^2 \\ \sigma_{12,t} \\ \sigma_{2,t}^2 \end{pmatrix} = \\ &= \begin{pmatrix} \gamma_{11} \\ \gamma_{12} \\ \gamma_{22} \end{pmatrix} + \underbrace{\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}}_{\mathbf{A}} \begin{pmatrix} \varepsilon_{1,t-1}^2 \\ \varepsilon_{1,t-1} \varepsilon_{2,t-1} \\ \varepsilon_{2,t-1}^2 \end{pmatrix} + \underbrace{\begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix}}_{\mathbf{B}} \begin{pmatrix} \sigma_{1,t-1}^2 \\ \sigma_{12,t-1} \\ \sigma_{2,t-1}^2 \end{pmatrix} \end{aligned}$$

<sup>27</sup>Die multivariate t-Verteilung wird umfassend in Kotz und Nadarajah (2004) dargestellt.

<sup>28</sup>Dieser Abschnitt fasst einen Teil der Inhalte verschiedener Übersichtsdarstellungen von MGARCH-Modellen zusammen. Diese sind: Bauwens et al. (2006), Jondeau et al. (2007, Kapitel 6.1), Lütkepohl (2005, Kapitel 16), Silvennoinen und Teräsvirta (2009).

<sup>29</sup>Bei der Vektorisierung einer Matrix  $\mathbf{A}$  werden die Spalten einer Matrix zu einem Spaltenvektor aneinandergereiht:

$$\text{vec} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = (a_{11} \ a_{21} \ a_{31} \ a_{12} \ a_{22} \ a_{32} \ a_{13} \ a_{23} \ a_{33})'$$

Ist die Matrix symmetrisch, so ist Bereich oberhalb der Diagonalen redundant zum Bereich unterhalb der Diagonalen. Bei der Halbvektorisierung (*vech*) einer Matrix wird deshalb nur die untere Dreiecksmatrix vektorisiert:

$$\text{vech} \begin{pmatrix} a_{11} & a_{21} & a_{31} \\ a_{21} & a_{22} & a_{32} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = (a_{11} \ a_{21} \ a_{31} \ a_{22} \ a_{32} \ a_{33})'$$

<sup>30</sup>Analog zu Jondeau et al. (2007, S. 198)

An diesem Modell lassen sich gut einige Probleme verdeutlichen, die die weiteren Modelle zu beheben versuchen.<sup>31</sup>

- *Parameterzahl:* Ein  $d$ -dimensionales Vech-GARCH(1,1)-Modell hat  $d(d+1)(d(d+1)+1)/2$  Parameter, so dass ein bivariates Modell 21, ein trivariates bereits 78 Parameter aufweist. Dies macht die Modellschätzung rechenaufwendig und setzt dem Austesten verschiedener Parameterkombinationen im Rahmen von Simulationsstudien Grenzen.
- *Interpretierbarkeit der Parameter:* Während die theoretische Bedeutung einzelner Parameter meist relativ klar ist, sind die Wirkungen und Wechselwirkungen einzelner Parameter auf die Korrelationsstruktur häufig schwer abzuschätzen. Es ist somit auch nicht leicht möglich, im Rahmen einer Simulation a priori Parameterkombinationen vorzugeben, die zu realistischen Zeitreihenverläufen oder zu bestimmten, interessierenden Verteilungseigenschaften führen.
- *Identifizierbarkeit:* Ähnlich wie bei Vektor-ARMA-Prozessen kann es bei MGARCH-Prozessen zu Problemen der Modellidentifikation kommen.
- *Gültige Kovarianzmatrix:* Beim Vech-Modell und einigen anderen Modellen ist nicht garantiert, dass der Prozess für alle Zeitpunkte  $t$  Kovarianzmatrizen  $\Sigma_t$  produziert, die positiv definit sind. In manchen Fällen lassen sich durch Parameterrestriktionen positiv definite Kovarianzmatrizen garantieren.
- *Momente und andere Eigenschaften der unbedingten Verteilung:* Performancemaße werden zumeist aus Momenten und anderen Funktionen der unbedingten Verteilung errechnet. Aber:
  - Es ist häufig unklar, welche unbedingten Momente des Prozesses bei einer bestimmten Parameterkombination existieren bzw. endlich sind.
  - Wenn die Momente existieren, gibt es häufig (noch) keine Formeln, sie aus den Parametern direkt zu berechnen. Da es sich um komplizierte Mischungen von Normalverteilungen oder anderer Störgrößenverteilungen handelt, bleibt für deren Ermittlung meist nur die Simulation übrig, wobei für jede Parameterkombination jeweils sehr viele Daten simuliert werden müssen. Für die Simulation von Renditen wiederum ist es wichtig, dass man bei der Generierung von Prozessen die für die Berechnung eines Performancemaßes wichtigen theoretischen Parameter exakt vorgeben kann – ansonsten sind im späteren Simulationsergebnis Verzerrung und Schätzvarianz einerseits und der Fehler aus der vorhergehenden Simulation nicht voneinander unterscheidbar.

---

<sup>31</sup>In dieser Arbeit wird multivariates GARCH ausschließlich zur Simulation von Zeitreihen benutzt, so dass besonders auf damit zusammenhängende Probleme hingewiesen wird, während andere wichtige Problemfelder wie die Parameterschätzung nicht angesprochen werden.

- Bei einigen Modellen sind bestimmte Werte unbedingter Momente nicht modellierbar. Will man zum Beispiel eine bivariate Zeitreihe aus einer Marktrendite und einer damit nahezu perfekt korrelierten Portfoliorendite (z.B. eines breit gestreuten Fonds auf demselben Markt) modellieren, so ist das in manchen Modellansätzen nicht möglich. Man kann zumeist erzwingen, dass sich die modellierten Renditen zu nahezu allen Zeitpunkten  $t$  in die selbe Richtung bewegen. Sind jedoch die zeitabhängigen unbedingten Varianzen der Zeitreihen nicht ebenso fast vollständig korreliert, dann wird die unbedingte Korrelation auch nicht nahezu perfekt (z.B.  $\rho_{MX} = 0,99$ ) sein.

Zur Lösung dieser Probleme wurden eine Reihe von Modellen vorgeschlagen: Bollerslev, Engle und Wooldridge regen bereits in ihrem oben genannten Artikel zum Vech-GARCH-Modell an, die Parametermatrizen  $\mathbf{A}$  und  $\mathbf{B}$  auf Diagonalmatrizen zu beschränken (DVEC). Dadurch sinkt die Anzahl zu schätzender Parameter um die Elemente außerhalb der Hauptdiagonalen von  $\mathbf{A}$  und  $\mathbf{B}$  und es lassen sich Bedingungen herleiten, unter der die positive Definitheit der Kovarianzmatrix garantiert ist. Ein weiterer Spezialfall des Vech-Modells ist das in Engle und Kroner (1995) vorgestellte BEKK-Modell<sup>32</sup>, das nochmals restriktiver als das DVEC-Modell ist, aber stets auf positiv definite Kovarianzmatrizen führt.

Ein anderer Literaturstrang folgt der Idee, kein genuin multivariates GARCH zu erzeugen, sondern univariate GARCH-Prozesse geeignet miteinander zu verknüpfen. Sehr einfach in der Handhabung ist Bollerslevs (1990) *CCC-GARCH*, wobei CCC für *constant conditional correlation* steht. Hierbei wird davon ausgegangen, dass die bedingte Korrelation zwischen den Zeitreihen konstant ist; anschließend wird die Korrelationsmatrix durch Multiplikation mit den zeitabhängigen Varianzen (bzw. Standardabweichungen) zu einer zeitabhängigen Kovarianzmatrix. Die Varianzen der einzelnen Zeitreihen entstammen jeweils univariaten GARCH-Prozessen. Dieser Ansatz benötigt relativ wenige, gut interpretierbare Parameter, garantiert unter einfachen Bedingungen positiv definite Kovarianzmatrizen und ist leicht zu schätzen. Problematisch für Simulationen ist, dass die unbedingten Korrelationen zwischen den Zeitreihen schwer zu berechnen sind. Weil die Varianzen der einzelnen Zeitreihen unkorreliert sind, sind sehr stark korrelierte Zeitreihen mit diesem Modell nicht darstellbar.

Eine weitere Modellklasse ist diejenige der Faktormodelle, die vor allem für die Modellierung hochdimensionaler Zeitreihen gut geeignet ist. Dabei wird die Varianz-Kovarianz-Struktur von  $d$  Zufallsvariablen auf  $k < d$  (möglicherweise unbeobachtbare) Faktoren zurückgeführt. Diebold und Nerlove (1989) veröffentlichten ein bedingt heteroskedastisches Faktormodell, welches nicht als GARCH-Modell, sondern eher in den Bereich der *Stochastic Volatility*-Modelle eingeordnet wird.<sup>33</sup> Die ersten Faktor-GARCH-Modelle gehen auf Engle, Ng und Rothschild (Engle (1987), Engle et al. (1990)) zurück. Sentana (1998) fasst

<sup>32</sup>Abkürzung der Nachnamen Baba, Engle, Kraft und Kroner

<sup>33</sup>siehe Silvennoinen und Teräsvirta (2009, S. 208)

die Modelle von Engle<sup>34</sup> und Diebold/Nerlove zu einem allgemeinen additiven Ansatz zusammen.

Die Dimensionsreduktion durch den Einsatz gemeinsamer Faktoren reduziert die Parameterzahl deutlich. Dennoch sind die Parameter in den vorgenannten Modellen nicht immer leicht zu interpretieren. Wichtig beim Einsatz in Simulationen ist zudem die exakte Modellierbarkeit der unbedingten Korrelationsmatrix der Zufallsvariablen. Zur Simulation wurde deshalb folgendes, einfaches Einfaktormodell eingesetzt:

$$\mathbf{X}_t = \mathbf{Z}_t \cdot \sqrt{h_t} + \boldsymbol{\mu}, \quad (2.1)$$

wobei

$$\begin{aligned} \mathbf{Z}_t &\stackrel{iid}{\sim} D_1(\mathbf{0}, \boldsymbol{\Sigma}), \\ h_t &= \alpha_0 + \alpha_1 m_{t-1}^2 + \beta_1 h_{t-1} \\ m_t &= \epsilon_t \sqrt{h_t} \\ \epsilon_t &\stackrel{iid}{\sim} D_2(0, 1) \end{aligned}$$

und  $\boldsymbol{\mu} \in \mathbb{R}^d$  ein konstanter Vektor.

Der Vektor  $\mathbf{Z}_t \in \mathbb{R}^d$  sei für alle  $t$  unabhängig von  $\epsilon_t$ .  $D_1$  ist eine multivariate Verteilung mit Erwartungswertvektor  $\mathbf{0}$  und Kovarianzmatrix  $\boldsymbol{\Sigma}$ ,  $D_2$  eine univariate Verteilung mit Erwartungswert 0 und Varianz 1.  $\alpha_0$ ,  $\alpha_1$  und  $\beta_1$  sind Parameter des Prozesses.

Die multivariate Rendite  $\mathbf{X}_t \in \mathbb{R}^d$  ist also das Produkt aus einer zeitlich unabhängig und identisch verteilten multivariaten Zufallsvariable und einem zeitabhängigen latenten Skalar, der die Standardabweichung eines GARCH-Prozesses darstellt. Die genaue Form der Verteilungen von  $\mathbf{Z}_t$  und  $\epsilon_t$  ist nicht vorgegeben. Die Konstante  $\alpha_0$  des GARCH-Varianzterms wird auf  $1 - \alpha_1 - \beta_1$  festgesetzt, damit  $h_t$  einen unbedingten Erwartungswert von 1 hat.<sup>35</sup>

Der GARCH(1,1)-Prozess  $m_t$  selbst wird nicht weiter verwendet, lediglich dessen zeitabhängige Varianz  $h_t$  wirkt auf die multivariate Zeitreihe ein. Durch die Normierung von  $E(h_t)$  auf 1 ist die unbedingte Kovarianzmatrix des Zielprozesses  $X_t$  die Kovarianzmatrix des i.i.d.-Prozesses  $Z_t$ :

LEMMA: Für die spezielle Wahl  $\alpha_0 = 1 - \alpha_1 - \beta_1$  in Modell (2.1) gilt:

$$\text{Cov}(\mathbf{X}) = \boldsymbol{\Sigma}$$

BEWEIS: Es soll bewiesen werden, dass  $\text{Cov}(\mathbf{X}) = \text{Cov}(\mathbf{Z})$ .

$E(\mathbf{Z}) = \mathbf{0}$ . Die Zufallsvariable  $H$  sei von  $\mathbf{Z}$  unabhängig und  $E(H)=1$ .  $Z_1, \dots, Z_n$  seien die Komponenten von  $\mathbf{Z}$ .

<sup>34</sup>Sentana bezieht sich bei der Definition seines Modells ausdrücklich auf Engles Modell aus dessen unveröffentlichtem Manuskript Engle (1987) und erst im weiteren Verlauf auf die Veröffentlichung Engle et al. (1990).

<sup>35</sup> $\sigma_{m_t}^2 = E(h_t) = \frac{\alpha_0}{1 - \alpha_1 - \beta_1}$  wird 1, wenn  $\alpha_0 = 1 - \alpha_1 - \beta_1$ .



$$\text{Cov}(\mathbf{X}) = \text{Cov}(\mathbf{Z} \cdot H + \boldsymbol{\mu}) = \text{Cov}(\mathbf{Z} \cdot H) = \begin{pmatrix} \sigma_{Z_1 H, Z_1 H} & \cdots & \sigma_{Z_1 H, Z_n H} \\ \vdots & \ddots & \vdots \\ \sigma_{Z_n H, Z_1 H} & \cdots & \sigma_{Z_n H, Z_n H} \end{pmatrix}$$

Es muss also gezeigt werden, dass  $\sigma_{Z_i H, Z_j H} = \sigma_{Z_i, Z_j} \forall i, j \in 1, \dots, n$ . Bohrnstedt und Goldberger (1969) liefern eine Formel für die Kovarianz von Produkten von Zufallsvariablen  $\text{Cov}(XY, UV)$ . Für den Spezialfall  $Y = V$ , in unserer Notation  $\sigma_{Z_i H, Z_j H}$ , ergibt sich:

$$\begin{aligned} \sigma_{Z_i H, Z_j H} &= \mu_{Z_i} \mu_{Z_j} \sigma_H^2 + \mu_{Z_i} \mu_H \sigma_{Z_i, H} + \mu_{Z_j} \mu_H \sigma_{Z_j, H} + (\mu_H)^2 \sigma_{Z_i, Z_j} \\ &\quad + E((Z_i - \mu_{Z_i})(Z_j - \mu_{Z_j})(H - \mu_H)^2) \\ &\quad + \mu_{Z_i} E((Z_j - \mu_{Z_j})(H - \mu_H)^2) + \mu_{Z_j} E((Z_i - \mu_{Z_i})(H - \mu_H)^2) \\ &\quad + 2\mu_H E((Z_i - \mu_{Z_i})(Z_j - \mu_{Z_j})(H - \mu_H)) - \sigma_{Z_i, H} \sigma_{Z_j, H} \end{aligned}$$

Eingesetzt  $\mu_{Z_i} = \mu_{Z_j} = 0$ ,  $\mu_H = 1$  und  $\sigma_{Z_i, H} = \sigma_{Z_j, H} = 0$  verbleibt:

$$\sigma_{Z_i H, Z_j H} = \sigma_{Z_i, Z_j} + E((Z_i Z_j (H - 1)^2) + 2 E((Z_i Z_j (H - 1)))$$

Da Funktionen unabhängiger Zufallsvariablen unabhängig sind und die letzten beiden Terme den Kovarianzen  $\sigma_{Z_i Z_j, H^2}$  bzw.  $\sigma_{Z_i Z_j, H}$  entsprechen, sind diese beiden Terme null. □

Wir haben also ein multivariates Renditemodell, dessen Komponenten GARCH-Varianzen besitzen und dessen Kovarianzmatrix beliebig bestimmbar ist. Streng genommen handelt es sich nicht um ein MGARCH-Modell, da die Volatilität zum Zeitpunkt  $t$  nicht von den (quadratischen) Renditen der Vorperiode abhängt, sondern von einer externen latenten Variablen vorgegeben wird. Demnach wäre es eher in die Klasse der Modelle stochastischer Volatilität einzuordnen. Ökonomisch gesehen werden hier  $n$  Portfoliorenditen mit fest definierten unbedingten Momenten modelliert. Die bedingten Mittelwerte und Korrelationen sind gleich den unbedingten. Die Varianzen der einzelnen Portfolios sind stochastisch, aber perfekt korreliert. Es wird implizit angenommen, dass die zeitliche Dynamik der Volatilität durch einen „Marktprozess“ vorgegeben wird. Erhöht sich die Volatilität am Markt, etwa durch makroökonomische Schocks, dann erhöhen sich die Volatilitäten aller betrachteten Portfolios in gleichem Maße. Das ist durchaus plausibel, wenn man etwa breit diversifizierte Investmentfonds innerhalb desselben Anlagesegments betrachtet: Spezifische Schocks einzelner Wertpapiere (z.B. Bilanzveröffentlichungen, Übernahmeangebote) sind fondsintern weitgehend wegdiversifiziert und wirken sich kaum auf die Fondsvolatilität in den Folgeperioden aus, die vor allem durch globale oder anlagesegmentspezifische Ereignisse bestimmt wird. Betrachtet man hingegen einzelne Aktien oder Portfolios aus verschiedenen Anlagegebieten, dann wird man deutliche Abweichungen von der Annahme proportionaler Varianzen feststellen können.

## 2.3 Empirische Renditen

Nach dem vorhergehenden Abschnitt über die Modellierung von Renditen werden nun die Eigenschaften von auf Finanzmärkten beobachteten Renditen untersucht. Im ersten Unterabschnitt werden Renditen von Wertpapieren behandelt, deren Preisfestsetzung in erster Linie über den laufenden Börsenhandel geschieht. Performancemessung wird jedoch vor allem auf die Renditen verwalteter (Wertpapier-) Portfolios angewendet, um die Qualität der Vermögensverwaltung zu messen. In den folgenden beiden Unterabschnitten werden deswegen mit (offenen) Investmentfonds und Hedgefonds zwei wichtige Formen verwalteter Finanzvermögen genauer betrachtet.

### 2.3.1 Aktien und andere börsengehandelte Wertpapiere

Häufigstes Forschungsobjekt, wenn man von Wertpapierzeitreihen spricht, sind Zeitreihen von Aktien. Viele Studien, die angeben, sich mit „financial data“, „financial time series“ oder „asset returns“ zu beschäftigen, behandeln hauptsächlich oder ausschließlich Aktienzeitreihen. Es folgt nun ein Überblick über einige stilisierte Fakten von Finanzmarktrenditen, basierend auf den Arbeiten von Pagan (1996), Cont (2001), Rydberg (2000), Schmid und Trede (2006), Sewell (2011a) und Taylor (2005, S. 51-96)<sup>36</sup>.

#### 2.3.1.1 Stilisierte Fakten

Den Begriff „Stilisierte Fakten“ kann man in diesem Kontext definieren als *„empirical findings that are so consistent across markets that they are accepted as truth.“*<sup>37</sup>

*Leptokurtosis:* Eine der am längsten bekannten<sup>38</sup> Beobachtungen bei Wertpapierrenditen ist die oben bereits angesprochene höhere Kurtosis empirischer Renditeverteilungen gegenüber der lange als Standard benutzten Normalverteilungsannahme. Sie äußert sich darin, dass extreme positive oder negative Renditen häufiger vorkommen als bei Annahme normalverteilter Renditen vorhergesagt (starke Flanken). Schmid und Trede<sup>39</sup> ermitteln für Tagesrenditen der DAX30-Aktien Werte der Kurtosis von 4,1 bis 8,4, mit zwei Ausreißern im zweistelligen Bereich.

---

<sup>36</sup>Pagan referiert ausführlich den seinerzeitigen Stand der Forschung. Cont listet stilisierte Fakten auf und bespricht sie (mit Datenbeispielen) anhand ausgewählter Literatur. Rydberg stellt die wichtigsten stilisierten Fakten anhand von Renditen (verschiedener Frequenzen) der IBM-Aktie dar. Schmid und Trede illustrieren im Rahmen ihres Lehrbuchs grundlegende stilisierte Fakten anhand von Tagesrenditen von DAX-30-Aktien. Die aktuelle Arbeit von Sewell sichtet sehr breit, aber relativ kursorisch die Literatur. Taylor betrachtet, neben einer zusammenhängenden Darstellung des Themas, einzelne Reihen von Tagesdaten vielerlei Anlageklassen.

<sup>37</sup>Sewell (2011a, S. 1)

<sup>38</sup>Sewell (2011a, S. 1) zitiert Veröffentlichungen beginnend mit dem Jahr 1915.

<sup>39</sup>Schmid und Trede (2006, S. 34)

*Asymmetrie:* Ebenfalls häufig beobachtet wird eine Linksschiefe von Renditeverteilungen, wobei hier der Befund nicht eindeutig ist.<sup>40</sup> Bei den DAX-Tagesrenditezeitreihen, für die Schmid und Trede<sup>41</sup> die Schiefe berechnen, weisen deutlich weniger als die Hälfte (11 von 31) eine negative Schiefe auf. In engem Zusammenhang mit der beobachteten Linksschiefe steht der Leverage-Effekt, dass die meisten Maße temporärer Volatilität negativ mit den Renditen korreliert sind. Pagan<sup>42</sup> macht hier folgende Beobachtung: Tests auf Symmetrie anhand der Schiefe lehnen die Nullhypothese  $\gamma_1 = 0$  deutlich ab. Monatsrenditen sind demnach signifikant rechtsschief ( $t$ -Statistik 3,04), Tagesrenditen signifikant linksschief ( $t$ -Statistik -13,1). Sodann führt er eine modifizierte Version des Tests durch, die robust ist gegenüber Abhängigkeiten von  $\hat{\gamma}_1$  mit den niedrigeren empirischen Momenten  $\hat{\mu}$  und  $\hat{\sigma}^2$ . Die Teststatistiken behalten ihr Vorzeichen, liegen aber betragsmäßig unter 1, sind also nicht signifikant. Dies kann an einer geringeren Testgüte des robusten Tests liegen, es kann aber, was das starke Absinken der Teststatistik bei Tagesrenditen betrifft, auch ein Ausfluss des Leverage-Effekts sein: Reagieren die Finanzmärkte stärker auf negative Nachrichten als auf positive und sind somit starke Renditeausschläge meist negativ, dann resultiert hieraus eine negative Schiefe.

*Existenz der Momente:* Sind die Flanken der Verteilung sehr stark besetzt, stellt sich die Frage, ob alle Momente der Verteilung existieren und ggf. welche nicht. Die univariaten Verteilungsmodelle haben klar definierte Eigenschaften betreffend der Existenz von Momenten. Unter Annahme der Normalverteilung sind alle Momente existent (und endlich). Im Falle der  $t$ -Verteilung mit  $\nu$  Freiheitsgraden existieren die ersten  $\nu - 1$  Momente, während für stabile Pareto-Verteilungen maximal das erste Moment definiert ist. In dieser Hinsicht beobachtet Rydberg<sup>43</sup>, es scheine einen allgemeinen Konsens zu geben, dass Tagesdaten endliche Varianz haben und stabile Pareto-Verteilungen dafür wenig geeignet seien. Er verweist dabei auf Eberlein und Keller (1995), die (S. 287) anhand von DAX30-Aktien zeigen, dass Tagesrenditen, für die der Formparameter  $\alpha$  auf deutlich unter 2 geschätzt wird, bei zunehmender Aggregation Schätzwerte von  $\alpha$  um 2 und darüber liefern, die Verteilung also nicht stabil ist.<sup>44</sup> Loretan und Phillips (1994) schätzen die Anzahl existierender Momente mittels eines Flankenindex auf zwei bis drei.<sup>45</sup> Mills und Markellos<sup>46</sup> kommen mit Tagesrenditen von FT30 und S&P500 zu ähnlichen Ergebnissen.

*Aggregation:* Je größer die Frequenz der Renditedaten ist, desto stärker nähert sich die Renditeverteilung der Normalverteilung an.<sup>47</sup> Rydberg vergleicht

<sup>40</sup>Cont (2001, S. 224) und Rydberg (2000, S. 238) zählen sie zu den stilisierten Fakten.

<sup>41</sup>Schmid und Trede (2006, S. 34)

<sup>42</sup>Pagan (1996, S. 35)

<sup>43</sup>Rydberg (2000, S. 237)

<sup>44</sup>Rachev et al. (2005) listen in ihrer Verteidigung des Stable-Paretian-Modells (gegenüber dem GARCH-Normal-Modell) den Punkt des Aggregationsverhaltens gerade nicht unter den von ihrem Modell abgebildeten stilisierten Fakten auf.

<sup>45</sup>Sie kommen dabei auf das etwas kontraintuitive Ergebnis, dass für Monatsrenditen tendenziell weniger Momente existierten als für Tagesrenditen, führen dies aber nicht weiter aus.

<sup>46</sup>Mills und Markellos (2008, S. 258)

<sup>47</sup>Cont (2001, S. 224)

die geschätzten Dichten von Aktienrenditen verschiedener Frequenzen (halbstündlich, täglich, wöchentlich, monatlich) miteinander<sup>48</sup>: Mit abnehmender Frequenz werden die Abweichungen von der Normalverteilung geringer. Jedoch sind auch bei Monatsdaten am unteren Rand der Verteilung noch deutliche Abweichungen von der Normalverteilung erkennbar.

*Autokorrelation der Renditen:* Theoretische Überlegungen sprechen gegen eine erhebliche Autokorrelation von Renditen börsengehandelter Wertpapiere, in erster Linie die sich ansonsten ergebenden Arbitragemöglichkeiten betreffend. Cont<sup>49</sup> ist der Ansicht, Autokorrelation sei nur im Falle von Hochfrequenzdaten relevant. Für Zeitintervalle größer als ca. 15 Minuten könne man für alle praktischen Zwecke eine Autokorrelation von null annehmen. Sewell<sup>50</sup> verweist hingegen auf eine größere Anzahl Studien, die quer über alle Renditefrequenzen positive oder negative, statistisch signifikante Renditeautokorrelationen entdecken, die aber ein sehr uneinheitliches Bild abgeben. Schmid und Trede<sup>51</sup> testen mittels Box-Pierce-, Ljung-Box- und Varianz-Quotienten-Tests die Tagesrenditen von DAX30-Aktien. Dabei wird die Nullhypothese, die (zentrierten) Renditen folgten einem White-Noise-Prozess, in deutlich mehr als 5 % der Fälle zum 5-%-Niveau abgelehnt, bei einigen Tests sogar in über der Hälfte der Fälle. Taylor<sup>52</sup> kommt, ebenfalls mit dem Box-Pierce-Test, zu einem ähnlichen Ergebnis. Grad und Richtung der Autokorrelation seien innerhalb der Wertpapierklassen unterschiedlich, bei Aktien etwa sei der Autokorrelationskoeffizient  $\rho_{t-1,t}$  bei großen Unternehmen höher als bei kleineren Unternehmen. Insofern muss man bei Tagesrenditen von geringer Autokorrelation ausgehen. Da die Autokorrelationen bei größeren Lag-Längen meist knapp um null streuen, dürfte der Effekt bei Monatsrenditen nur noch sehr gering sein.

*Autokorrelation quadrierter bzw. absoluter Renditen:* Dass die quadrierten Renditen  $r^2$  bzw. die Absolutwerte  $|r|$  der Renditen eine deutliche positive Autokorrelation aufweisen, und zwar über sehr viele Wertpapierklassen und Renditefrequenzen hinweg, kann als unstrittig gelten und wird in allen o.g. Veröffentlichungen behandelt<sup>53</sup>. Dieses Phänomen führte u. a. zur Entwicklung der im vorigen Abschnitt beschriebenen GARCH-Modelle. Die Frage bleibt, inwieweit ein Standardansatz, etwa ein GARCH(1,1)-Modell mit normalverteilten Störtermen, Finanzzeitreihen hinreichend beschreibt.

Curto et al. (2009) untersuchen die Anpassung von GARCH-Modellen mit unterschiedlichen Verteilungsannahmen für die Störterme (normal-,  $t$ - und stabile Pareto-Verteilung) anhand der Tagesschlusskurse dreier Aktienindizes. Die Anpassung an die Daten ist, gemessen an den Informationskriterien von Akaike und Schwarz, bei normalverteilten Störtermen schlechter als bei Modellen mit stabil oder  $t$ -verteilten Störtermen. Bei der out-of-sample-Prognose schneidet die  $t$ -Verteilung (mit geschätzter Zahl der Freiheitsgrade  $\nu$  zwischen 6 und 10)

---

<sup>48</sup>Rydberg (2000, S. 240 f.)

<sup>49</sup>Cont (2001, S. 229)

<sup>50</sup>Sewell (2011a, S. 2)

<sup>51</sup>Schmid und Trede (2006, S. 158-161)

<sup>52</sup>Taylor (2005, S. 77 ff.)

<sup>53</sup>z.B. Cont (2001, S.230 f.), Schmid und Trede (2006, S. 167 ff.), Rydberg (2000, S. 239 f.)

besser ab als die beiden Alternativen. Engle und Bollerslev (1986) kommen für wöchentliche USD-CHF-Wechselkursveränderungen auf Werte von  $\nu$  um 10 herum.

Ein weiterer Punkt ist die Persistenz der Autokorrelation der quadrierten Renditen. Ein GARCH(1,1)-Modell impliziert, dass der Autokorrelationskoeffizient  $\rho_{r_t^2, r_{t+s}^2}$  mit steigendem  $s$  exponentiell abfällt. Studien zeigen jedoch, dass diese Korrelation auch für große  $s$  noch systematisch größer als null ist. Rydberg (2000, S. 243) und Taylor (2005, S. 89) zeigen, dass die absoluten Tagesrenditen auch bei Lags von 150 - 200 Tagen noch signifikant positiv korreliert sind. Wertpapierrenditen haben also häufig ein „langes Gedächtnis“ (long memory).

*Kovarianz und Tail Dependence:* Die (unbedingte) Kovarianzmatrix der Renditen von Wertpapieren wird als Maß für die Abhängigkeit der Renditen kritisiert, weil sie von konstanter Korrelation zwischen den einzelnen Komponenten ausgeht, was empirisch nicht haltbar ist. Kommt es in einer Marktphase zu starken Kursausschlägen, sind die Korrelationen zwischen verschiedenen Wertpapieren häufig sehr hoch<sup>54</sup>.

*Weitere Merkmale:* Sewell (2011a, S. 8 ff.) listet eine Reihe von uhrzeit- und datumsbedingten Anomalien auf: unterschiedliche mittlere Renditen verschiedener Tageszeiten, Wochenend- und Feiertageeffekte, Renditeunterschiede zwischen Kalendermonaten („Januareffekt“), die für sich genommen Indizien für eine Nichtstationarität von Renditezeitreihen darstellen. Die Ausmaße dieser Phänomene sind aber sehr begrenzt und sie können als exogene Störungen aufgefasst werden.

Taylor (2005) weist darauf hin, dass viele der in diesem Abschnitt aufgeführten Effekte in ihrem Ausmaß stark von Ausreißern abhängen. Er führt einen Großteil seiner Berechnungen ein weiteres Mal, diesmal ohne die im zehn Jahre umfassenden Ursprungsdatensatz enthaltenen Daten der Woche vom 19. - 23. Oktober 1987 durch. In diese Woche fiel der „Schwarze Montag“, an dem der Dow Jones Industrial Average um 22,6 % verlor. Die Ergebnisse verändern sich qualitativ nicht, aber die Schätzwerte verschiedener Größen weichen teilweise stark ab.

### 2.3.1.2 Renten und Derivate

Die zweite große Anlageklasse börsengehandelter Wertpapiere sind Anleihen verschiedener Art, also handelbare Schuldverschreibungen bestimmter Emittenten, häufig auch Rentenpapiere genannt. Sie verbiefen klar definierte, nicht stochastische Zahlungsansprüche zu einem oder mehreren zukünftigen Zeitpunkten gegenüber dem Emittenten. Beim vorzeitigen Verkauf ist der erzielbare Verkaufspreis/Kurs der Anleihe abhängig vom Barwert der Zahlungsreihe, die die zeitlich gestaffelten Zins- und Tilgungsansprüche gegenüber dem Emittenten bilden. Der Barwert wird anhand der aktuellen Zinsstruktur ermittelt. Das Risiko besteht darin, dass sich die Marktzinssätze gegenüber dem Stand zur Emission bzw. zum Kaufzeitpunkt ändern. Dieses Zinsänderungsrisiko wird mit der Zeit kleiner,

<sup>54</sup>Cont (2001, S. 232)

da sich mit Herannahen der Fälligkeit der Barwert der Tilgungszahlung deren Nennwert annähert. Die Renditezeitreihe ist also nicht stationär. Hinzu kommt das Risiko, dass der Emittent die vereinbarten Zahlungen nicht oder nicht in vollem Umfang leisten kann. In diesem Fall der Insolvenz des Emittenten erhält der Investor die Zahlungen nicht bzw. vermindert und/oder verspätet. Steigt das Risiko eines solchen Zahlungsausfalls, sinkt der Kurs der Anleihe. Dieses Bonitätsrisiko wird häufig qualitativ in Form von Ratings renommierter Agenturen (i.d.R. Moody's, Standard & Poor's oder Fitch) angegeben. Zusammengefasst erscheint es wenig zweckmäßig, die Rendite einer Anleihe durch ein stationäres Verteilungsmodell zu modellieren, da die Rendite fast ausschließlich Ausfluss von Veränderung der Zinsstruktur und veränderten Bonitätseinschätzungen ist und sich die Eigenschaften der Rendite mit sinkender Restlaufzeit der Anleihe ändern.<sup>55</sup>

Die Preisentwicklung von Derivaten von Wertpapieren (Optionen, Futures, Swaps, Zertifikate,...) hängt im Wesentlichen von der jeweiligen Ausgestaltung und der Art sowie der aktuellen Marktsituation der zugrunde liegenden Basiswerte ab. Häufig, etwa bei Optionen, hängen die Eigenschaften der Rendite vom zeitlichen Abstand zu einem bestimmten Bewertungstag (z.B. Verfalltag bei Optionen) ab, so dass die Renditezeitreihe nicht stationär ist. Die Anpassung eines Renditeverteilungsmodells auf Klassen solcher Wertpapiere macht wenig Sinn. Die Preise solcher Instrumente sind in der Regel Resultate mathematischer Bewertungsmodelle<sup>56</sup>, die von den Marktteilnehmern eingesetzt werden.

In dieser Arbeit sollen also keine Renditemodelle für Rentenpapiere und Derivate verwendet werden. Diese Wertpapiere sind jedoch wichtige Anlageinstrumente für viele Investment- und Hedgefonds, die Thema der beiden folgenden Unterabschnitte sind.

### 2.3.2 Investmentfonds

Die Methoden der Performancemessung werden in der Regel nicht auf die Wertentwicklung einzelner Aktien, Anleihen und Derivate angewandt, sondern zur Bewertung der Wertentwicklung eines Anlageportfolios und somit der Leistung des verantwortlichen Portfoliomanagements. Bei einem Fonds verwaltet ein Vermögensverwalter (bzw. ein Vermögensverwaltungsunternehmen) eine Vermögensmasse, die mehreren Anlegern zu definierten Anteilen gehört. Man kann Fonds grob nach folgenden Kriterien einteilen:

- *Geschlossene und offene Fonds*: Bei geschlossenen Fonds werden vorab Zusagen für Anlagesummen eingesammelt. Bei Fondsstart werden diese Summen eingezahlt und der Fonds wird für weitere Einzahlungen geschlossen. Auch Mittelentnahmen sind in der Regel vor dem Ende der Fondslaufzeit nicht möglich. Die Laufzeit des Fonds ist vielfach vorab festgelegt. Diese

---

<sup>55</sup>Eine ausführliche Übersicht über Bewertungsmodelle für Anleihen bieten Fabozzi und Mann (2005)

<sup>56</sup>Siehe z.B. Hull (2011), Deutsch (2008), Franke et al. (2001), Steiner und Uhlir (2001, S. 215 ff.)

Eigenschaften erfordern, dass die Anleger ihr Geld für die gesamte Fondslaufzeit zur Verfügung stellen.<sup>57</sup> Andererseits hat die Fondsverwaltung durch das Schließen des Fonds eine kalkulierbare, langfristig verfügbare Anlagesumme und eine genaue Wertbestimmung der Fondsbestandteile ist nur in größeren Abständen bzw. am Ende der Laufzeit erforderlich. Bei offenen Fonds kann auch nach Fondsstart zu festgelegten Zeitpunkten, in der Regel börsentäglich, Geld eingezahlt oder aus dem Fonds entnommen werden. Offene Fonds bieten den Anlegern somit eine höhere Flexibilität bzw. Liquidität ihrer Mittel. Um nachträgliche Ein- und Auszahlungen zu ermöglichen, muss zu jedem Ein- und Auszahlungszeitpunkt der Wert des Fondsvermögens festgestellt werden. In der Regel geschieht dies mittels der Ausgabe von Fondsanteilen, die nicht auf einen bestimmten Bruchteil des Fondsvermögens lauten. Zu einem Ein- und Auszahlungszeitpunkt beläuft sich der Wert eines Fondsanteils auf den Wert des Fondsvermögens, geteilt durch die Anzahl umlaufender Fondsanteile. Zahlt ein Anleger in den Fonds ein, werden zum jeweiligen Anteilswert Fondsanteile neu geschaffen; durch Auszahlungen sinkt analog die Anzahl ausgegebener Fondsanteile, der Wert eines einzelnen Fondsanteils bleibt unberührt. In der Regel werden dabei börsentäglich Ausgabe- und Rückgabekurse veröffentlicht. Die ständig bestehende Möglichkeit, dass Fondsmittel abgezogen werden oder neu hinzukommende Gelder angelegt werden müssen, stellt hohe Ansprüche an die Liquidität des Fondsvermögens. Ausgabeaufschläge dienen neben der Abdeckung von Vertriebskosten häufig auch dazu, den Fonds für kurzfristige Anleger unattraktiv zu machen.

- *Börsennotierung der Anlageobjekte:* Fonds, die ausschließlich in börsennotierte Wertpapiere (und Bankguthaben) investieren, sind stets in der Lage, ihr gesamtes Fondsvermögen anhand der aktuellen Börsendaten zu bewerten. Enthält der Fonds nicht börsennotierte (oder nicht fungible) Anlageobjekte wie z.B. Immobilien, ist deren Wertbestimmung nur auf ungefährender Basis möglich. Da Neubewertungen vielfach nur in größeren Abständen erfolgen, kann die Renditezeitreihe daher deutlich glatter erscheinen.<sup>58</sup> Wertpapierfonds konzentrieren sich häufig auf einzelne Anlagesegmente wie Aktien oder Anleihen bestimmter Länder, Regionen oder Branchen.
- *Aktiv gemanagte und passive Fonds:* Der klassische Managementgedanke bei Fonds ist, dass ein fähiges Fondsmanagement durch ausführliche Recherche, gute Wertpapierausswahl und geschickte Wahl der Kauf- und Ver-

<sup>57</sup> Es gibt prinzipiell die Möglichkeit, Fondsanteile privat an Dritte zu veräußern, jedoch ist dieser Sekundärmarkt häufig kaum entwickelt.

<sup>58</sup> Dies sieht man deutlich beim Vergleich der Renditen von börsennotierten REITs (Real Estate Investment Trusts) und offenen Immobilienfonds nach deutschem Recht. Beide investieren ihre Mittel ähnlich; die Renditen von REITs ähneln dabei eher Aktienrenditen, während offene Immobilienfonds häufig jahrzehntelang fast wie eine Festgeldanlage rentieren. In der Immobilienkrise ab 2008 jedoch waren viele Fonds zu massiven Abschreibungen auf ihren Bestand gezwungen, so dass z.B. der seit 1972 bestehende Fonds Degi Europa (ISIN DE0009807800) von Ende 2009 bis Mitte 2011 etwa 30 % seines Wertes einbüßte.

kaufszeitpunkte aktiv Anlageerfolge im Sinne der Zielsetzung des Fonds erstrebt. Die Kosten dafür werden durch eine Verwaltungsgebühr auf die Anleger umgelegt. Diese ist vom Anlageschwerpunkt des Fonds abhängig: Bei Geldmarktfonds liegt sie häufig unter 0,5 % des Nettofondsvermögens<sup>59</sup> pro Jahr, bei auf bestimmte Branchen oder Regionen spezialisierten Aktienfonds kann sie 1,5 % p.a. deutlich überschreiten. Hinzu kommen häufig noch weitergereichte Kosten, z.B. in Form einer Depotbankgebühr; für den Vertrieb fallen zusätzliche Gebühren an, meist in Form von einmaligen Ausgabeaufschlägen von i.d.R. bis zu 6 %, in Ausnahmefällen bis zu 10 % der Anlagesumme. Aktiv gemanagte Investmentfonds sind ein Hauptanwendungsgebiet der Performancemessung, sowohl als Bewertungsinstrument für das Portfoliomanagement als auch als Werkzeug zur Auswahl möglicher Anlageobjekte.

In den letzten Jahren sind Fonds und fondsähnliche Anlageinstrumente populär geworden, die zu deutlich reduzierten Kosten „passive“ Anlagemöglichkeiten bieten. Für deutlich geringere Gebühren (meist unter 1 % der Anlagesumme p.a., bei liquiden Anlagebereichen häufig ca. 0,3 %) wird lediglich versucht, einen bestimmten Marktindex möglichst genau nachzubilden.<sup>60</sup> Der Vertrieb erfolgt bei vielen Fonds über die Börse (Exchange Traded Funds), so dass keine Ausgabeaufschläge anfallen. Die „Leistung“ der Fondsverwaltung besteht hier mehr in technischer Abwicklung als in Anlagepolitik und die Wertentwicklung wird fast vollständig durch den abzubildenden Marktindex bestimmt, so dass Methoden der Performancemessung hier wenig Sinn machen.

Ein erheblicher Teil der Literatur über Investmentfonds beschäftigt sich mit der Frage, ob (manche) aktiv gemanagte Fonds eine systematisch höhere Rendite erreichen können als der Markt und wenn ja, ob diese Überrendite über mehrere Perioden hinweg Bestand hat. Einen kurzen Überblick über Literaturquellen hierzu findet man in Sewell (2011b).

- *Publikumsfonds und institutionelle Fonds*: Publikumsfonds sind Fonds, die auch an Privatanleger vertrieben werden können. Ihre Zulassung zum Vertrieb an Privatanleger geschieht unter Berücksichtigung von Verbraucherschutzbestimmungen, in Deutschland etwa dem Investmentgesetz (InvG, seit 01. Januar 2004). Andere Fonds dürfen nur an institutionelle Anleger (z.B. Banken, Versicherungen, Pensionsfonds) vertrieben werden. Darunter fallen auch die weiter unten behandelten Hedgefonds.

Kernanwendungsgebiet der Performancemessung ist somit der *aktiv gemanagte Wertpapierfonds*, wobei sich der Rest dieses Abschnitts auf offene Publikumsfonds beschränkt.

Die Renditen dieser Fonds werden auf Basis der (Rücknahme-)Kurse errechnet, die börsentäglich von den Fondsgesellschaften veröffentlicht werden. Es existiert

<sup>59</sup>Nettofondsvermögen = Aktiva des Fonds abzüglich Verbindlichkeiten

<sup>60</sup>Häufig erfolgt eine (bis auf Gebühren) genaue Replikation dadurch, dass der Fonds ein Portfolio von Wertpapieren beliebiger Art erwirbt und dessen Wertentwicklung durch ein Swapgeschäft mit einem Dritten gegen die Wertentwicklung des Zielindexes tauscht.



zwar bei vielen Fonds auch ein Börsenhandel, dieser dient aber in erster Linie dazu, Ausgabeaufschläge beim Erwerb dieser Fonds zu umgehen. Die Rendite des Fonds über eine Periode ist Ergebnis der Rendite der vom Fonds innerhalb dieser Periode gehaltenen Wertpapiere. Dementsprechend kann man erwarten, dass sich die statistischen Eigenschaften der Fondsbestandteile in den statistischen Eigenschaften der Renditen niederschlagen.

Vrontos et al. (2011) untersuchen Monatsrenditen von 108 großen US-Investmentfonds der Jahre 1990 bis 2005. Fast in jedem Fall liegt die Kurtosis der Fonds über drei, im Durchschnitt beträgt sie 6,34. Der Jarque-Bera-Test lehnt bis auf einen Fonds in allen Fällen die Nullhypothese normalverteilter Renditen ab. Während Autokorrelation der Renditen mittels des Ljung-Box-Tests<sup>61</sup> nur bei 4 % aller Fonds festgestellt wurde, wurde bei quadrierten bzw. absoluten Renditen die Nullhypothese „keine Autokorrelation“ in etwa der Hälfte bzw. drei Viertel der Fälle abgelehnt.

Zum Vergleich wurde diese Analyse auf etwas breiterer Datenbasis repliziert. Dazu wurden aus der CRSP Mutual Funds-Datenbank alle verfügbaren Monatsrenditen von Investmentfonds von Januar 2001 bis März 2011 heruntergeladen, die zu den jeweiligen Zeitpunkten ein Nettofondsvermögen von mindestens 50 Mio. USD sowie eine komplette Renditegeschichte von 2001 bis 2011 aufweisen konnten. Unter den 2946 US-amerikanischen Fonds befanden sich auch in kurzfristige Zinsanlagen investierende Fonds, deren Renditekurve der Entwicklung der kurzfristigen Zinsen folgte und deren Monatsrenditen eine fast perfekte Autokorrelation erster Ordnung bei minimaler Varianz aufwiesen. Deshalb wurden nur diejenigen Fonds weiter betrachtet, bei denen die Nullhypothese einer Einheitswurzel bei einem Einheitswurzeltest abgelehnt wurde.<sup>62</sup> Es verbleiben 2378 Fonds, aus deren Daten einige Parameter berechnet wurden, deren Verteilung in Tabelle 2.1 skizziert wird.

Mittelwert und Standardabweichung der einzelnen Renditezeitreihen im betrachteten Zeitraum hängen stark von der Marktphase bzw. vom Anlagespektrum des jeweiligen Fonds ab. Interessanter sind die skaleninvarianten Maße Schiefe und Kurtosis. Die Schiefe ist erwartungsgemäß bei der Mehrheit der Fonds leicht negativ, die Kurtosis etwas größer als die der Normalverteilung und im Mittel ähnlich hoch wie bei Vrontos et al.. Eine Kurtosis von 4,3 (1. Quartil) bis 6,3 (3. Quartil) entspräche bei  $t$ -verteilten Daten etwa 5,8 bis 8,6 Freiheitsgraden<sup>63</sup>. Als Tests auf Normalverteilung der Renditen werden der Anderson-Darling- und der Jarque-Bera-Test verwendet. Beide Tests testen die Nullhypothese normalverteilter Daten. Dabei prüft der Anderson-Darling-Test<sup>64</sup> die

<sup>61</sup>jeweils mit Anzahl Lags = 12 bei  $\alpha = 5\%$

<sup>62</sup>Dickey-Fuller-Test mit Trend und Drift,  $\alpha = 0,01$ . Siehe Dickey und Fuller (1981), erläutert in Greene (2008, S. 745-754). Auf den Einsatz eines Augmented-Dickey-Fuller-Tests wurde verzichtet, da Autokorrelationen höherer Ordnung im Datensatz sehr gering sind und der Einbezug höherer Ordnungen angesichts der geringen Zeitreihenlängen ( $n = 123$ ) die Güte des Tests wohl deutlich verschlechtert hätte.

<sup>63</sup>nach der Formel  $\gamma_2 = 6/(\nu - 4) + 3 \Leftrightarrow \nu = 6/(\gamma_2 - 3) + 4$

<sup>64</sup>Anderson und Darling (1954), erläutert in Thode (2002, S.104) oder Sachs und Hedderich (2006, S. 342).

Maß	Mittelwert	Min.	1. Quart.	Median	3. Quart.	Max.
$\hat{\mu}^1$	0,00366	-0,01092	0,00230	0,00338	0,00483	0,02145
$\hat{\sigma}$	0,04433 <sup>2</sup>	0,00174	0,01551	0,04522	0,05478	0,16160
Schiefe	-0,9275	-6,8970	-1,0890	-0,8900	-0,7191	1,1290
Kurtosis	5,971	2,700	4,349	5,098	6,220	57,530
$\hat{\rho}_{r_t, r_{t-1}}$	0,1780	-0,2112	0,1220	0,2077	0,2477	0,6262
$\hat{\rho}_{r_t, r_{t-1}}(01-05)$	0,0746	-0,2152	-0,0004	0,0863	0,1430	0,9704
$\hat{\rho}_{r_t, r_{t-1}}(06-11)$	0,2340	-0,2967	0,1780	0,2705	0,3173	0,6688
$\hat{\rho}_{ r_t ,  r_{t-1} }$	0,2169	-0,0949	0,1404	0,2209	0,2859	0,7164
$\hat{\rho}_{r_t^2, r_{t-1}^2}$	0,2372	-0,1099	0,1482	0,2504	0,3184	0,7162
relative Ablehnhäufigkeit bei Tests auf $H_0$ : Normalverteilung						
Anderson-Darling, $\alpha = 5\%$ :			92,34 %		Anderson-Darling, $\alpha = 1\%$ :	
					83,26 %	
Jarque-Bera, $\alpha = 5\%$ :			96,59 %		Jarque-Bera, $\alpha = 1\%$ :	
					93,78 %	
relative Ablehnhäufigkeit Ljung-Box-Test auf $H_0$ : keine Autokorrelation für $r_t$						
Lags = 1, $\alpha = 5\%$ :			66,61 %		Lags = 12, $\alpha = 5\%$ :	
					28,93 %	
relative Ablehnhäufigkeit Ljung-Box-Test auf $H_0$ : keine Autokorrelation für $ r_t $						
Lags = 1, $\alpha = 5\%$ :			64,68 %		Lags = 12, $\alpha = 5\%$ :	
					79,64 %	
relative Ablehnhäufigkeit Ljung-Box-Test auf $H_0$ : keine Autokorrelation für $r_t^2$						
Lags = 1, $\alpha = 5\%$ :			70,02 %		Lags = 12, $\alpha = 5\%$ :	
					73,04 %	

<sup>1</sup> Die Zahlen zur mittleren Rendite sind mit Vorsicht zu betrachten, da nur Fonds mit vollständiger Renditegeschichte einbezogen wurden. Daraus ergeben sich die in Abschnitt 2.3.3 unter dem Stichwort *Survivorship Bias* beschriebenen Verzerrungen.

<sup>2</sup> Wurzel des Mittelwerts der Varianzen  
Quelle: eigene Berechnung aus Daten von CRSP

Tabelle 2.1: Kennzahlen der Monatsrenditen von 2 378 amerikanischen Investmentfonds 2001-2011

Anpassung der empirischen Verteilungsfunktion an die Verteilungsfunktion der Normalverteilung in ihrem gesamten Verlauf, hat aber im Verhältnis zum verbreiteten Kolmogorow-Smirnow-Test eine höhere Güte im Falle von Abweichungen an den Flanken. Der Jarque-Bera-Test testet auf Abweichungen der Schiefe und Kurtosis einer Verteilung gegenüber den Werten der Normalverteilung von 0 bzw. 3. Aufgrund der Verwendung höherer Momente ist er sehr ausreißer anfällig.<sup>65</sup> Beide Tests führen in der Regel zur Ablehnung der Nullhypothese auch auf dem 1 % -Niveau, in besonderem Maße beim Jarque-Bera-Test. Es kann also davon ausgegangen werden, dass das Normalverteilungsmodell insbesondere starke Flanken nicht ausreichend abbildet.

Etwas überraschend sind die Ergebnisse für die Autokorrelation erster Ordnung  $\rho_{r_t, r_{t-1}}$ . Anders als bei Vrontos et al. lässt sich bei den meisten Renditezeitreihen eine deutlich positive Autokorrelation feststellen. Im Median liegt der Koeffizient

<sup>65</sup> Benannt nach dem Artikel Jarque und Bera (1980), erläutert in Hendry und Nielsen (2007, S. 128). Die Testidee ist wesentlich älter, siehe hierzu Thode (2002, S. 41 - 68).

der Autokorrelation erster Ordnung bei 0,2077, also deutlich über dem erwarteten Wert von 0. Der Ljung-Box-Test<sup>66</sup> mit einem Lag, der also nur auf Autokorrelation erster Ordnung prüft, lehnt bei  $\alpha = 5\%$  in zwei Drittel der Fälle die Nullhypothese  $\rho_{r_t, r_{t-1}} = 0$  ab. Werden analog zu Vrontos et al. 12 Lags, entsprechend einem Jahr, einbezogen, sinkt der Anteil auf unter 30 %. Da die Nullhypothese in letzterem Test noch restriktiver ist ( $\rho_{r_t, r_{t-1}} = \rho_{r_t, r_{t-2}} = \dots = \rho_{r_t, r_{t-12}} = 0$ ), liegt dieses Ergebnis an der geringeren Güte des Tests bei höheren Laglängen. Man kann darauf schließen, dass Lags höherer Ordnung kaum oder keine Korrelation mit den aktuellen Fondsrenditen haben. Das Phänomen höherer Autokorrelation scheint neueren Datums zu sein. Teilt man den 10-Jahres-Datensatz in zwei Datensätze (2001-2005 und 2006-2011) auf, ist die Autokorrelation erster Ordnung im ersten Zeitabschnitt (Median: 0,0863) weitaus kleiner als im zweiten Zeitabschnitt (Median: 0,2705). Die Autokorrelationen erster Ordnung der absoluten und der quadrierten Renditen sind für den Zeitraum 2001-2011 nur leicht höher als die Autokorrelation der Originalrenditen. Der Ljung-Box-Test mit Laglänge 1 lehnt ähnlich häufig ab wie oben (65 % bzw. 70 % bei  $\alpha = 0,05$ ). Erhöht man hier die Laglänge auf 12, steigt die Ablehnungsquote auf 80 % bzw. 73 %, was in etwa den von Vrontos et al. genannten Werten entspricht. Die höhere Ablehnungsquote bei größerer Laglänge spricht dafür, dass auch Autokorrelationen über mehr als einen Monat hinweg nicht gleich null sind.

Um der Frage nachzugehen, ob sich die Charakteristika der Fondsrenditen hinsichtlich der Autokorrelation seit der Analyse von Vrontos et al. geändert haben, wurde aus CRSP ein Datensatz heruntergeladen, der in wesentlichen Merkmalen den Daten von Vrontos et al. entspricht. In Tabelle 2.2 sind deshalb die obigen Ergebnisse, die Angaben aus Vrontos et al. (2011) und eine Vergleichsanalyse nebeneinander gestellt. Dafür wurden sämtliche Fondsmonatsrenditen aus dem von Vrontos et al. gewählten Betrachtungszeitraum (April 1990 bis Dezember 2005) heruntergeladen – ohne Beschränkungen hinsichtlich des Fondsvolumens – und mit exakt denselben Methoden analysiert wie in Tabelle 2.1. Es wurden also nur diejenigen Fonds einbezogen, für die über den gesamten Zeitraum Renditen vorliegen und für deren Renditezeitreihe der Dickey-Fuller-Test eine Einheitswurzel ablehnt. Dies traf auf 1426 Fonds zu.

Die Ergebnisse sind nicht vollständig vergleichbar, weil zum einen die Grundgesamtheiten der drei Untersuchungen unterschiedlich sind und zum zweiten die Stichprobe bei Vrontos et al. erheblich kleiner ist. Zum dritten ist die Zeitreihenlänge in der dritten Untersuchung etwas kleiner, wodurch die Güte und damit die Ablehnhäufigkeit der Tests geringer ist. Dennoch kann man erkennen, dass die Abweichung von der Normalverteilung von 1990 - 2005 auf 2001 - 2011 zugenommen hat. Die durchschnittliche Autokorrelation der Renditen hat ebenfalls deutlich zugenommen, was man sowohl an den Werten für  $\bar{\rho}_{r_t, r_{t-1}}$  als auch an den Ablehnhäufigkeiten der Ljung-Box-Tests sieht. Hinsichtlich der Autokorrelation der absoluten bzw. quadrierten Renditen stellt man fest, dass diese für den Zeitraum 1990 - 2005 unerwartet niedrig sind, auch wenn die Testergebnisse

<sup>66</sup>Siehe Ljung und Box (1979), erläutert in Greene (2008, S. 729).

	Vrontos et al. <sup>1</sup>	1990 - 2005 <sup>2</sup>	2001 - 11 <sup>3</sup>
Anzahl Fonds	108	1426	2378
Zeitreihenlänge	189	189	123
Schiefe ( $\hat{\gamma}_1$ )	-	-0,3426	-0,9275
Kurtosis ( $\hat{\gamma}_2$ )	6,34	4,646	5,971
Ablehnung AD-Test ( $\alpha = 5\%$ )	-	58,56 %	92,34 %
Ablehnung JB-Test ( $\alpha = 5\%$ )	-	49,58 %	96,59 %
$\hat{\rho}_{r_t, r_{t-1}}$	-	0,0789	0,1780
Ablehnung LB-Test für $r_t$ ( $\alpha = 5\%$ , Lags 1)	-	20,90 %	66,61 %
( $\alpha = 5\%$ , Lags 12)	3,7 %	11,57 %	28,93 %
$\hat{\rho}_{ r_t ,  r_{t-1} }$	-	0,1000	0,2169
Ablehnung LB-Test für $ r_t $ ( $\alpha = 5\%$ , Lags 1)	-	25,88 %	64,68 %
( $\alpha = 5\%$ , Lags 12)	77,8 %	48,74 %	79,64 %
$\hat{\rho}_{r_t^2, r_{t-1}^2}$	-	0,0850	0,2372
Ablehnung LB-Test für $r_t^2$ ( $\alpha = 5\%$ , Lags 1)	-	20,62 %	70,02 %
( $\alpha = 5\%$ , Lags 12)	45,4 %	42,78 %	73,04 %

<sup>1</sup> Angaben aus Vrontos et al. (2011, S. 298) entnommen, beziehen sich lt. Autoren auf 108 Aktienfonds mit Nettofondsvermögen > 500 Mio USD aus CRSP, die für den Betrachtungszeitraum (1990 - 2005) eine vollständige Renditehistorie aufweisen konnten.

<sup>2</sup> Alle Monatsrenditen von Fonds, die von 1990 - 2005 in CRSP eine vollständige Renditehistorie hinterlegt haben und nach DF-Test keine Einheitswurzel aufwiesen.

<sup>3</sup> Alle Monatsrenditen von Fonds, die von 2001 - 2011 in CRSP eine vollständige Renditehistorie hinterlegt haben, ständig ein Nettofondsvermögen > 50 Mio. USD und nach DF-Test keine Einheitswurzel aufwiesen.

Tabelle 2.2: Vergleich der Kennzahlen für Investmentfondsmonatsrenditen

in etwa mit denen bei Vrontos et al. korrespondieren. Es ist allerdings auch zu erwarten, dass die Autokorrelation der absoluten und der quadrierten Renditen bei Monatsdaten geringer ist als bei den häufiger untersuchten Tagesdaten. Externe Schocks kommen in der Regel plötzlich und bei höherfrequenten Daten können sich kurzfristige Volatilitätsspitzen über mehrere Perioden ausbilden. Nach einem oder mehreren Monaten dagegen ist der Schock häufig schon verarbeitet. Offenbar ist die Autokorrelation der absoluten/quadrerten Renditen im Zeitraum 2001 - 2011 höher als im früher liegenden Zeitraum.

Insgesamt können wir also weder die Annahme normalverteilter Renditen noch die Annahme unabhängig und identisch verteilter Renditen als unproblematisch für Monatsrenditen von Investmentfonds betrachten. Die Abweichungen sind aber auch nicht zwangsläufig gravierend: Schiefe und Kurtosis weichen nicht stark von den Werten der Normalverteilung ab. Dass in beiden, recht großen Gesamtheiten das ausreißeranfällige arithmetische Mittel der Kurtosis klar im

einstelligen Bereich blieb, deutet darauf hin, dass mindestens das vierte Moment der Renditeverteilung endlich sein dürfte: Hätte die wahre Renditeverteilung kein endliches viertes Moment, wären einige Ausreißer zu erwarten gewesen, die das arithmetische Mittel der Kurtosis stark aufgebläht hätten. Ein Autokorrelationskoeffizient erster Ordnung von 0,178 entspräche bei einer Regression einer Rendite auf die Rendite der Vorperiode einem  $R^2$  von nur 0,0317.

### 2.3.3 Hedgefonds

Auf die Frage nach einer Definition des Begriffs „Hedgefonds“ gibt es die halb scherzhafte Antwort: „*Anything that charges two and twenty*“. Diese Definition nimmt vordergründig Bezug auf die typische Vergütungsstruktur von Hedgefonds, die in der Regel einen fixen Prozentsatz des Fondsvermögens, häufig 2 % p.a., einschließt, sowie eine erfolgsabhängige Gebühr, häufig 20 % des über ein definiertes Renditeziel hinausgehenden Ertrags. Der eigentliche Punkt der Definition liegt aber darin, dass demnach ansonsten nichts als gemeinsames Merkmal von Hedgefonds geeignet sei.

Der Name der Investmentklasse kommt vom englischen *to hedge*, ursprünglich das Anlegen von Hecken (*hedges*) zum Schutz von Weiden (dt.: einhegen). Das Konzept der als erster Hedgefonds betrachteten Anlagegesellschaft von Alfred W. Jones in den 1950er Jahren war es, durch Kauf „günstiger“ und gleichzeitigen Leerverkauf „überbewerteter“ Aktien einerseits Gewinne zu machen und andererseits das Portfolio gegenüber dem Auf und Ab des Gesamtmarktes zu schützen. Heute verfolgen Hedgefonds alle erdenklichen Anlagestrategien; gehalten hat sich bei vielen jedoch der Anspruch, unabhängig von der Entwicklung des Gesamtmarktes beständig positive Renditen erzielen zu wollen. Dies ist jedoch nicht strikt der Fall: Neben Hedgefonds, die explizit eine positive (oder negative) Korrelation mit dem Markt anstreben, haben die Mehrzahl der Hedgefondsrenditen eine positive Korrelation mit dem Aktienmarkt.<sup>67</sup>

Ein Hedgefonds ist letztlich eine Vermögenmasse, die von Kapitalgebern an einen oder mehrere von ihnen für ausreichend befähigt gehaltene Vermögensmanager zur Verwaltung übergeben wurde, wobei regelmäßig erwartet wird, dass die Vermögensmanager selbst einen Teil ihres Privatvermögens in den Fonds investieren. Dabei wird dem Management in der Regel sehr große Entscheidungsfreiheit hinsichtlich Anlagestrategie und einsetzbarer Finanzinstrumente gelassen. Das Management soll möglichst wenig durch staatliche Regulierung in ihrer Handlungsfreiheit beschränkt oder durch Berichtspflichten belastet werden. Dies und steuerliche Gesichtspunkte führen dazu, dass Hedgefonds meist in der Rechtsform der *Limited Partnership* oder *Limited Liability Company* organisiert sind<sup>68</sup> und ihren rechtlichen Sitz in angelsächsischen Ländern oder in Offshore-Finanzplätzen wie den Kaimaninseln oder den Britischen Jungferninseln haben. Bezeichnend für die Scheu vor Regulierung ist die Situation in Deutschland. Seit dem Inkrafttreten des Investment-Modernisierungsgesetzes

---

<sup>67</sup>vgl. z.B. Lo (2008, S. 38)

<sup>68</sup>Siehe Lhabitant (2002, S. 16)

am 1. Januar 2004 sind „Sondervermögen mit zusätzlichen Risiken“ auch in Deutschland zum Vertrieb zugelassen, wobei eine Aufsicht durch die Bundesanstalt für Finanzdienstleistungsaufsicht (BaFin) erfolgt. Wohl auch aufgrund dieser vorgeschriebenen Regulierung waren in Deutschland mit Stand 28. September 2011 lediglich 12 „Sondervermögen mit zusätzlichen Risiken gemäß § 112 InvG“ genehmigt.<sup>69</sup> Weltweit zählte der Hedgefonds-Dienstleister HFR im Jahr 2009 jedoch 6644 Hedgefonds.<sup>70</sup> Eine auch nur annähernd genaue Zahl der Hedgefonds weltweit lässt sich nicht ermitteln, da keine Registrierungspflicht besteht und die Grenzen zu anderen Anlagekategorien wie Private-Equity-Fonds fließend sind. Starke Regulierungen erfahren Hedgefonds lediglich dahingehend, dass Werbung sowie der Vertrieb an kleine und mittlere Privatanleger in den meisten Ländern verboten sind.

Kapitalgeber von Hedgefonds haben im Vergleich zur Anlage in Publikumsfonds einige Einschränkungen in Kauf zu nehmen. Neben hohen Mindestanlagensummen ist ein Einstieg in Hedgefonds nicht jederzeit möglich, sondern nur bei Auflage des Fonds, häufig auch zu festgelegten späteren Einstiegsterminen. Fast immer gibt es eine Mindestanlagedauer (*lockup period*), danach können Anleger in der Regel unter Einhaltung bestimmter Vorankündigungsfristen zu festgelegten Zeitpunkten Kapital abziehen. Durch diese Regelungen erhält das Management Planungssicherheit für ihre Anlagestrategie. Anders als bei Publikumsfonds muss nicht ständig Liquidität für kurzfristig abziehbare Gelder vorgehalten werden. Eine Vermögensbewertung und -aufstellung ist nur zu den Ein- und Auszahlungsterminen erforderlich, häufig einmal im Monat oder im Quartal. Hauptkapitalgeber von Hedgefonds sind sehr wohlhabende Einzelpersonen und Familien sowie institutionelle Anleger.

Ein weiterer häufig angeführter Unterschied von Hedgefonds gegenüber Publikumsfonds ist der, dass die Fondsgröße einen (nichtlinearen) Einfluss auf den Erfolg der Anlagestrategie haben kann. Einerseits ist ein gewisses Mindestvermögen günstig, um eine erkannte Anlagechance voll auszunutzen, ohne das Risikoprofil des Fonds zu stark zu belasten. Andererseits können Manager sehr großer Fonds Schwierigkeiten bekommen, das Kapital effektiv einzusetzen, da einzelne Anlagechancen in der Regel nur eine limitierte Einstiegssumme erlauben, ohne dass die zusätzliche Nachfrage des einsteigenden Hedgefonds das Preisgefüge am Markt so verändert, dass die Anlagechance verschwindet. Die Fondsgröße kann durch Einzahlungsstopps und Kapitalausschüttungen nach oben begrenzt werden.

Publikumsfonds unterliegen strengen Publizitätspflichten. Sie veröffentlichen börsentäglich Anteilspreise und müssen jährlich, meist auch zusätzlich in kürzeren Abständen, Berichte veröffentlichen, in denen sie über die von ihnen gehaltenen Wertpapiere sowie getätigte Transaktionen Rechenschaft ablegen. Bei Hedgefonds ist dies nicht der Fall. Bestände und Transaktionen werden meist als Geschäftsgeheimnisse behandelt, da diese Daten Rückschlüsse auf die Anlagestrategie des Fondsmanagements zulassen. Angaben zu Fondsvolumen und

---

<sup>69</sup>siehe BaFin (2011)

<sup>70</sup>ohne Dachfonds, lt. HFR (2009), zitiert nach Kaiser (2009, S. 14)

Wertentwicklung werden von vielen Hedgefonds nur an Kapitalgeber und ausgewählte Personenkreise weitergereicht. Zwecks Neukundengewinnung reichen viele Hedgefonds monatlich ihre Renditedaten auf freiwilliger Basis an Fondsdatenanbieter wie Center for International Securities and Derivatives Markets (CISDM), Hedge Fund Research (HFR), BarclayHedge, HedgeFund.net (HFN), Dow Jones Credit Suisse (DJCS) oder Lipper TASS weiter, die diese in aufbereiteter Form an interessierte Nachfrager verkaufen. Außerdem berechnen viele Anbieter Indizes für den Hedgefonds-Gesamtmarkt, bestimmte Teilmärkte und Strategiegruppen.<sup>71</sup> Die von Hedgefonds-Datenanbietern vertriebenen Datensätze sind die Datengrundlage für empirische Studien in Wissenschaft und Finanzwirtschaft. Will man Daten dieser Anbieter für die Analyse des Hedgefonds-Gesamtmarktes oder typischer Eigenschaften von Hedgefondsdaten nutzen, muss man sich einer Reihe potentiell störender Effekte bewusst sein, die zu großen Teilen auf der Freiwilligkeit der Datenlieferung durch die jeweiligen Hedgefonds beruhen.<sup>72</sup>

- *Survivorship Bias*: Publikums- und Hedgefonds mit schlechter Wertentwicklung stellen häufig aus diesem Grund – bzw. aufgrund der damit häufig verbundenen Mittelabflüsse – ihren Betrieb ein und liefern keine Renditedaten mehr. Sich gut entwickelnde Fonds werden nur selten aufgelöst. Berücksichtigt man bei der Untersuchung eines Zeitraums nur Fonds, für die für den gesamten Zeitraum Renditedaten vorliegen, wie ich es in Abschnitt 2.3.2 gemacht habe, oder nur Fonds, die zum Zeitpunkt des Datenabrufs noch existieren, dann „vergisst“ man dabei Fonds, die sich schlecht entwickelt haben und überschätzt damit die Marktentwicklung. Die von Kaiser (2009) zitierten Studien quantifizieren den Survivorship Bias auf 0,16 - 6,22 % pro Jahr, wobei drei Viertel der Angaben Werte zwischen 1 % und 4% p.a. nannten. Den Survivorship Bias kann man verhältnismäßig einfach vermeiden, indem man die Renditedaten aufgelöster Fonds in der Datenbank belässt (*fund graveyard*).
- *Self-Selection Bias*: Da die Hedgefondsbetreiber ihre Fondsrenditen zum großen Teil zum Zweck der Kundenakquise an die Datenbanken melden, kann man sich vorstellen, dass sie versucht sein könnten, Renditen sich schlecht entwickelnder Fonds nicht mehr weiterzugeben, um die Reputation der Gesellschaft oder der Fondsmanager zu schützen. Dieser Effekt wirkt ähnlich wie der Survivorship Bias, beruht aber auf dem freiwilligen Meldeverhalten des Fondsmanagements. Hier ist auch ein umgekehrter Effekt denkbar: Erfolgreiche Fonds ziehen neue Kapitalmittel an; wie

<sup>71</sup>Hierbei ordnen sich die Fonds selbst bestimmten Strategiegruppen zu, die je nach Datenanbieter unterschiedlich definiert sein können. Beispiele sind: Equity Market Neutral (Spekulation mit Aktien, möglichst keine Korrelation von Fonds- und Marktentwicklung), Global Macro (Makroökonomische Spekulation, z.B. mit Währungen), Event Driven (Spekulation im Umfeld von Übernahmen und anderen Ereignissen), Convertible Arbitrage (Arbitrage bei Wandelanleihen) und Statistical Arbitrage (Mathematisch fundierte Arbitragestrategien)

<sup>72</sup>Siehe z.B. Fung und Hsieh (2000); Kaiser (2009, S. 206-215) liefert neben einer Zusammenfassung einen ausführlichen Literaturüberblick.

oben dargestellt, ist ein Wachstum des Fonds über eine gewisse Größe hinaus oft gar nicht gewollt, so dass das Management das externe Reporting der Renditen einstellt. Aiken et al. (2012) untersuchen Daten der US-amerikanischen Börsenaufsicht SEC über die Anlageklasse *fund of hedge funds* und stellen fest, dass Fonds, die in einem bestimmten Quartal an Datenanbieter berichten, eine höhere Durchschnittsrendite aufweisen als solche, die gerade nicht berichten. Ein Blick auf die Verteilung der Renditen berichtender und nicht berichtender Fonds zeigt, dass bei letzteren stark negative Renditen deutlich häufiger auftauchen.

- *Instant History Bias / Stop-Reporting Bias*: Wenn ein Fonds erstmals seine Renditen an eine Datenbank meldet, sind meist schon einige Monate oder Jahre seit seiner Initiierung vergangen und häufig liefert der Fonds seine bisherige Renditegeschichte nach.<sup>73</sup> Dabei kann es sein, dass der Fonds seine Renditehistorie nur dann anbietet, wenn diese eindrucksvoll ist oder dass er Anfangsmonate streicht, wenn sie weniger gute Ergebnisse brachten. Schließlich kann es sein, dass ein Fonds, der sich von Beginn an schlecht entwickelt, nie an die Datenbanken berichtet. Dieser Effekt kann die berichtete Rendite des Fonds bzw. Gesamtmarkts künstlich aufblähen. Die von Kaiser (2009) aufgelisteten Studien beziffern dessen verzerrenden Einfluss auf 0,05 % bis 5,47 % pro Jahr. In einer neueren Studie identifizieren Aggarwal und Jorion (2010) ebenfalls hohe Werte für den Instant History Bias, kommen aber darüber hinaus zu dem Ergebnis, dass jüngere Fonds generell eine etwas höhere Rendite aufweisen. Die Verzerrung kann dadurch beseitigt werden, dass man nur Renditen einbezieht, die nach Beginn der Berichterstattung an die Datenbank entstanden sind, wofür dieses Beginndatum bekannt sein muss. Am anderen Ende der Lebensspanne eines Fonds ist anzunehmen, dass schlechte Renditen, die zur Überlegung oder dem Entschluss führen, den Fonds zu schließen, nicht mehr gemeldet werden.<sup>74</sup> Insbesondere ist fraglich, ob Einbrüche nach katastrophalen Fehlspekulationen, die unmittelbar zum Konkurs oder einer Abwicklung des Fonds führen, noch an die Datenbankanbieter gemeldet werden.<sup>75</sup> In der oben genannten Vergleichsstudie von Aiken et al. (2012) wird genau dieser Effekt beobachtet: Anlass für den Abbruch der Berichterstattung sind häufig stark negative Renditen.<sup>76</sup> Diese Fehlerquelle lässt sich nicht

<sup>73</sup>Nach Ibbotson et al. (2011, Tabelle A.2) besteht die TASS-Datenbank für die Jahre 1995 bis 2009 je nach Strategiegruppe zu 40 - 50 % aus „aufgefüllten“ Daten.

<sup>74</sup>Siehe hierzu ausführlich Grecu et al. (2007)

<sup>75</sup>Xu et al. (2010), die diesen Effekt näher untersuchen, nennen ihn *Hidden Survivorship Bias*.

<sup>76</sup>Zudem machen sie die Beobachtung, dass die Renditen von Fonds, die die Berichterstattung abbrechen, schon in den unmittelbar vorausgehenden Perioden tendenziell schwächer sind. Das könnte darauf hinweisen, dass Manager schlecht laufender Hedgefonds ihre Renditen zunächst etwas schönen, vielleicht in der Hoffnung, dass die tatsächlichen Verluste durch spätere Gewinne wieder aufgeholt werden können. Wird diese Hoffnung nicht erfüllt, würden die tatsächlich aufgelaufenen Verluste ausgewiesen, aber nicht mehr an Dritte berichtet. Diese Verzögerung bzw. Glättung von Verlusten fällt unter den unten erläuterten *Data Quality Bias*.



über die Auswertung der Stammdaten eines Fonds korrigieren.<sup>77</sup>

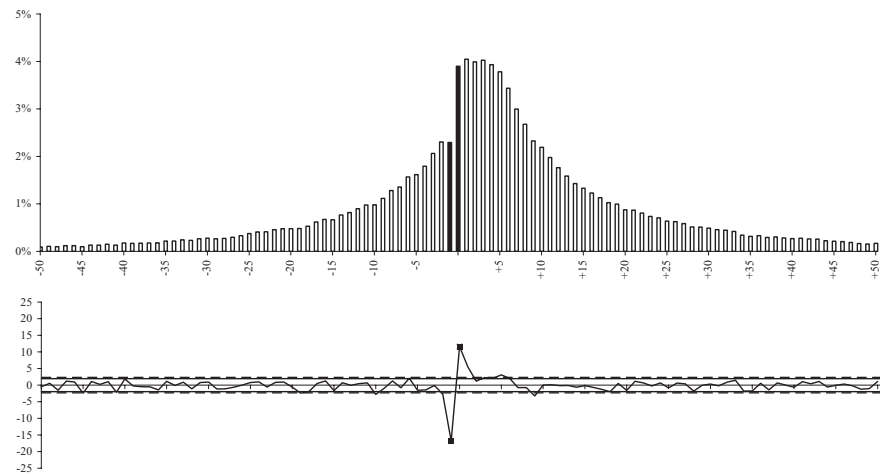
- *Existenz mehrerer Datenbankanbieter*: Deutliche Ungenauigkeiten ergeben sich daraus, dass es mehrere Datenbankanbieter gibt und Hedgefonds nach eigener Wahl an einen, mehrere, alle oder überhaupt keinen Datenbankanbieter berichten. Häufig wechseln Fonds den Datenbankanbieter<sup>78</sup>, ohne dass der bisherige oder der neue Anbieter vom jeweils anderen weiß; auch die übermittelten Daten müssen nicht übereinstimmen. Dies führt zu Fehlerquellen innerhalb der Datenbanken. Zum anderen enthält jede Datenbank einen anderen Ausschnitt aus dem Fondsuniversum. Fung und Hsieh (2006) untersuchen hierzu fünf große Datenbanken<sup>79</sup>. Von den von ihnen identifizierten Fonds sind 64 % jeweils nur in einer Datenbank enthalten, lediglich 3 % in allen fünf; die restlichen 33 % sind in zwei bis vier der fünf Datenbanken enthalten. Dies führt dazu, dass die Ergebnisse empirischer Studien zu Hedgefonds stark von der/den jeweils benutzten Datenbank(en) abhängen.
- *Data Quality Bias*: Die Datenbankanbieter können die ihnen von den Hedgefonds gemeldeten Renditedaten nicht selbst überprüfen, was Gelegenheit zum Verdacht gibt, Hedgefonds könnten ihre gemeldeten Renditen manipulieren. Häufiger Anhaltspunkt ist die recht hohe Autokorrelation von Hedgefondsrenditen, die dazu führt, dass Renditen „geglättet“ werden und dadurch geringere Schwankungen aufweisen, somit also risikoärmer aussehen als sie eigentlich sind. Dagegen wird angeführt, dass viele Anlageinstrumente von Hedgefonds illiquide und schwer zu bewerten seien, so dass ihr Wertansatz bei monatlichen Bewertungen nur zögernd angepasst werde.<sup>80</sup> Die Autokorrelation entstehe somit unbeabsichtigt. Einige Studien finden jedoch Hinweise auf gezielte Steuerung von Renditen: Straumann (2009) führt Tests auf fünf numerische Eigenschaften von Hedgefondsrenditen durch (z.B. sich wiederholende Renditesequenzen, empirische Verteilung der Ziffer an der zweiten Nachkommastelle) und findet erheblich mehr solcher Auffälligkeiten als in der Vergleichsgruppe (Aktienrenditen). Agarwal et al. (2007) beobachten ungewöhnlich hohe Renditen im Dezember. Cici et al. (2010) vergleichen die Bewertungen von Aktienpositionen, die Hedgefonds an die Wertpapieraufsicht melden, mit den tatsächlichen Kursen der entsprechenden Aktien am jeweiligen Bewertungsstichtag. Sie finden teilweise erhebliche Abweichungen, die z.T. auf bestimmte Manipulationsmotive hinweisen. Einen eindeutigen Hinweis auf mutwillige Steuerung von Hedgefondsrenditen liefern Bollen und Pool (2009): Hedgefondsmanager melden offenbar lieber kleine positive Renditen als kleine negative Renditen, wie das Histogramm in Abbildung 2.1 zeigt.

<sup>77</sup>Zusätzlich wird in der Literatur (z.B. Kaiser (2009, S. 212)) ein *Backdelete Bias* erwähnt, der entsteht, wenn Hedgefondsmanager nachträglich die Löschung ihres Fonds aus einer Datenbank verlangen können.

<sup>78</sup>Fung und Hsieh (2009) erörtern dieses Problem.

<sup>79</sup>CISDM, EUR, HFR, MSCI und TASS

<sup>80</sup>vgl. Lo (2008, S. 66 f.)



Oben: Ausschnitt aus einem Histogramm der Monatsrenditen aller Hedgefonds aus der Datenbank CISDM 1994 - 2005 ( $n = 215930$ ), ohne Nullrenditen, mit Balkenbreite von 19 Basispunkten. Die beiden schwarzen Balken markieren die Intervalle um null herum.

Unten: Teststatistik und Annahmehereich eines Tests auf Glätte des Histogramms aus: Bollen und Pool (2009).

Abbildung 2.1: Histogramm der Hedgefondsrenditen 1994 - 2005 aus CISDM

- *Zeit- und Wertgewichtung:* Für gewöhnlich erfolgt die zeitliche Aggregation von Renditen ohne Einbeziehung der jeweiligen Ein- und Auszahlungen des Fonds, man spricht hier von zeitgewichteter (*time-weighted*) Aggregation. Eine andere Möglichkeit ist die Betrachtung des Fonds als Investitionsprojekt mit Ein- und Auszahlungen, die wertgewichtete (*money-weighted*) Aggregation.<sup>81</sup> Dichev und Yu (2011) zeigen, dass die Rendite von Hedgefonds bei wertgewichteter Betrachtung erheblich niedriger liegt als bei zeitgewichteter „Buy-and-hold“-Betrachtung. Ihre Abbildung des Hedgefonds-Gesamtmarktes von 1980 bis 2008 (einschließlich) zeigt eine zeitgewichtete Rendite von 13,8 % bei einer wertgewichteten Rendite-

<sup>81</sup>Ein stark vereinfachtes Beispiel zur Zeit- und Wertgewichtung:

Ein Fonds hat zum 1. Januar eines Jahres ein Vermögen von 50 € und kann dies durch geschickte Anlage bis zum 1. Juli des Jahres auf 100 € verdoppeln. Zu diesem Termin werden weitere 100 € in den Fonds eingezahlt. Bis zum Jahresende erleidet der Fonds 40 % Verlust und hat ein Vermögen von 120 €.

Zeitgewichtet betrug die Rendite des Jahres  $2,0 \cdot 0,6 - 1 = 0,2$ , also + 20 %. Wer zu Beginn des Jahres 1 € in den Fonds angelegt hatte und ihn bis zum Jahresende hielt, hat also 20 % Gewinn gemacht.

Nach wertgewichteter Betrachtung wurden in den Fonds  $50 + 100 = 150$  € eingezahlt und 120 € ausgezahlt. Die Berechnung des internen Zinsfußes auf das eingesetzte Kapital ergibt

$$50 \cdot (1 + r) + 100 \cdot (1 + r)^{0,5} = 120 \Leftrightarrow r = -28,78\%.$$

Dies entspricht der Rendite des durchschnittlichen in diesem Jahr in den Fonds investierten Euros. Für nähere Erläuterungen zu verschiedenen Methoden der zeit- und wertgewichteten Renditeberechnung siehe Bacon (2004, S. 5 ff.)

te von 6,1 %. Dieser Unterschied hat zwei Ursachen. Erstens gewichtet der wertgewichtete Ansatz die erfolgreichen 1980er Jahre, als Hedgefonds noch recht exotische Anlagen waren, weit schwächer als die entwickelten Märkte seit 2000, in die das Krisenjahr 2008 fiel. Zweitens ziehen Fonds mit überdurchschnittlicher Wertentwicklung neue Anlegergelder an, nach deren Einzahlungen sich die Fonds nicht ebenso gut weiterentwickeln müssen. Bei diesen Fonds liegt die wertgewichtete Rendite deutlich unter der zeitgewichteten.

Die Datenanbieter entwickeln zunehmend Methoden, um den genannten Verzerrungen bei der Berechnung von Marktindizes entgegenzuwirken. Ein anderer Ansatz zur Bewertung der Entwicklung des Hedgefonds-Gesamtmarktes besteht in der Auswertung der Renditen von Dach-Hedgefonds, also Fonds, die ihr Kapital in Hedgefonds investieren. Deren Renditen spiegeln reale Anlageergebnisse der Fonds wider.

Hinsichtlich der statistischen Eigenschaften von Hedgefondsrenditen sind allgemeine Aussagen kaum sinnvoll, da sie vom Betätigungsfeld und der Anlagestrategie des einzelnen Hedgefonds abhängen. Schon zwischen den Strategieguppen gibt es große Unterschiede. Getmansky et al. (2004) untersuchen die Daten des Anbieters TASS von November 1977 bis Januar 2001 (insgesamt 909 einzelne Fonds) in ihrer Gesamtheit und aufgeteilt auf einzelne Strategieguppen.

Parameter	Gesamtmittelwert	Minimum	Maximum
$\hat{\mu}_{annual}$	13,72 % ( $\sigma = 9,46$ %)	0,33 % (10,26 %) Dedicated Shortseller	22,53% (10,80 %) US Equity Hedge
$\hat{\sigma}_{annual}$	16,06 % (12,21 %)	8,19 % (5,05 %) Nondirectional	26,60% (14,44 %) Pure Emerging Markets
Schiefe	-0,15 (1,38)	-0,94 (1,29) Fixed Income Directional	0,53 (1,03) Global Macro
Kurtosis	8,05 (7,56)	4,25 (1,49) Natural Resources	11,26 (11,68) Event Driven
$\hat{\rho}_{t,t-1}$	0,121 (0,168)	-0,001 (0,129) Pure Managed Futures	0,225 (0,123) Convertible Fund (Long)

Durchschnittswerte und Standardabweichungen für verschiedene Verteilungsparameter von Hedgefondsrenditen aus der TASS-Datenbank, Nov. 1977 - Jan. 2001 aus Getmansky et al. (2004). Es wurden nur Fonds mit mindestens fünf Jahren ununterbrochener Renditehistorie einbezogen. Die untersuchten Parameter sind zeilenweise angeordnet. In der ersten Datenspalte steht der Mittelwert der Gesamtheit der 909 Fonds. In der zweiten und dritten Spalte sind die Mittelwerte für diejenige Strategieguppe aufgeführt, die den kleinsten bzw. größten Mittelwert für den betreffenden Parameter aufweist; dabei wurden zwei nur aus jeweils einem Fonds bestehende Strategieguppen nicht berücksichtigt. Die empirischen Standardabweichungen der Werte stehen jeweils in Klammern.

Tabelle 2.3: Kennzahlen für Hedgefondsmonatsrenditen nach Getmansky et al. (2004)

Die Ergebnisse in Tabelle 2.3 weisen für den Hedgefondsgesamtmarkt eine außerordentlich hohe Jahresrendite von 13,7 % aus. Durch die Beschränkung auf Fonds mit mindestens fünfjähriger ununterbrochener Renditehistorie ist der Survivorship Bias in den Daten trotz der Einbeziehung nicht überlebender Fonds nicht ausgeschlossen. Die durchschnittlichen Renditen unterscheiden sich deutlich zwischen den Strategieguppen und sind auch innerhalb der Strategieguppen sehr verschieden, wie man an den Standardabweichungen innerhalb der Gruppen (in der Tabelle in Klammern gesetzt) sieht. Für die Standardabweichungen der Hedgefondsrenditen lässt sich dasselbe sagen. Die Schiefe ist über alle Strategien hinweg kaum ausgeprägt, die Kurtosis ist dagegen gegenüber der Normalverteilung und auch den Investmentfondsdaten aus Abschnitt 2.3.2 deutlich erhöht. Über fast alle Strategieguppen hinweg lässt sich eine positive Autokorrelation 1. Ordnung feststellen, die jedoch nur in drei von sechzehn Gruppen (knapp) über 0,2 liegt.<sup>82</sup> Brooks und Kat (2002) untersuchen durch verschiedene Datenanbieter zur Verfügung gestellte Hedgefondsindizes und finden neben den in Getmansky et al. (2004) beschriebenen Eigenschaften deutliche Unterschiede zwischen den jeweiligen Indizes verschiedener Anbieter, die dieselbe Hedgefonds-Strategieguppe abbilden sollen.

Schließlich noch ein Blick auf die Auswirkungen der Finanzkrise 2008 auf die Hedgefondsbranche: Hinsichtlich der sichtbaren Auswirkungen war das Jahr 2008 zwar das Jahr mit den höchsten Verlusten der Hedgefondsbranche seit der planmäßigen Sammlung ihrer Renditedaten, also etwa 1977. Ibbotson et al. (2011) beziffern, wiederum aufgrund von Daten der TASS-Datenbank, die Gesamtrendite der Hedgefondsbranche im Jahr 2008 auf  $-16,7\%$ . Erstmals in dieser Zeit sank das von Hedgefonds verwaltete Vermögen, auch durch erstmals auftretende Nettoabflüsse von Kundengeldern.<sup>83</sup> Im Vergleich zu anderen Anlageformen, z.B. Aktien, schnitten Hedgefonds nicht auffällig schlecht ab. Ibbotson et al. (2011) führen die Verluste im Jahr 2008 in erster Linie auf Auswirkungen der Marktentwicklung („*systematic beta*“) zurück. Auch sind spektakuläre Insolvenzen von Hedgefonds mit systemischen Auswirkungen wie in der Krise um den Zusammenbruch des Hedgefonds LTCM 1998 ausgeblieben.<sup>84</sup> Es ist aber nicht

---

<sup>82</sup>Weitere, teilweise aktuellere Auswertungen zu Verteilungsparametern von Hedgefondsrenditen findet man in Posthuma und van der Sluis (2003), Malkiel und Saha (2005) und Lo (2008, S. 44 ff.), die alle ebenfalls auf Basis der TASS-Datenbank berechnet wurden und ihn ihren Ergebnissen kaum abweichen. Fung et al. (2002) nutzen Daten von MAR, ihre (obigen Aussagen nicht widersprechenden) Ergebnisse beruhen jedoch auf lediglich 115 Fondsdatenreihen. Die Forschung zu Hedgefonds und deren Renditen erlebte um das Jahr 2000 herum einen starken Aufschwung. Eine Übersicht über die damals veröffentlichte Literatur findet man in Géhin (2004). Eine bündige Zusammenfassung wichtiger Forschungsergebnisse liefert Kat (2003). Einen qualitativen Blick auf den Bereich der Vermögensverwaltung für institutionelle Anleger und Großinvestoren ermöglicht Drobny (2010), das dreizehn ausführliche Interviews mit meist anonymen Vermögensverwaltern und anderen Brancheninsidern enthält.

<sup>83</sup>siehe Kaiser (2009, S. 14-16)

<sup>84</sup>Der in der Krise 2008 zusammengebrochene Madoff-Fonds war ein auf Betrug basierendes Schneeballsystem, das nach außen hin nach Art eines Hedgefonds auftrat. Die Finanzkrise führte nicht durch Vermögensverluste dieses Fonds, sondern durch die starken Mittelabflüsse aus der Hedgefondsbranche zum Platzen dieses Systems.

klar, ob die Verluste der Hedgefondsbranche vollständig in den Datenbanken abgebildet werden. Xu et al. (2010) stellen für die Jahre 2007 und insbesondere 2008 anhand der CISDM-Datenbank eine außergewöhnlich starke Zunahme der Zahl der Berichterstattung abbrechender Fonds fest. Während in den Jahren 1997 bis 2006 stets zwischen 10,1 % und 13,7 % aller Fonds, die im jeweiligen Jahr Renditen berichten, ihre Berichterstattung abbrechen, sind es 2007 14,7 % und 2008 31 %. Es ist also gut möglich, dass viele Fonds verheerende Anlageergebnisse erst gar nicht berichtet haben.

Die für Anlageinteressenten und Wissenschaft zugänglichen Hedgefondsrenditen sind also mit einer ganzen Reihe von Problemen behaftet. Will man ein Verteilungsmodell für die Rendite eines Hedgefonds auswählen, kommen im Vergleich zu offenen Investmentfonds noch einige weitere Punkte hinzu. Bei offenen Investmentfonds ist die Anlagestrategie des Fonds in wesentlichen Teilen vorgegeben, öffentlich transparent und kann nicht kurzfristig grundlegend verändert werden. Ein geeignetes Verteilungsmodell für einen bestimmten Fonds kann anhand seines Anlagespektrums, für vergleichbare Fonds geeigneten Renditemodellen und ggf. bereits beobachteten Renditen des Fonds sinnvoll bestimmt und für die Modellierung seiner zukünftigen Renditeeigenschaften verwendet werden. Bei Hedgefonds ist das nicht so einfach möglich. Erstens hat jeder Hedgefonds eine individuelle Investmentstrategie, so dass Vergleiche mit anderen Fonds nicht ohne weiteres möglich sind. Zweitens haben Außenstehende keinen Einblick in die Anlagestrategie und die Transaktionen eines Hedgefonds und können das mit der Strategie verbundene Risikoprofil nicht vollständig erfassen. Möglicherweise haben auch die Fondsmanager selbst nicht zu jedem Zeitpunkt eine adäquate Vorstellung über die vollständige Risikoposition des Gesamtfonds. Drittens ist das Fondsmanagement in der Lage, den Kapitaleinsatz und die Risikopositionen schnell stark zu verändern. Dadurch kann es sein, dass ein Modell für die Rendite des Fondsportfolios, das für einen bestimmten Zeitpunkt geeignet ist, nach starken Portfolioumschichtungen ungeeignet wird.

In diesem Kapitel wurden die im weiteren Verlauf der Arbeit verwendeten Renditemodelle vorgestellt und die Eigenschaften verschiedener Arten von Finanzmarktrenditen diskutiert. Für Monatsrenditen von offenen Publikumsfonds erscheint die Annahme unabhängig und identisch normalverteilter Renditen vertretbar. Es kann sich aber immer noch lohnen, starke Flanken (über die  $t$ -Verteilung) oder starke Flanken und abhängige Varianzen (über GARCH-Prozesse) zu modellieren. Bei feineren Renditefrequenzen treten diese Phänomene verstärkt auf und sollten berücksichtigt werden. Hedgefondsrenditen sind mit spezifischen Problemen bei der Datenbeschaffung und -bereinigung verbunden und können sehr unterschiedliche Eigenschaften aufweisen, darunter auch Autokorrelation.



## Kapitel 3

# Performancemessung

Ziel der Performancemessung ist es, die für Kapitalgeber relevanten Eigenschaften der (gemessenen oder erwarteten) Wertentwicklungen von Vermögensgegenständen oder Portfolios so zu Kennzahlen zu verdichten, dass Aussagen zur Vorteilhaftigkeit von Anlagegegenständen bzw. zur optimalen Auswahl aus mehreren dieser möglich werden. Die Kennzahlen werden Performancemaße genannt.

Performancemessung soll bei folgenden Fragestellungen Entscheidungen unterstützen:

- Soll man eine bestimmte Vermögensanlage kaufen, nicht kaufen, abstoßen oder sogar leerverkaufen?
- Welche Vermögensanlage aus einer Menge möglicher Anlagen ist am vorteilhaftesten?
- Hat ein Vermögensverwalter eine gute oder schwache Leistung im Sinne der Kapitalgeber erbracht?

Für die Bewertung der Vorteilhaftigkeit für die Kapitalgeber muss zunächst die Frage beantwortet werden, was diese unter „Vorteilhaftigkeit“ bzw. „Nutzen“ verstehen, was im ersten Abschnitt dieses Kapitels kurz erläutert wird. Der zweite Abschnitt behandelt zunächst Performancemaße im Allgemeinen. Anschließend werden jene Performancemaße vorgestellt, für die in dieser Arbeit Inferenzstatistik betrieben wird. In den beiden darauffolgenden Abschnitten werden die Punktschätzung dieser Performancemaße erläutert und einige praktische Probleme bei der Berechnung von Performancemaßen thematisiert. Im letzten Abschnitt folgt ein Überblick über Studien zum Vergleich von Performancemaßen.

### 3.1 Rendite, Risiko und Nutzen

Im Rahmen der Performancemessung wird der Nutzen einer Vermögensanlage allein auf die statistischen Eigenschaften ihrer Wertentwicklung, ggf. im Zusammenhang mit der Wertentwicklung anderer Vermögensgegenstände, reduziert.

Darüber hinausgehende Eigenschaften der Anlageobjekte bleiben unberücksichtigt. Darunter fallen emotionale, gesellschaftspolitische und ethische Präferenzen der Kapitalgeber. Beispiele hierfür sind die Wertschätzung und Förderung bestimmter Regionen, Branchen und Produktinnovationen, die Einhaltung ethischer Standards wie Menschen-, Bürger- und Arbeitnehmerrechte sowie öffentliche Güter wie Umweltaspekte oder Finanzmarktstabilität.<sup>1</sup> Ethische, weltanschauliche und religiöse Grundsätze können bestimmte Anlageobjekte vollständig ausschließen, sowohl hinsichtlich einzelner Branchen und Unternehmen (z.B. Biotechnologie, Alkohol, Waffen) als auch Anlageklassen (Zinsverbot im Islam). Prinzipiell wäre es möglich, derartige Präferenzen in die Performancemessung einzubeziehen, z.B. indem man nichtmonetäre Nutzenarten in die Berechnung der Wertentwicklung miteinbezieht. Eine einfachere Möglichkeit besteht darin, anhand nichtmonetärer Kriterien eine Vorauswahl potentieller Anlageobjekte zu treffen und jene mittels Performancemessung zu analysieren. Eine solche Vorauswahl bieten Anbieter ethisch bzw. religiös orientierter Investments bzw. Investmentfonds an. Andererseits kann eine verbreitete Ausrichtung der Anlagepolitik an Performancekennzahlen zu Rückwirkungen auf die Realwirtschaft führen, wenn dadurch unternehmerischen Entscheidern implizit vorgegeben wird, sich ausschließlich an der Verbesserung von Wertentwicklungskennzahlen auszurichten.

Der neben der Rendite und dem Risiko dritte wichtige Aspekt der Vermögensanlage, die Liquidität des Vermögens, wird in der Performancemessung nicht berücksichtigt. Es wird davon ausgegangen, dass das in Rede stehende Anlagevermögen für die betrachtete Anlageperiode nicht benötigt wird oder im Ernstfall zügig über die Börse in liquide Mittel umgewandelt werden kann. Auch steuerliche Aspekte werden hier ausgeblendet. Wenn die Erträge verschiedener Anlageobjekte steuerlich unterschiedlich behandelt werden, sollte das in der Kalkulation der Renditen berücksichtigt werden.

Gehen wir im Folgenden davon aus, dass sich die Präferenzen eines Kapitalanlegers hinsichtlich Anlageobjekten vollständig durch deren Wertentwicklungseigenschaften abbilden lassen.

Der einfachste Fall der Performancemessung ist derjenige, dass die Wertentwicklung aller Anlageobjekte deterministisch ist. Nimmt man dann an, dass der Nutzen  $U$  des Kapitalanlegers streng monoton mit dem Anlageergebnis  $x$  wächst ( $dU/dx > 0$ ), so wird er diejenige Anlagemöglichkeit mit der höchsten Rendite auswählen.

Sind die Renditen risikobehaftet, muss das Nutzenkalkül des jeweiligen Anlegers genauer betrachtet werden. Ist der Nutzen  $U$  eine lineare Funktion

$$U(X) = b + a \cdot X$$

der Wertentwicklung  $X$ , ist der Anleger risikoneutral: Eine Auszahlungsveränderung um  $k$  Geldeinheiten führt unabhängig vom bereits erreichten Nutzenniveau

---

<sup>1</sup>Man denke an bestimmte Hedgefonds-Kategorien, deren Handelsstrategien destabilisierende Wirkungen auf Finanzmärkte haben können.



$U_0$  stets zu einer Nutzenänderung von  $a \cdot k$ . Die Entscheidung zwischen verschiedenen risikobehafteten Alternativen kann alleine anhand des Erwartungswerts der Auszahlungen getroffen werden. Der Anleger ist zwischen einer sicheren Auszahlung  $k$  und einer risikobehafteten Auszahlung  $X$  mit  $E(X) = k$  indifferent. In diesem Fall würde die (erwartete) Rendite einer Anlage also zur Charakterisierung der Performance einer Anlage ausreichen.

In der Tat ist die erwartete oder beobachtete Rendite einer Anlage ein verbreitetes Performancemaß. Sie ist – bis auf die Dauer der einbezogenen Anlageperiode<sup>2</sup> – fest definiert, ist verhältnismäßig einfach zu berechnen und hat (mit „% p.a.“) eine einfach interpretierbare Einheit.

Risikoneutralität dürfte für die Mehrheit der Anleger keine geeignete Verhaltensannahme darstellen. Eine mehr oder weniger ausgeprägte Risikoaversion lässt sich aus der Theorie des von Neumann-Morgenstern begründen. Nimmt der Nutzenzuwachs durch eine zusätzliche Geldeinheit mit steigendem Wohlstand ab, ist die Nutzenfunktion konkav ( $d^2U/d^2x < 0$ ). Dies impliziert für unsichere Auszahlungen  $X$

$$U(E(X)) = U\left(\int xf(x)dx\right) > \int U(x)f(x)dx = E(U(X)).$$

Man kann zum Vergleich nun diejenige sichere Zahlung heranziehen, die dem Entscheider denselben Nutzen bringt wie die unsichere Zahlung  $X$ . Diese heißt Sicherheitsäquivalent von  $X$ , kurz:  $SE_X$ .<sup>3</sup>

$$U(SE_X) = E(U(X))$$

Ist der Entscheider risikoavers, dann ist das Sicherheitsäquivalent von  $X$  kleiner als der Erwartungswert von  $X$ . Die Differenz zwischen den beiden Werten ist die Risikoprämie  $RP_X$ :

$$E(X) = SE_X + RP_X$$

Die Höhe von  $RP_X$  ist sowohl von der Verteilung von  $X$  als auch vom genauen Verlauf der Nutzenfunktion  $U(X)$  abhängig und soll hier keine weitere Rolle spielen. Wichtig ist die Erkenntnis, dass das Risiko einer Anlage deren Nutzen schmälert und eine Berücksichtigung des Risikos von Anlagen bei der Performancemessung sinnvoll erscheint.

Für die Umsetzung des Begriffs „Risiko“ in quantitative Konzepte gibt es mehrere Ansätze. Die zwei bekanntesten sind:

- a) *Risiko als Unsicherheit*: Als Risiko wird die Streuung der Auszahlung  $X$  bewertet. Die Lagekomponente von  $X$  geht nicht in die Risikobewertung ein.
- b) *Risiko als Drohverlust*: Als Risiko wird die Wahrscheinlichkeit (oder Gefahr) angesehen, dass die Auszahlung  $X$  ein bestimmtes Niveau (z.B. den Anlagebetrag) unterschreitet (engl. *Downside Risk*).

<sup>2</sup>Dieser Parameter ist bei der Ex-post-Performanceberichterstattung nicht zu vernachlässigen. Interessengeleitete Berichterstatter können einen Berichtszeitraum wählen, für den sich eine besonders hohe Durchschnittsrendite ergibt.

<sup>3</sup>vgl. Eisenführ et al. (2010, S. 261 ff.)

Betrachten wir als Beispiel die Anlagen  $A$  und  $B$ , die keine Einzahlungen beinhalten und als Auszahlungen  $X_A$  bzw.  $X_B$  liefern. Hierbei seien

$$X_A = -10 \quad \text{und} \quad X_B = \begin{cases} 0 & \text{mit } p = 50\% \\ 10 & \text{mit } p = 50\% \end{cases} .$$

Bewertet man das Risiko der beiden Anlagen mit dem Kriterium der Unsicherheit, so beinhaltet  $A$  kein Risiko, da die Auszahlung von  $-10$  sicher ist.  $X_B$  hingegen ist unsicher,  $B$  ist demnach risikobehaftet. Die Verlustträchtigkeit der Anlage  $A$  wird separat über die Lagekomponente  $E(X) = -10$  abgebildet.

Definiert man hingegen Risiko als drohenden Verlust, so ist  $A$  risikoreich, weil hier (mit Wahrscheinlichkeit 1) ein Verlust droht. Da  $X_B \geq 0$  gilt, ist hier kein Verlust möglich, Anlage  $B$  also risikolos.

Hinsichtlich erwünschter Eigenschaften von Risikomaßen gibt es dementsprechend zwei unterschiedliche Axiomensysteme. Drohverlust-Risikomaße wurden von Artzner et al. (1999) axiomatisiert, Unsicherheits-Risikomaße von Rockafellar et al. (2002)<sup>4</sup>. Beide Systeme haben folgende Axiome gemeinsam, wobei von einem reellwertigen Risikomaß  $R(X)$  ausgegangen wird:

- *Positive Homogenität:* Investiert man in ein Vielfaches der Anlage  $X$ , so vervielfacht sich das Risikomaß in gleicher Weise.

$$R(a \cdot X) = a \cdot R(X)$$

für  $a \geq 0$  (Artzner et al.) bzw.  $a > 0$  (Rockafellar et al.).

- *Subadditivität:* Das zusammengefasste Risiko zweier Anlagen  $X_1$  und  $X_2$  darf nicht größer sein als die Summe der Einzelrisiken.

$$R(X_1 + X_2) \leq R(X_1) + R(X_2)$$

Rockafellar et al. stellen darüber hinaus folgende Anforderungen:

- *Nichtnegativität:*  $R(X) = 0$ , wenn  $X$  eine Konstante ist, sonst  $R(X) > 0$ .
- *Lageinvarianz:* Wenn zu  $X$  eine Konstante  $a \in \mathbb{R}$  addiert wird, ändert sich das Risikomaß nicht.

$$R(X + a) = R(X)$$

Artzner et al. fordern hingegen:

- *Monotonie:* Ist in jedem möglichen Zustand die Auszahlung von  $X_1$  mindestens so groß wie die Auszahlung von  $X_2$ , dann ist das Risiko von  $X_2$  mindestens so groß wie das von  $X_1$ .

$$X_1 \geq X_2 \quad \Rightarrow \quad R(X_1) \leq R(X_2)$$

---

<sup>4</sup>Dieses Axiomensystem baut auf Vorarbeiten von Pedersen und Satchell (1998) und Kijima und Ohnishi (1993) auf. Eine Zusammenfassung verschiedener Axiomensysteme für Risikomaße bietet Albrecht (2003).

- *Translationsäquivalenz*: Wenn zu  $X$  eine Konstante  $a \in \mathbb{R}$  addiert wird, wird  $a$  vom Risikomaß subtrahiert.

$$R(X + a) = R(X) - a$$

Man erkennt, dass sich Lage- und Translationsinvarianz gegenseitig ausschließen, die beiden Axiomensysteme also nicht kompatibel sind.

Auch Rendite und Risiko, erfasst durch Mittelwert und Standardabweichung, bilden das Erwartungsnutzenkalkül nur unter bestimmten Annahmen an die Nutzenfunktion oder die Renditeverteilung vollständig ab. Hinreichend dafür sind:<sup>5</sup>

- eine quadratische Nutzenfunktion  $U(X) = aX^2 + bX + c$  des Entscheiders oder
- die Annahme gemeinsam normalverteilter Renditen oder, allgemeiner,
- dass die Renditeverteilungen aller Anlageobjekte derselben Lage-Skalen-Familie angehören.<sup>6</sup>

Während an der Annahme einer quadratischen Nutzenfunktion Kritik geübt wird<sup>7</sup>, ist die Annahme, dass die verschiedenen Renditen einer Lage-Skalen-Familie angehören, zumindest für den Vergleich von Anlageobjekten aus einer gemeinsamen Investmentklasse eher akzeptabel.

Es stellt sich die Frage, ob höhere Momente der Renditeverteilungen ebenfalls bei der Entscheidungsfindung berücksichtigt werden sollten, da es durchaus Phänomene gibt, die mit dem Rendite-Risiko-Kalkül nicht vereinbar sind. Man denke an die ungebrochene Beliebtheit von Glücksspielen wie dem Lotto, wo der Spieler eine sichere Geldsumme investiert, um dafür eine äußerst unsichere Auszahlung zu erhalten, deren Erwartungswert deutlich unterhalb des Einsatzes liegt. Offenbar wird einer positiven Schiefe der Auszahlungsverteilung ein positiver Nutzen zugewiesen. Dies ist theoretisch begründbar: Nehmen wir an, dass die Nutzenfunktion eines Investors monoton steigend mit abnehmenden Grenznutzen sei und der Nutzenbewertung höherer Momente bis zur Ordnung  $k$  ein eindeutiges Vorzeichen zugewiesen werden kann. Für diesen Fall weisen Scott und Horvath (1980) nach, dass dann die Nutzenbewertung des Investors für höhere zentrale Momente ungerader Ordnung  $k$  (z.B. Schiefe,  $k = 3$ ) positiv und für höhere zentrale Momente gerader Ordnung  $k$  (z.B. Kurtosis,  $k = 4$ ) negativ ist.

---

<sup>5</sup>vgl. Meyer und Rasche (1992)

<sup>6</sup>vgl. Sinn (1989) und Meyer (1987)

<sup>7</sup>vgl. Eisenführ et al. (2010, S. 266 f.)

## 3.2 Performancemaße

### 3.2.1 Überblick und Klassifikation

Unter Performancemaßen wollen wir hier also Maßzahlen verstehen, die Rendite und Risiko in einer für Entscheidungen hilfreichen Art zu einer Zahl verdichten.

Die ersten Performancemaße wurden in den sechziger Jahren des letzten Jahrhunderts entwickelt, als die Portfoliotheorie zunehmend Verbreitung fand. Damals entstanden die drei 'klassischen', heute allgemein verbreiteten Performancemaße *Sharpe-Ratio*, *Treynor-Ratio* und *Jensens Alpha*. In dem seitdem vergangenen halben Jahrhundert wurde eine schwer überschaubare Menge verschiedenster weiterer Performancemaße vorgestellt. Entsprechend gaben Cogneau und Hübner einem 2009 erschienenen zweiteiligen Übersichtsartikel den Namen „*The (more than) 100 Ways to Measure Portfolio Performance*“.<sup>8</sup>

Ausgangspunkt der Performanceberechnung ist zunächst die Rendite  $R_X$  bzw. deren Erwartungswert  $\mu_{R_X}$ . Sie ist der Mehrertrag gegenüber der Hortung von Geld in Form von Bargeld oder auf einem unverzinsten Bankkonto. In den meisten Fällen werden Anleger jedoch eine verzinsliche Anlageform finden, die sie als mindestens genauso sicher empfinden wie ein Bankkonto oder die Bargeldhortung. Der sichere Zins  $R_f$  dieser alternativen, risikolosen Anlage wird deshalb anstelle der Nullrendite der eigentliche Vergleichsmaßstab für die Bewertung einer risikobehafteten Anlagemöglichkeit sein. Deshalb wird in der Performancemessung in der Regel die Überrendite  $\mu_X = \mu_{R_X} - R_f$ , die Differenz zwischen der erwarteten (Roh-)Rendite  $\mu_{R_X}$  einer Anlagemöglichkeit und dem sicheren Zins  $R_f$ , als Renditemaß verwendet.<sup>9</sup>

Die in dieser Arbeit eingehender beschriebenen Performancemaße kann man in zwei Typen einteilen:

- **Typ I** (Quotientenmaße): Die Überrendite  $\mu_X$  des Portfolios wird durch ein Risikomaß dividiert.

$$\text{Performancemaß Typ I} = \frac{\mu_X}{\text{Risikomaß}}$$

Performancemaße von Typ I haben die Benennung „Renditeeinheit je Risikoeinheit“. Rendite und Risiko werden in der Regel nicht in derselben Einheit gemessen oder unterschiedlich zeitlich skaliert. Deshalb ist der Zahlenwert des Performancemaßes nicht leicht interpretierbar – oft nur im Vergleich von Portfolios untereinander – und kann beim Wechsel auf eine andere Zeiteinheit nicht immer linear umgerechnet werden (siehe dazu Abschnitt 3.4.3).

Da das Risikomaß in der Regel strikt positiv ist, hängt das Vorzeichen des Performancemaßes vom Vorzeichen der Rendite ab. Ist die Rendite

<sup>8</sup>Cogneau und Hübner (2009a) und Cogneau und Hübner (2009b) versuchen, möglichst viele Ansätze zu erwähnen und kurz einzuordnen. Der etwas strukturiertere Überblick von Le Sourd (2007) beschränkt sich auf gut 30 Maße.

<sup>9</sup>Zur Problematik der risikolosen Anlage siehe Abschnitt 3.4.1

und damit das Vorzeichen positiv, gibt es keine weiteren Interpretationsprobleme. Ist die Rendite negativ, ist der Wert des Performancemaßes um so höher (= weniger negativ), je höher der Wert des Risikomaßes ist. Diese Eigenschaft entspricht zwar manchmal der Logik eines bestimmten Performancemaßes, ist aber für sich genommen problematisch: Bei einer negativen Rendite wirkt sich eine Risikoerhöhung positiv auf die Performancebewertung aus. Insofern gibt es an der Stelle  $Performancemaß = 0$  einen Bruch in der Interpretation.

- **Typ II** (Prozentmaße): Von der ggf. um einen dimensionslosen Faktor  $a$  adjustierten Überrendite wird ein Term  $b$  abgezogen, der dieselbe Einheit wie die Überrendite hat.

$$\text{Performancemaß Typ II} = \mu_X \cdot a - b$$

Performancemaße von Typ II stellen (hypothetische) Renditen dar und werden in aller Regel in der Einheit Prozent ausgedrückt bzw. sind (bei Verwendung stetiger Renditen) einfach in Prozente umrechenbar.

Durch die Terme  $a$  und  $b$  sollen die Renditen verschiedener Portfolios vergleichbar gemacht werden. Die Zahlenwerte sind als Renditen oder Renditedifferenzen einfach interpretierbar und es gibt keinen „Bruch“ zwischen der Interpretation von positiven und negativen Zahlenwerten.

Im weiteren Verlauf dieses Abschnitts werden zunächst die drei „klassischen“ Performancemaße behandelt, die in den 1960er Jahren veröffentlicht wurden. Nicht zuletzt durch ihre Rolle in Portfoliotheorie und Capital Asset Pricing Model (CAPM) sind sie bis heute die bekanntesten und am meisten verwendeten Performancemaße. Verstärkt seit den 1990er Jahren wurde eine große Zahl neuer Ansätze zur Performancemessung veröffentlicht. Hierzu zählen die Risk-Adjusted Performance von Franco und Leah Modigliani, das eher in der Tradition der klassischen Maße steht, sowie eine Reihe von Maßen, die das Risiko einer Anlage explizit als Verlustrisiko modellieren.

## 3.2.2 Die klassischen Performancemaße

### 3.2.2.1 Das Sharpe-Ratio

Auflistungen von Performancemaßen beginnen in der Regel mit dem Sharpe-Ratio, veröffentlicht von William Sharpe im Jahr 1966.<sup>10</sup> Es ist zwar nicht das älteste Performancemaß, heute aber das in der Praxis wohl am häufigsten berichtete. Selbst in der allgemeinen Wirtschaftspresse findet man Fondsvergleiche anhand des Sharpe-Ratios.

Das Sharpe-Ratio eines Portfolios  $X$  ist der Quotient aus dessen Überrendite und der Standardabweichung der Rendite:

$$Sh_X = \frac{\mu_{R_X} - R_f}{\sigma_X} = \frac{\mu_X}{\sigma_X}$$

<sup>10</sup>Sharpe (1966), eine Zusammenfassung und Diskussion durch den Erfinder des Maßes ist in Sharpe (1994) zu finden.

Steht eine risikobehaftete Anlage  $X$  zur Verfügung, dann wird der Anleger nur dann darin investieren, wenn er eine positive Überrendite als Ausgleich für das einzugehende Risiko erwarten kann. Das Sharpe-Ratio drückt nun die Überrendite je eingegangener Risikoeinheit aus.

Kennt man das Sharpe-Ratio mehrerer Anlagen, sind Vergleichsaussagen möglich. Angenommen, die Portfolios  $X$  und  $Y$  haben Sharpe-Ratios von  $Sh_X = 0,5$  und  $Sh_Y = 0,4$ . Wenn man nun soviel in Portfolio  $X$  investiert, dass man dasselbe Risiko eingeht wie bei einer Investition einer bestimmten Summe in Portfolio  $Y$ , kann man die  $\frac{0,5}{0,4} = 1,25$ -fache erwartete Rendite der Investition in  $Y$  erhoffen. Insofern ist es für rationale Anleger vorteilhafter, in  $X$  statt in  $Y$  zu investieren.

Das Sharpe-Ratio hat als Performancemaß vom Typ I bei negativen Überrenditen die Eigenschaft, risikoarme Portfolios schlechter zu bewerten als risikoreiche Portfolios mit gleicher negativer Überrendite. Wenn es um eine Anlageentscheidung geht (*ex-ante*-Fall) und es die Möglichkeit zu Leerverkäufen (Short Selling) gibt, lässt sich das sinnvoll interpretieren. Bei einem risikobehafteten Portfolio mit negativer erwarteter Überrendite kann man eine positive Rendite erzielen, wenn man es leerverkauft. Will man eine bestimmte Risikogrenze einhalten, sollte man jenes Portfolio leerverkaufen, welches je Risikoeinheit die am stärksten negative Überrendite – bzw. das niedrigste Risiko bei einer bestimmten negativen Überrendite – und somit das niedrigste Sharpe-Ratio aufweist.<sup>11</sup> Bei der nachträglichen Bewertung der Wertentwicklung eines Portfolios (*ex-post*-Fall) ist es jedoch nicht ersichtlich, weshalb man bei schlechter Wertentwicklung hohe eingegangene Risiken positiv bewerten soll.

Annahmegemäß richtet sich das Nutzenkalkül eines rationalen Anlegers an der Wertentwicklung und dem Risiko seines gesamten Vermögens aus. Wenn nur ein Teil des Vermögens neu angelegt werden soll, reicht es nicht aus, nur das Sharpe-Ratio des potentiellen Anlagegegenstandes zu betrachten, denn das Gesamtrisiko des Vermögens hängt auch von der Korrelation der Wertentwicklung des Anlageobjekts mit der des sonstigen Vermögens ab. Um das Sharpe-Ratio für solche Anlageentscheidungen sinnvoll anwenden zu können, müsste man jeweils Portfolios aus dem sonstigen Vermögen und den einzelnen in Frage kommenden Anlagemöglichkeiten bilden und diese dann anhand des Sharpe-Ratios vergleichen. Insofern sind für Anleger, die nicht ihr gesamtes Vermögen in ein Anlageobjekt investieren wollen, Angaben zu Sharpe-Ratios von Anlagemöglichkeiten, welche alleine aus deren erwarteten Renditen und Standardabweichungen berechnet wurden, nur von eingeschränktem Nutzen.

### 3.2.2.2 Das Treynor-Ratio

Das Treynor-Ratio, veröffentlicht von Jack Treynor (1965), stellt eine Kennzahl dar, die für die Beurteilung einzelner (potentieller) Teilinvestments geeignet ist.

<sup>11</sup>Bei dieser Interpretation ist zu bedenken, dass der Anleger bei Wertpapierkäufen auf Kredit einen höheren Zins zahlen muss, als er für eigene Anlagen erhält, sowie bei Leerverkäufen Gebühren zahlen und/oder Sicherheiten leisten muss und sich somit die Überrendite und entsprechend das Sharpe-Ratio der Anlagestrategie ändert.

Zentraler Gedanke dahinter ist die Möglichkeit zur Diversifizierung bei der Vermögensanlage gemäß der Portfoliotheorie: Das Risiko eines Portfolios mehrerer (nicht exakt positiv korrelierter) Anlageobjekte ist kleiner als die Summe der Einzelrisiken dieser Anlageobjekte. Gewinne und Verluste verschiedener Einzelobjekte gleichen sich – zumindest teilweise – aus. Je mehr verschiedene Anlageobjekte in einem Portfolio enthalten sind, desto stärker wirkt der risikomindernde Diversifikationseffekt. Aber auch nach noch so guter Diversifikation verbleibt ein Restrisiko, das von globalen Faktoren ausgelöst wird, die in unterschiedlichem Maße auf alle Anlageobjekte gleichzeitig einwirken. Dieses Restrisiko wird systematisches Risiko oder Marktrisiko genannt im Gegensatz zum unsystematischen Risiko der einzelnen Anlageobjekte.

Treynor geht nun davon aus, dass der betreffende Anleger über ein perfekt diversifiziertes Anlageportfolio verfügt und nunmehr eine im Verhältnis zum Gesamtportfolio sehr kleine Summe neu investieren will. Hinsichtlich der Rendite ist sein erster Vergleichsmaßstab die risikolose Anlage, so dass wiederum die Überrendite über einen risikolosen Zinssatz als Renditemaß verwendet wird. Das Risiko betreffend ist für ihn das Einzelrisiko der zusätzlichen Anlage nicht relevant. Wichtig ist, in welchem Maße die zusätzliche Anlage auf die Marktbewegung reagiert und somit dem bestehenden Portfolio Marktrisiko hinzufügt. Diese Sensitivität gegenüber dem Marktrisiko lässt sich mittels eines Regressionsansatzes ausdrücken:<sup>12</sup>

$$\begin{aligned} R_X &= (R_f + \alpha_X) + \beta_X \cdot (R_M - R_f) + \epsilon_X \\ \text{bzw. } \mu_{R_X} &= (R_f + \alpha_X) + \beta_X \cdot (\mu_{R_M} - R_f) , \\ \text{kurz: } \mu_X &= \alpha_X + \beta_X \cdot \mu_M \end{aligned} \quad (3.1)$$

Die erwartete Überrendite eines Anlageobjekts  $X$  setzt sich additiv zusammen aus einem Achsenabschnitt  $\alpha_X$  und dem systematischen Risiko  $\beta_X \cdot \mu_M$ , wobei  $\mu_M$  die Überrendite eines geeigneten Global- oder Marktportfolios ist.<sup>13</sup> Der Koeffizient  $\beta_X$  gibt an, um wie viele Einheiten sich die Erwartung der Rendite des Anlageobjekts  $X$  ändert, wenn die Marktrendite um eine Einheit steigt. Entsprechend der Formel für den Regressionskoeffizienten in der linearen Einfachregression von  $X$  auf  $M$  gilt  $\beta_X = \frac{\sigma_{X,M}}{\sigma_M^2}$ . Da das unsystematische Risiko durch die anderen Anlageobjekte im Gesamtportfolio wegdiversifiziert wird und die Marktrendite für den Anleger nicht beeinflussbar ist, ist der Wert des Koeffizienten  $\beta$  für das Anlageobjekt  $X$  das relevante Maß seines Risikos.

Entsprechend lautet das Treynor-Ratio

$$Tr_X = \frac{\mu_{R_X} - R_f}{\beta_X} = \frac{\mu_X}{\frac{\sigma_{X,M}}{\sigma_M^2}} = \frac{\mu_X \sigma_M^2}{\sigma_{X,M}}$$

Ein Anleger sollte aus mehreren Anlageobjekten dasjenige mit dem höchsten (positiven) Treynor-Ratio auswählen, also jenes, das für eine bestimmte Menge

<sup>12</sup>Hierbei ist  $R_M$  eine beobachtete Rendite des Marktportfolios,  $\epsilon_X$  ein Störterm und  $\mu_{R_X}$  der Erwartungswert der Rohrendite des Marktportfolios.  $\sigma_{X,M}$  ist die Kovarianz der Überrenditen von Portfolio  $X$  und dem Marktportfolio  $M$ .

<sup>13</sup>Die Frage nach einem geeigneten Marktportfolio wird in Abschnitt 3.4.2 aufgegriffen.

systematischen Risikos die höchste Überrendite erwirtschaftet. Ist die Überrendite negativ, bieten sich - solange das  $\beta$  positiv ist - analog zum Sharpe-Ratio Leerverkäufe an. Es ist zu beachten, dass das Risikomaß  $\beta$  nicht positiv sein muss.<sup>14</sup> Es gibt Anlageobjekte, etwa Derivate und bestimmte Hedgefonds, die sich mehr oder weniger stark gegenläufig zum Markt bewegen. Anlageobjekte mit positiver Überrendite und negativem  $\beta$  weisen ein negatives Treynor-Ratio auf, obwohl sie ideale Anlageobjekte wären: Sie erwirtschaften mehr als den risikolosen Zinssatz, während sie gleichzeitig das Risiko des Gesamtportfolios mindern. Sind Überrendite und  $\beta$  negativ, ist das Treynor-Ratio positiv. In diesem Fall wäre aber ein niedrigeres Treynor-Ratio vorteilhafter, denn man erhielte eine leicht negative Überrendite bei deutlicher Risikoreduktion.

Das Treynor-Ratio ist anders als das Sharpe-Ratio also ein geeignetes Maß für die Bewertung einzelner Anlageobjekte, nicht jedoch für das Gesamtportfolio eines Investors. Seine Angabe kann stets nur bezogen auf ein bestimmtes Marktportfolio erfolgen. Wenn für ein Anlageobjekt ein Treynor-Ratio angegeben wird, kann ein Investor es für seine Anlageentscheidung übernehmen, wenn er das bei der Berechnung verwendete Marktportfolio als für seine Belange geeignet ansieht.

### 3.2.2.3 Jensens Alpha

Eine weitere Möglichkeit der Performancemessung ist der Vergleich eines Anlageergebnisses mit dem eines passiven Investors. Ein passiver Investor, der sich nicht über die Eigenschaften einzelner Investitionsobjekte informiert und auf eine adäquate Preisfindung durch den Markt hofft, investiert den Anteil  $c \in \mathbb{R}$  seines Kapitals in das Marktportfolio und den Anteil  $1 - c$  in die risikolose Anlage.<sup>15</sup> Sein eingegangenes Risiko, gemessen mit der Standardabweichung seiner Rendite, ist  $\sigma_X = |c| \cdot \sigma_M$  und seine erwartete Rendite beträgt  $R_f + c \cdot \mu_M$ . Den Wert  $c$  wählt er anhand seiner Rendite-Risiko-Präferenz. Will man die Performance eines Portfolios  $X$  mit jener einer passiven Anlage mit gleicher Exposition  $\beta_X$  zum Marktrisiko vergleichen, kann man einfach die Differenz der erwarteten Renditen bilden:

$$\mu_X - \beta_X \mu_M = \alpha_X$$

Dieses Maß  $\alpha_X$  wird Jensens Alpha genannt<sup>16</sup> und ist gleich dem Achsenabschnitt aus der Regressionsgleichung 3.1. Jensens Alpha gehört zu den Performancemaßen vom Typ II und hat als Benennung die Einheit, in der auch die Rendite ausgedrückt wird. Hat ein Portfolio ein Alpha von einem Prozent, so ist seine Rendite um einen Prozentpunkt höher als jene einer passiven Anlage mit demselben Marktrisiko. Hierbei wird ebenso wie beim Treynor-Ratio das  $\beta_X$  des Portfolios als Risikomaß benutzt. Das setzt auch hier voraus, dass der jeweilige

<sup>14</sup> $\beta$  ist nach keinem der beiden oben genannten Axiomensysteme ein kohärentes Risikomaß, da es weder die Bedingung der Nichtnegativität (Rockafellar et al.) noch die der Monotonie (Artzner et al.) erfüllt.

<sup>15</sup>Bei  $c \in \mathbb{R}$  sind der Leerverkauf des Marktportfolios sowie der Kauf des Marktportfolios auf Kredit erlaubt; will man dies nicht zulassen, gilt  $c \in [0, 1]$ .

<sup>16</sup>Nach Michael C. Jensen (1968)



Investor ein im Verhältnis zur Anlage in  $X$  sehr großes, perfekt diversifiziertes Portfolio besitzt.

Für den Fall, dass Jensens Alpha für die Ex-post-Bewertung von Portfoliomanagern eingesetzt wird, ist zu bedenken, dass jene das Beta ihres Portfolios durch ihre Anlageentscheidungen beeinflussen können. Ingersoll et al. (2007) machen auf die Manipulierbarkeit dieser Maßzahl (und anderer Performancemaße) aufmerksam.<sup>17</sup> In entscheidungstheoretischer Hinsicht muss man bedenken, dass Anlagen mit gleichem Alpha zwar dieselbe (erwartete) marktrenditebereinigte Überrendite haben, aber mit dem Alpha keine Aussage zum absoluten oder systematischen Risiko verbunden ist. Deswegen – und wegen des Hebeleffektes von Beta – wird davon abgeraten, Ranglisten von Portfolios mit unterschiedlichen Betas anhand von Jensens Alpha aufzustellen.<sup>18</sup>

### 3.2.3 Neuere Performancemaße

#### 3.2.3.1 Risk-Adjusted Performance nach Modigliani / Modigliani

Den zuletzt angesprochenen Mangel wollen Franco und Leah Modigliani beheben, indem sie ein Typ-II-Performancemaß vorschlagen, das Risikounterschiede zwischen Portfolios berücksichtigt. Es wurde von Modigliani/Modigliani *Risk-Adjusted Performance*, abgekürzt *RAP* genannt. In der Literatur findet man es oft unter der neutraleren Bezeichnung  $M^2$ , nach dem gemeinsamen Nachnamen der beiden Autoren. Um Verwechslungen mit dem Quadrat einer Marktrendite zu vermeiden, wird hier im folgenden das Kürzel *MM* verwendet. Zudem gibt es neben der Version von *MM* gemäß dem Originalartikel noch eine weitere Version (s. u.), die teilweise ohne weiteren Kommentar als *MM* verwendet wird. Zur Unterscheidung nennen wir die Originalversion  $MM_{ori}$  und die alternative Formulierung  $MM_{alt}$ .

$MM_{ori}$  ist für ein Portfolio  $X$  und eine Benchmark  $M$  gemäß Modigliani und Modigliani (1997) wie folgt definiert:

$$MM_{ori} = \frac{\sigma_M}{\sigma_X}(\mu_{R_X} - R_f) + R_f = \frac{\sigma_M}{\sigma_X}\mu_X + R_f$$

Die Überrendite  $\mu_X$  wird also mit dem Verhältnis von Markt- zu Portfoliorisiko skaliert und zur sicheren Rendite addiert. Das Ergebnis kann man als die Rendite interpretieren, die man erhielte, wenn man genau soviel in Anlage  $X$  investieren würde, dass das so gebildete Mischportfolio aus  $X$  und der risikolosen Anlage dasselbe Risiko aufweist wie eine Investition des gesamten Anlagebetrags in das Marktportfolio.

Neben dieser im Vergleich zum Sharpe-Ratio aussagekräftigeren Interpretation hat es bei der Reihung mehrerer Portfolios im Vergleich zu Jensens Alpha den Vorteil, dass alle Portfolios auf dasselbe Risiko skaliert werden. Diese Eigenschaft

<sup>17</sup>Haben Portfoliomanager z.B. in der ersten Hälfte einer Bewertungsperiode ein positives Alpha erzielt, können sie das Alpha zum Ende der Bewertungsperiode durch Senkung des Betas (Desinvestition, Leerverkauf des Marktportfolios) in die Höhe treiben.

<sup>18</sup>vgl. z.B. Wittrock (2000, S. 81)

kann man durch die Beziehung von  $MM_{ori}$  zum Sharpe-Ratio illustrieren<sup>19</sup>:

$$MM_{ori} = \frac{\mu_X \sigma_M}{\sigma_X} + R_f = \sigma_M \cdot Sh_X + R_f$$

$MM_{ori}$  ist also eine lineare Transformation des Sharpe-Ratios. Solange man mehrere Portfolios auf dasselbe Marktportfolio bezieht, ergibt sich eine identische Rangfolge wie beim Sharpe-Ratio, und dies unabhängig davon, welches konkrete Marktportfolio man wählt.

Das Performancemaß  $MM$  taucht in der Literatur manchmal in einer leicht abgewandelten Form auf, die hier mit  $MM_{alt}$  abgekürzt wird<sup>20</sup>:

$$MM_{alt} = \frac{\sigma_M}{\sigma_X} \mu_X - \mu_M = MM_{ori} - \mu_{RM} \quad (3.2)$$

Hierbei wird also von der risikoadjustierten Rendite  $MM_{ori}$  die Rendite des Marktportfolios abgezogen, es handelt sich also lediglich um eine Verschiebung von  $MM_{ori}$  um eine Konstante. Der Vorteil dieser Variante ist, dass man bereits am Vorzeichen von  $MM_{alt}$  sieht, ob sich das betreffende Portfolio risikoadjustiert besser oder schlechter als das Marktportfolio entwickelt. Ist  $MM_{alt}(X)$  positiv, ist die risikoadjustierte Wertentwicklung des Portfolios  $X$  besser als ein passives Investment in das Marktportfolio; für  $MM_{alt}(X) < 0$  ist sie schlechter und man sollte demnach nicht in  $X$  investieren.

Das Performancemaß  $MM$  ist eine Kennzahl aus Rendite und absolutem Risiko der Anlage, das Marktportfolio wird nur zur Skalierung des absoluten Risikos verwendet. Das Anwendungsgebiet von  $MM$  ist also die Bewertung des gesamten Anlageportfolios eines Investors und die bereits bei der Diskussion des Sharpe-Ratios genannten Hinweise gelten entsprechend auch hier. Eine risikoadjustierte Rendite auf Basis des systematischen Risikos haben Wilkens und Scholz (1999) in vergleichbarer Form vorgestellt ( $MRAP = \text{Market Risk-Adjusted Performance}$ ).

### 3.2.3.2 Performancemaße auf Basis partieller Momente

Bei den bisher vorgestellten Performancemaßen wurde Risiko immer als Unsicherheit oder Schwankungsanfälligkeit aufgefasst. Alle Performancemaße enthielten, direkt oder als Bestandteil des Betas, die Varianz bzw. die Standardabweichung der Rendite. Ein gleichwertiges Risikokonzept ist, wie auf S. 45 dargestellt, die Vorstellung des Risikos als Möglichkeit negativer Abweichungen von einem Bezugswert.

Diese Risikokonzeption kann durch *partielle Momente* quantitativ umgesetzt werden. Ein partielles Moment  $PM$  ist für eine stetige Zufallsvariable  $X$  (in allgemeiner Form) wie folgt definiert:

$$PM(a, b, c) = \int_a^b x^c f(x) dx$$

<sup>19</sup>vgl. Fischer (2010, S.463)

<sup>20</sup>z.B. Schmid und Schmidt (2006)

mit  $-\infty \leq a < b \leq \infty$  und  $c \in \mathbb{N}_0$ . Der Unterschied zu einem „normalen“ Moment besteht lediglich darin, dass nicht über ganz  $\mathbb{R}$  integriert wird, sondern nur über ein Teilintervall  $[a, b]$ .

Von *oberen partiellen Momenten* (UPM) bzw. *unteren partiellen Momenten* (LPM)<sup>21</sup> von  $X$  bezüglich eines Parameters  $a \in \mathbb{R}$  spricht man, wenn über ganz  $\mathbb{R}$  rechts bzw. links von  $a$  integriert wird:

$$UPM_X(a, c) = \int_a^{\infty} x^c f(x) dx, \quad LPM_X(a, c) = \int_{-\infty}^a x^c f(x) dx$$

Im Rahmen des Konzepts von Risiko als Drohverlust definieren wir nun einen Verlust als eine Rendite unterhalb eines Wertes  $z \in \mathbb{R}$ . Der Parameter  $z$  kann in der Praxis der risikolose Zinssatz, ein Eigenkapitalkostensatz oder eine andere extern vorgegebene Größe sein. Die Gewinn-/Verlust-Verteilung ist nun um  $z$  zentriert. Die partiellen Momente dieser Verteilung sind:

$$UPM_X(z, c) = \int_z^{\infty} (x - z)^c f(x) dx = E(((X - z)_+)^c)$$

$$LPM_X(z, c) = \int_{-\infty}^z (z - x)^c f(x) dx = E(((z - X)_+)^c),$$

wobei bei dem unteren partiellen Moment  $(z - x)$  anstatt  $(x - z)$  integriert wird, damit dessen Vorzeichen stets positiv ist. Die Darstellung  $E(((\dots)_+)^c)$  bringt zum Ausdruck, dass die Werte innerhalb der inneren Klammer nur dann in die Berechnung des Erwartungswerts eingehen, wenn sie positiv sind. Im Folgenden schreiben wir kürzer  $E((\dots)_+^c)$ .

Die ersten unteren partiellen Momente kann man wie folgt interpretieren:

- $LPM_X(z, 0) = \int_{-\infty}^z f(x) dx$ : Wahrscheinlichkeit eines Verlustes,
- $LPM_X(z, 1) = \int_{-\infty}^z (z - x) f(x) dx$ : erwarteter Verlust,
- $LPM_X(z, 2) = \int_{-\infty}^z (z - x)^2 f(x) dx$ : erwarteter quadrierter Verlust.

Untere partielle Momente können als Risikomaße für die Performancemessung verwendet werden. Das erste Performancemaß dieser Art stellten Sortino und van der Meer (1991) vor, es wird heute **Sortino-Ratio** genannt. Seine Definition lautet:

$$Sor_X(z) = \frac{E(X - z)}{\sqrt{LPM_X(z, 2)}} = \frac{E(X - z)}{\sqrt{\int_{-\infty}^z (z - x)^2 f(x) dx}}$$

<sup>21</sup>UPM steht für *upper partial moment*, LPM für *lower partial moment*

Das Sortino-Ratio ist ein Performancemaß vom Typ I und das Downside-Risk-Äquivalent des Sharpe-Ratios: An der Stelle der Wurzel des zweiten zentralen Moments steht die Wurzel des zweiten partiellen Moments, außerdem erfolgt die Korrektur der erwarteten Rendite  $E(X)$  nicht zwingend durch den risikolosen Zinssatz, sondern bleibt als zu wählender Parameter  $z$  offen. Ist für den Entscheider das Downside Risk das relevante Risikokonzept, ist für ihn das Sortino-Ratio aussagekräftiger als das Sharpe-Ratio.

Ein weiteres Performancemaß auf Basis partieller Momente ist das **Omega** von Keating und Shadwick (2002):

$$\text{Omega} = \frac{E((X - z)_+)}{E((z - X)_+)} = \frac{UPM_X(z, 1)}{LPM_X(z, 1)}$$

Hier steht über dem Bruchstrich der (partielle) Erwartungswert des Gewinns und darunter der (partielle) Erwartungswert des Verlusts, bezogen auf eine Rendite  $z$ . Anders als bei den bisher vorgestellten Performancemaßen ist der Wertebereich von Omega auf  $\mathbb{R} \geq 0$  beschränkt. Ein Wert von 1 bedeutet, das sich erwarteter Gewinn und erwarteter Verlust die Waage halten. Ebenfalls ist auffällig, dass hier erstmals nicht die erwartete Überrendite im Zähler steht, sondern ein anderer Ausdruck für den Anlageerfolg. Das ist aber nur vordergründig der Fall. Wenn  $X$  wie bei den klassischen Performancemaßen als Überrendite interpretiert wird bzw.  $z = R_f$  gilt, lässt sich Omega mittels der Beziehung

$$\begin{aligned} \mu_X &= \int_{-\infty}^{\infty} xf(x) dx = \int_z^{\infty} xf(x) dx + \int_{-\infty}^z xf(x) dx \\ &= E((X - z)_+) - E((z - X)_+) \end{aligned}$$

in einen Ausdruck mit dem Erwartungswert von  $X$  im Zähler umformen:

$$\begin{aligned} \text{Omega} &= \frac{E((X - z)_+)}{E((z - X)_+)} = \frac{\mu_X + E((z - X)_+)}{E((z - X)_+)} \\ &= \frac{\mu_X}{E((z - X)_+)} + 1. \end{aligned}$$

Kaplan und Knowles (2004) haben mit ihrem Maß **Kappa** ein verallgemeinertes Performancemaß auf Basis partieller Momente vorgestellt:

$$\text{Kappa}(\alpha, z) = \frac{E(X - z)}{\sqrt[\alpha]{E((z - X)_+^\alpha)}}$$

Im Zähler steht die um das Anspruchsniveau verminderte erwartete Rendite, im Nenner die  $\alpha$ -te Wurzel des  $\alpha$ -ten partiellen Moments. Der Parameter  $\alpha$  steuert die Gewichtung der als Verlust betrachteten Renditen untereinander. Bei  $\alpha = 1$  werden alle Renditen gleich gewichtet, bei  $\alpha = 2$  werden die Renditen quadriert, also mit ihrem eigenen Wert gewichtet. Stärkere Verluste erhalten so ein höheres Gewicht. Wenn man  $\alpha$  weiter erhöht, steigt die Gewichtung extremer Verluste weiter an.

Das Sortino-Ratio ist leicht als Spezialfall  $\alpha = 2$  von Kappa zu erkennen. Omega ist gleich Kappa ( $\alpha = 1$ ) plus eins. Im weiteren Verlauf der Arbeit wird Kappa als der Repräsentant für Performancemaße auf Basis partieller Momente verwendet.

Schließlich kann man auch noch ein weiteres Maß in die Kategorie der Performancemaße auf Basis partieller Momente einordnen, nämlich den **Mean on Conditional Value-at-Risk (M-CVaR)** nach Agarwal und Naik (2004). Er setzt die Überrendite ins Verhältnis zum Expected Shortfall einer Anlage. Der Expected Shortfall basiert auf einem partiellen Moment einer Zufallsvariablen nach folgender Berechnungsvorschrift:

$$ExS_X(\alpha) = \frac{1}{\alpha} \int_{-\infty}^{x_\alpha} x f(x) dx$$

Integriert wird die untere Flanke einer Verteilung bis zum  $\alpha$ -Quantil  $x_\alpha$  der Verteilung. Der Parameter  $\alpha \in ]0; 1[$  ist extern vorgegeben, meist 0,05 oder kleiner. Das Ergebnis wird mit  $1/\alpha$  multipliziert und kann als der Wert von  $X$  interpretiert werden, der zu erwarten ist, falls  $X$  unterhalb des  $\alpha$ -Quantils der Verteilung liegt. In dieses Risikomaß gehen also nur die  $100 \cdot \alpha$  % extremen Risiken der Anlage ein. Der Expected Shortfall ist das einzige hier vorgestellte Risikomaß, das gemäß den Axiomen für Drohverlust-Risikomaße von Artzner et al. (1999) ein konsistentes Risikomaß ist. Er wird gelegentlich auch *Conditional Value-at-Risk* genannt, obwohl er ein anderes Konzept darstellt als der Value-at-Risk, der im Folgeabschnitt vorgestellt wird. Das Performancemaß M-CVaR<sup>22</sup> ist definiert als

$$\text{M-CVaR} = \frac{\mu_{R_X} - R_f}{\frac{1}{\alpha} \int_{-\infty}^{x_\alpha} x f(x) dx} = \frac{\mu_X}{ExS_X(\alpha)}.$$

### 3.2.3.3 Performancemaße auf Basis des Value-at-Risk

Der Value-at-Risk (VaR) ist ein verbreitet angewendetes Risikomaß in der Finanzwirtschaft und er kann auch in der Performancemessung als Risikomaß dienen. Beim Value-at-Risk, der immer in Bezug auf ein Niveau  $(1 - \alpha)$  angegeben werden muss, handelt es sich um das  $\alpha$ -Quantil der Verteilung einer Zufallsvariable, z.B. einer Portfoliorendite.<sup>23</sup> Dieser zumeist negative Wert wird für gewöhnlich mit einem positiven Vorzeichen versehen und drückt dann das Verlustniveau aus, das mit einer Wahrscheinlichkeit  $(1 - \alpha)$  nicht überschritten wird. Übliche Parameterwerte für  $\alpha$  sind 10 %, 5 % oder weniger. Der 99%-VaR (oder  $VaR_{\alpha=0,01}$ ) ist dementsprechend das 0,01-Quantil der Renditeverteilung und wird zwecks positiven Vorzeichens mit  $-1$  multipliziert. Der Value-at-Risk gibt keine Hinweise auf mögliche extreme Verluste. Ein Vorteil des Value-at-Risks ist, dass für den ex-ante-Fall eine Vielzahl von exakten und näherungsweise Berechnungsmethoden existiert. Will man ex ante den Value-at-Risk für ein aus

<sup>22</sup>In Eling und Schuhmacher (2007) wird es *Conditional Sharpe Ratio* genannt.

<sup>23</sup>Man beachte, dass der Buchstabe  $\alpha$  bei den Maßen Jensens Alpha, Kappa und den Value-at-Risk-basierten Maßen unterschiedliche Bedeutungen hat. In allen Fällen wurde die Notation aus der Literatur übernommen.

verschiedenen Anlagetiteln zusammengesetztes Portfolio ermitteln, stehen z.B. die Delta-Normal- bzw. die Delta-Gamma-Normal-Methode<sup>24</sup> oder verschiedene Simulationsansätze zur Verfügung.<sup>25</sup>

Das zugehörige Performancemaß ist der **Expected Return on Value-at-Risk (ERVaR)**:

$$ERVaR_X(\alpha) = \frac{\mu_X}{VaR_\alpha}$$

Aus den verschiedenen Berechnungsansätzen des Value-at-Risk bzw. bestimmten Verteilungsannahmen für die Portfoliorenditen sind verschiedene Unterformen des ERVaR entstanden. Nimmt man zum Beispiel an, dass die Portfoliorendite normalverteilt ist, was z.B. das Delta-Normal-Modell impliziert, dann gilt (mit  $u_\alpha = \alpha$ -Quantil der Standardnormalverteilung)  $VaR_{\alpha, norm.} = -(\mu_X + u_\alpha \cdot \sigma_X)$  und der ERVaR hat die Form

$$\text{Normal-ERVaR}_X(\alpha) = \frac{\mu_X}{-(\mu_X + u_\alpha \cdot \sigma_X)}$$

Diese Form des Performancemaßes ergibt sich aus einer (ex-ante-)Entscheidungsregel in Dowd (2000) über die Vorteilhaftigkeit einer Investition in ein zusätzliches Wertpapier. Aus der ex-post-Perspektive hat diese Variante des ERVaR jedoch die Eigenschaft, dass sie aufgrund der Normalverteilungsannahme (für  $\alpha < 0,5$ ) lediglich eine Funktion des Sharpe-Ratios ist.

$$\begin{aligned} \text{Normal-ERVaR}_X(\alpha) &= \frac{\mu_X}{-(\mu_X + u_\alpha \sigma_X)} = \frac{\mu_X}{-(\mu_X + u_\alpha \frac{\mu_X \sigma_X}{\mu_X})} \quad (\text{für } \mu_X \neq 0) \\ &= \frac{\mu_X}{-(\mu_X + u_\alpha \frac{\mu_X}{Sh_X})} = -\frac{1}{1 + \frac{u_\alpha}{Sh_X}} \\ &= -\frac{Sh_X}{Sh_X + u_\alpha} = -\frac{Sh_X}{Sh_X - u_{1-\alpha}} = \frac{Sh_X}{u_{1-\alpha} - Sh_X} \end{aligned}$$

Im Bereich  $Sh_X < u_{1-\alpha}$ <sup>26</sup> ist der Normal-ERVaR also eine monoton steigende, nichtlineare Funktion des Sharpe-Ratios unter Beibehaltung des Vorzeichens. Vorteilhaftigkeits- und Rangfolgeentscheidungen sind also nach Sharpe-Ratio und Normal-ERVaR identisch. Es besteht also auch dann, wenn die Annahme normalverteilter Renditen unproblematisch erscheint, kein Grund, das Sharpe-Ratio durch den komplizierteren Normal-ERVaR zu ersetzen.

Favre und Galeano (2002) haben einen *Modified Value-at-Risk*

$$\begin{aligned} MVaR(\alpha) &= \mu_X + q_\alpha^* \cdot \sigma_X \quad \text{mit} \\ q_\alpha^* &= u_\alpha + \frac{u_\alpha^2 - 1}{6} \gamma_{1,X} + \frac{u_\alpha^3 - 3u_\alpha}{24} (\gamma_{2,X} - 3) - \frac{2u_\alpha^3 - 5u_\alpha}{36} \gamma_{1,X}^2 \end{aligned}$$

<sup>24</sup>Siehe zusammengefasst in Hull (2011, S. 478 ff.) oder Jorion (2007, S. 247 ff.) und ausführlich in Read (1998).

<sup>25</sup>Zu nennen sind die historische Simulation anhand von beobachteten Renditedaten und die Monte-Carlo-Simulation anhand von Verteilungsmodellen, beide ebenfalls u.a. in Hull (2011) und Jorion (2007) beschrieben.

<sup>26</sup>Für die üblichen Werte von  $\alpha$  – meist 5 % oder kleiner – dürfte dieser Bereich alle ökonomisch relevanten Fälle umfassen.

entwickelt, der mittels Cornish-Fisher-Entwicklung für das Quantil  $q_\alpha^*$  die Schiefe  $\gamma_{1,X}$  und die Kurtosis  $\gamma_{2,X}$  der Renditeverteilung in ein parametrisches Modell einbezieht. Gregoriou und Gueyie (2003) haben das Verhältnis  $\frac{\mu_X}{MVaR_X(\alpha)}$  unter dem Namen *Modified Sharpe Ratio* als Performancemaß eingeführt.

Wir werden uns in der Folge mit der Schätzung von Performancemaßen aus beobachteten (Portfolio-)Renditedaten beschäftigen. Dabei wird der Value-at-Risk direkt als Quantil aus den Daten geschätzt und die parametrischen Ansätze zur Ermittlung des Value-at-Risk nicht weiter verfolgt.

### 3.3 Punktschätzung von Performancemaßen

In diesem Abschnitt wird die Punktschätzung der im vorigen Abschnitt dargestellten Performancemaße behandelt.

Die meisten der hier vorgestellten Performancemaße lassen sich als Funktion von Momenten oder partiellen Momenten darstellen. Es liegt nahe, die Punktschätzung von Performancemaßen aus empirischen Daten mittels Ersetzung der theoretischen Momente durch die empirischen Momente vorzunehmen.

Empirische Momente sind unter relativ allgemeinen Voraussetzungen konsistente und asymptotisch normalverteilte Schätzer der theoretischen Momente. Im Falle einfacher Stichproben reicht die Existenz und die Endlichkeit des  $2k$ -ten Moments aus, damit das  $k$ -te empirische Moment ein konsistenter und asymptotisch normalverteilter Schätzer des  $k$ -ten theoretischen Moments ist.<sup>27</sup>

In diesem Abschnitt wird stets mit Überrenditen unter weitgehendem Verzicht auf zusätzliche Indexierung gerechnet.

- Schätzer für den Erwartungswert  $E(X)$  ist das arithmetische Mittel

$$\hat{\mu}_X = \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i.$$

- Schätzer für die Varianz  $\sigma_X^2$  ist - als erwartungstreuer Schätzer - die korrigierte Stichprobenvarianz

$$\hat{\sigma}_X^2 = S^{*2} = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2.$$

- Entsprechend wird die Standardabweichung  $\sigma_X$  mit der korrigierten Stichprobenstandardabweichung  $S^*$ , also der Wurzel aus der korrigierten Stichprobenvarianz, geschätzt.
- Schätzer für die Kovarianz  $\sigma_{X,Y}$  ist die korrigierte Stichprobenkovarianz

$$\hat{\sigma}_{X,Y} = S_{X,Y}^* = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y}).$$

---

<sup>27</sup> vgl. Lehmann (1999, S. 297)

- Der Schätzer des Korrelationskoeffizienten  $\rho_{X,Y}$  lautet entsprechend

$$\widehat{\rho}_{X,Y} = \frac{S_{X,Y}^*}{S_X^* S_Y^*} = \frac{S_{X,Y}}{S_X S_Y} = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2}}.$$

Daraus ergeben sich für die klassischen Performancemaße die in Tabelle 3.1 angegebenen Punktschätzer.

Maß	Definition	Schätzer
Sharpe	$\frac{\mu_X}{\sigma_X}$	$\frac{\bar{X}}{S_X^*}$
Treynor	$\frac{\mu_X}{\beta_X} = \frac{\mu_X \sigma_M^2}{\sigma_{X,M}}$	$\frac{\bar{X} S_M^2}{S_{X,M}} = \frac{\bar{X} S_M}{\widehat{\rho}_{X,M} S_X}$
Jensen	$\mu_X - \beta_X \mu_M = \mu_X - \frac{\mu_M \sigma_{X,M}}{\sigma_M^2}$	$\bar{X} - \frac{\bar{M} S_{X,M}}{S_M^2} = \bar{X} - \frac{\bar{M} \widehat{\rho}_{X,M} S_X}{S_M}$
Modigliani	$\mu_X \frac{\sigma_M}{\sigma_X} - \mu_M$	$\bar{X} \frac{S_M}{S_X} - \bar{M}$

Tabelle 3.1: Punktschätzer für klassische Performancemaße

Diese Schätzer sind in der Regel nicht erwartungstreu. Zum Beispiel ist nach Miller und Gehr (1978) der Erwartungswert des Schätzers  $\widehat{Sh}_X = \bar{X}/S_X^*$  bei unabhängig und identisch normalverteilten Renditen

$$E(\widehat{Sh}_X) = \left( \frac{\Gamma\left(\frac{n-2}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)} \sqrt{\frac{n-1}{2}} \right) Sh_X,$$

wobei  $\Gamma$  in dieser Gleichung für die Gammafunktion steht. Da der Korrekturterm größer als eins ist, werden positive Sharpe-Ratios überschätzt und negative Sharpe-Ratios unterschätzt, und zwar jeweils um einen bestimmten, nur von der Stichprobengröße  $n$  abhängigen Faktor. Für  $n = 12$  beträgt die Verzerrung beispielsweise ca. 7,5 %, für  $n = 60$  ca. 1,29 % und konvergiert mit etwa der Rate  $n^{-1}$  gegen null.

Jobson und Korkie (1981) leiten mit der unten in Abschnitt 4.1 dargestellten Delta-Methode eine Näherung für  $E(\widehat{Sh}_X)$  bei unabhängig und identisch beliebig verteilten Renditen her:

$$E(\widehat{Sh}_X) \approx \left( 1 + \frac{3}{4n} + \frac{100}{128n^2} \right) Sh_X$$



Daraus ergeben sich Verzerrungen von 6,8 % für  $n = 12$  und 1,27 % für  $n = 60$ ; für größere  $n$  liefern die beiden Methoden fast identische Ergebnisse. Der Erwartungswert von  $\widehat{Tr}_X$  lautet<sup>28</sup>

$$E(\widehat{Tr}_X) \approx \left( 1 + \frac{\frac{1}{\rho_{X,M}^2} - 1}{n} + \frac{1 - \frac{4}{\rho_{X,M}^2} + \frac{3}{\rho_{X,M}^4}}{n^2} \right) Tr_X$$

und ist damit auch vom Parameter  $\rho_{X,M}$  abhängig, der in der Regel unbekannt ist, so dass eine Angabe der Verzerrung beim Treynor-Ratio dann nicht a priori möglich ist.

*Partielle Momente* kann man durch die empirischen partiellen Momente schätzen:

$$\widehat{E}(X - z)_+^\alpha = \frac{1}{n} \sum_{i=1}^n (\max(X_i - z, 0))^\alpha$$

Entsprechend erfolgt die Schätzung der Performancemaße des Kappa-Typs

$$\kappa_{\alpha,z} = \frac{E(X - z)}{\sqrt[\alpha]{E(z - X)_+^\alpha}}$$

mit der Formel

$$\widehat{\kappa}_{\alpha,z} = \frac{\bar{X} - z}{\sqrt[\alpha]{\frac{1}{n} \sum_{i=1}^n (\max(z - X_i, 0))^\alpha}}.$$

Die Parameterwahl hat hier Auswirkungen auf die Funktionalität des Schätzers. Wählt man die Schranke  $z$  zu niedrig, kann es besonders bei kurzen Zeitreihen vorkommen, dass sie in der Stichprobe überhaupt nicht unterschritten wird. Dann wird das untere partielle Moment auf den Wert 0 geschätzt und der Wert des Performancemaßschätzers, in dessen Nenner das partielle Moment steht, ist nicht definiert. Beispielsweise wird für  $z$  häufig der Wert 0 eingesetzt, so dass das partielle Moment eine Funktion des Risikos nominaler Vermögensverluste ist. Bei offenen Immobilienfonds und bei Rentenfonds, die in Anleihen mit kurzer Restlaufzeit investieren, sind negative Monatsrenditen selten, so dass häufig kein Wert für  $\text{Kappa}(\alpha, 0)$  bestimmbar ist.

Nicht ganz eindeutig ist im Falle des *Excess Return on Value at Risk (ER-VaR)* die Schätzung des Value-at-Risk (VaR). Da der  $(1 - \alpha)$ -VaR das  $\alpha$ -Quantil  $x_\alpha$  einer Renditeverteilung darstellt, ist das empirische Quantil  $\tilde{x}_\alpha$  dessen „natürlicher“ Schätzer, für das jedoch unterschiedliche Schätzverfahren verbreitet sind.

Unstrittig ist die Bestimmung des empirischen Quantils, wenn das Produkt aus  $\alpha$  und der Stichprobengröße  $n$  ganzzahlig ist. Das 0,05-Quantil aus 60 Einzeldaten ist der drittkleinste Wert  $x_{(3)}$ , da  $0,05 \cdot 60 = 3$ . Ist das Produkt nicht ganzzahlig, gibt es die beiden Möglichkeiten, (I) einen der beiden Nachbarwerte

<sup>28</sup>Jobson und Korkie (1981, S. 892)

$x_{(\lfloor n\alpha \rfloor)}$  oder  $x_{(\lceil n\alpha \rceil)}$  auszuwählen<sup>29</sup> oder (II) zwischen diesen beiden Werten geeignet zu interpolieren. Möglichkeit (I) hat die vielfach erwünschte Eigenschaft, dass  $\tilde{x}_\alpha$  gleich einem der beobachteten Werte ist. Die empirische Quantilfunktion  $Q(\alpha|x_{(1)}, \dots, x_{(n)})$  ist dann eine Treppenfunktion. Möglichkeit (II) hat dagegen den Vorteil, dass  $Q(\alpha|x_{(1)}, \dots, x_{(n)})$  für  $\alpha \in ]0; 1[$  stetig ist. Für die Anwendung auf den ERVaR ist die zweite Eigenschaft von Bedeutung: Es soll bei Variationen von  $\alpha$  (bzw.  $n$ ) nicht zu plötzlichen Sprüngen des Quantilschätzers kommen. Würde z.B. ein Schätzer gewählt, der bei der Parameterkombination ( $\alpha = 0,05$ ,  $n = 60$ ) den drittkleinsten Wert, bei ( $\alpha = 0,0501$ ,  $n = 60$ ) bzw. ( $\alpha = 0,05$ ,  $n = 61$ ) jedoch den viertkleinsten Wert wählt<sup>30</sup>, würde sich das zu erwartende Schätzergebnis des VaR bei nur geringfügigen Parametervariationen deutlich verändern.

Eine Diskussion verschiedener Schätzer findet man in Hyndman und Fan (1996).<sup>31</sup> In der vorliegenden Arbeit wird eine dort unter Nummer 7 aufgeführte Methode der linearen Interpolation zwischen benachbarten Ordnungsstatistiken benutzt. Die Quantilfunktion  $Q(p)$  ist dabei eine stückweise lineare Funktion, die in der grafischen Darstellung der Quantilfunktion die Punkte  $((0, x_{(1)}), (1/(n-1), x_{(2)}), \dots, ((k-1)/(n-1), x_{(k)}), \dots, (1, x_{(n)}))$  miteinander verbindet, wobei  $x_{(k)}$  die  $k$ -t kleinste Rendite ist. Diese Methode hat den Vorteil, dass die Quantilfunktion streng monoton wächst, sofern keine Werte mehrfach auftauchen. Das Konzept des Schätzers geht auf Gumbel (1939) zurück.

## 3.4 Risikoloser Zins, Marktportfolio und Zeitskala

In diesem Abschnitt sollen kurz einige Problemfelder skizziert werden, die bei der Implementierung bzw. der Schätzung von Performancemaßen regelmäßig relevant werden. In den ersten beiden Unterabschnitten werden Probleme bei der Wahl von risikolosem Zinssatz und Marktportfolio vorgestellt. Der dritte Teil dieses Abschnitts handelt von zeitlichen Dimension der Performancemessung: die gewünschte Anlagedauer und die Frequenz der Renditedaten zur Schätzung der Parameter bzw. der Performancemaße.

### 3.4.1 Der risikolose Zins

Die meisten vorgestellten Performancemaße beinhalten einen Term  $R_f$ , den „risikolosen Zinssatz“, der von der Rohrendite des Portfolios zur Ermittlung der Überrendite abgezogen wird. Diesen Zinssatz kann man als Opportunitätskostensatz auffassen. Anstatt das eingesetzte Kapital in ein risikobehaftetes Portfolio zu investieren, könnte der Investor das Geld auch möglichst risikolos verzinslich anlegen und dadurch einen vorab festgelegten, sicheren Ertrag erzielen.

<sup>29</sup>Zur Notation:  $\lfloor a \rfloor$  ist die größte ganze Zahl kleiner oder gleich  $a$ ;  $\lceil a \rceil$  ist die kleinste ganze Zahl größer oder gleich  $a$ .

<sup>30</sup>Dies wäre bei der empirischen Quantilfunktion  $Q(\alpha) = \min\{x|\widehat{F}(x) \geq \alpha\}$  der Fall.

<sup>31</sup>Rob Hyndman ist zusammen mit Ivan Frohne Autor der R-Funktion *quantile*, in der alle in Hyndman und Fan (1996) beschriebenen Methoden implementiert sind.

Wenn ein Festgeldkonto bei seiner Hausbank für einen Investor die für ihn relevante Alternative zu einer Kapitalmarktanlage darstellt, kann er den Zinssatz, den ihm seine Bank anbietet, als risikolosen Zinssatz  $R_f$  verwenden. Wird die Performancemessung dagegen nicht nur für die eigene Anlageentscheidung oder ex post durchgeführt, braucht man eine „sichere Rendite“ (oder eine Renditezeitreihe) aus einer externen Quelle.

Zunächst bezieht sich ein Zinssatz stets auf eine bestimmte Währung und einen Anlagezeitraum. Als Währung bietet sich aus entscheidungstheoretischer Sicht die Währung des Landes an, in dem der Investor lebt bzw. die Währung, in der er seinen zukünftigen Konsum zu bezahlen beabsichtigt. Werden Portfolios untersucht, die nur in Wertpapiere einer bestimmten Währung investieren, kann man auch jene Währung heranziehen, wobei dann das Währungsrisiko des Investors unberücksichtigt bleibt. Schließlich kann man auch eine große Anlage- und Reservewährung wie Euro oder US-Dollar verwenden. Die beste Lösung für die Laufzeit des risikolosen Zinssatzes wäre die geplante Anlagedauer des Investors bzw. die geplante Zeitspanne bis zur nächsten Neubewertung und Vermögensumschichtung des Investors. Soll seine jetzige Anlageentscheidung z.B. die Vermögensstruktur für zwei Jahre festlegen, stellt der aktuelle zweijährige Zinssatz die relevante Anlagealternative dar.

Praktikerbücher zur Performancemessung geben zu diesem Thema zum Teil wenig Auskunft. Fischer (2010, S. 456) beschränkt sich auf die Angabe, dass in der Praxis in der Regel ein „Geldmarktzinssatz wie etwa das Dreimonatsgeld“ als risikoloser Zinssatz verwendet wird. Auch Bacon (2004, S. 57) gibt lediglich an, der risikolose Zinssatz sei „the return I should expect on a riskless asset (e.g., the interest rate on cash or Treasury bills).“ In der Neufassung des Buches (Bacon (2013), S. 46 f.) behandelt er das Thema in einem Unterabschnitt, gibt aber auch keine konkreten Empfehlungen für die Wahl des Zinssatzes. Er sieht die Wahl eines bestimmten Zinssatzes als Aufgabe des individuellen Investors unter Berücksichtigung der Währung und seines Anlagezeitraums. Beim Vergleich mehrerer Portfolios sei wichtig, dass die Methodik für alle Portfolios konsistent sei. Die konkrete Wahl des Zinssatzes sei nicht besonders bedeutend, weil bei einer Änderung des risikolosen Zinses alle Portfolios gleichermaßen betroffen sind: Wird z.B. ein um 0,5 Prozentpunkte höherer Zinssatz gewählt, sinkt die Überrendite aller Portfolios gleichmäßig um 0,5 Prozentpunkte. Christopherson et al. (2009, S. 57 ff.) sehen den risikolosen Zinssatz als „the minimum that an investor could expect in the most secure investment“, wozu sie kurz laufende Staatspapiere solventer Industriestaaten (USA, Vereinigtes Königreich, Japan, Deutschland, Schweiz,...) zählen.

Wenn man sich empirische Studien anschaut, findet man ganz verschiedene Entscheidungen in dieser Frage: Sharpe (1966) selbst setzt einen risikolosen Zinssatz von 3 % p.a. an und bezieht sich dabei auf die ungefähre Rendite langlaufender US-Staatsanleihen zu Beginn der von ihm untersuchten Anlageperiode. Eling und Schuhmacher (2007, siehe dort Fußnote 10) wählen eine konstante risikolose Monatsrendite (hier: 0,35 %) entsprechend der Rendite der zehnjährigen US-

Staatsanleihen am Ende ihres Beobachtungszeitraums.<sup>32</sup> Schmid und Schmidt (2006) benutzen den von *Datastream* bereitgestellten 12-Monatsszinssatz ECW-CM1Y, der auf der Verzinsung von Staatsanleihen basiert. Manche Studien machen überhaupt keine Angabe zum konkret verwendeten Zinssatz (z.B. Kosowski et al. (2007)).

Weitere Kandidaten für den Euro-Raum sind die EURIBOR- und LIBOR-Zinssätze, die Referenzzinssätze für unbesicherte Interbankenverbindlichkeiten darstellen und ein wichtiger Orientierungspunkt für die kurzfristigen Zinssätze im Laufzeitspektrum von einer Woche bis zu einem Jahr sind.<sup>33</sup> Die Deutsche Börse veröffentlicht börsentäglich REX-Renditen, die die Verzinsung von handelbaren deutschen Bundeswertpapieren im Laufzeitbereich von einem Jahr bis zehn Jahren abbilden. Für den US-Dollar gibt es *US Treasury Bonds Rates* für US-amerikanische Staatspapiere mit Laufzeiten von drei Monaten bis zu dreißig Jahren.

Solche regelmäßig und nach öffentlich bekannt gemachter Methodik erstellten Referenzzinssätze eignen sich gut als Datengrundlage für Schätzungen von Performancemaßen. Es bleibt die Frage, inwieweit die genannten Zinssätze wirklich risikolose Anlagen repräsentieren. Bis zum Beginn der Finanzkrise ab 2007 wurde dieser Aspekt kaum beachtet, auch weil die Zinsdifferenzen bei identischen Laufzeiten ziemlich gering waren. In der Folge sind die Renditeabstände zwischen erstklassigen Staatspapieren und unbesicherten Interbankenkrediten angestiegen.<sup>34</sup> Mittlerweile ist die Tatsache, dass keine Geldanlage absolut sicher ist, allgemein bewusst und es ist unklar, in welchem Maße in den als am sichersten geltenden Anlagen (Schuldverschreibungen der USA bzw. mitteleuropäischer Euro-Länder) Risikoprämien eingepreist sind. Möchte man diese Überlegung berücksichtigen, könnte man von den Renditen dieser Staatspapiere Prämien für entsprechende Kreditversicherungen abziehen. Absolute Sicherheit der „risikolosen“ Anlage ist aber weder für die Performanceanalyse notwendig noch in der Realität erzielbar.

### 3.4.2 Das Marktportfolio

Bei Treynor-Ratio, Jensens Alpha und MM wird die Performance von Portfolios unter Zuhilfenahme eines Marktportfolios bewertet. Zur Schätzung dieser Maße muss auch hier eine Entscheidung getroffen werden, welche konkrete Renditezeitreihe dafür herangezogen wird.

Zumindest bei den ersten beiden der genannten Maße wird davon ausgegangen, dass der Investor über ein perfekt diversifiziertes Portfolio verfügt, sodass für

---

<sup>32</sup>Sie berichten, auch andere risikolose Zinssätze ausprobiert zu haben, aber – was die Rankingreihenfolge von Portfolios nach unterschiedlichen Performancemaßen angeht – zu fast identischen Ergebnissen gekommen zu sein.

<sup>33</sup>Der Nachteil dieser beiden Zinssätze ist, dass sie nicht auf Marktdaten beruhen, sondern auf Meldungen ausgewählter Banken. Durch die im LIBOR-Skandal 2012 aufgedeckten Manipulationen hat das Vertrauen in diese Referenzzinssätze stark gelitten.

<sup>34</sup>Der 12-Monats-EURIBOR und die einjährige REX-Rendite z.B. unterschieden sich in den Jahren nach 2000 zumeist nur um ein bis zwei Zehntelprozente. Seit etwa 2007 beträgt der Unterschied etwa 0,5 bis 1,5 Prozentpunkte.

ihn das systematische Risiko die relevante Risikoart ist, wenn er neue Investitionsentscheidungen treffen will.<sup>35</sup> Sein Marktportfolio wäre für ihn dann die Gesamtheit der Vermögensgegenstände, in die er realistischerweise investieren kann (engl.: Investment Opportunity Set).

Für einen großen Investor wären das die weltweiten Aktien- und Rentenmärkte, Immobilien, Derivate und sowie alle anderen Anlagemöglichkeiten, in die er in der Lage und bereit ist zu investieren. Lediglich Vermögenswerte, in die er nicht oder nur zu prohibitiv hohen Kosten investieren kann (z.B. nicht öffentlich notierte Unternehmen, Staatsunternehmen, Vermögenswerte in Ländern mit Beschränkungen für den Besitz durch Ausländer), sind nicht entscheidungsrelevant und sollten nicht im Marktportfolio berücksichtigt werden. Für solche Investoren wäre ein möglichst breit gefasstes Marktportfolio sinnvoll, das sehr viele Vermögensklassen und Anlageregionen beinhaltet. Ein klarer Standard hierfür hat sich nicht entwickelt. Angesichts der Vielzahl mehr oder weniger zugänglicher Anlagebereiche weltweit gibt es keine optimale Lösung. Eine Möglichkeit ist die Konstruktion eines Portfolios mittels weit gefasster Indizes für einzelne Vermögensklassen. Kandidaten wären z.B. der *MSCI All Country World Index (ACWI)* für Aktien oder der *Barclays Multiverse Index* für Renten. Die Firma Dow Jones bietet die *Dow Jones Relative Risk Indices* an, die Mischungen aus weltweiten Aktien-, Renten- und Geldmarktindizes für verschiedene Risikoprofile darstellen. Einen besonders weiten Ansatz beschreiben Clark und Kassimatis (2011), die anhand makroökonomischer Daten einen Proxy für das Weltvermögen oder Welt-Marktportfolio konstruieren. Kleinere Investoren, die ihr Anlagespektrum auf Aktien ihres Heimatmarktes, Bankeinlagen und bis zur Fälligkeit gehaltene Anleihen<sup>36</sup> begrenzen, können einen passenden Aktienindex verwenden.

Soll die Leistung von Fondsmanagern mittels Performancemessung bewertet werden, dann sollte das Marktportfolio ihrem vorgegebenen Investment Opportunity Set entsprechen. Manager von offenen Investmentfonds sind in der Regel durch ihre Anlagerichtlinien auf bestimmte Marktsegmente festgelegt. Indizes dieser Marktsegmente können dann als Benchmarks oder Marktportfolios dienen (z.B. bei einem Fonds für europäische Standardwerte der STOXX Europe 50). Mit solchen Indizes berechnete Performancekennzahlen können im Prinzip auch für Anleger aussagekräftig sein, wenn sie Rückschlüsse auf die Qualität des Fondsmanagements erlauben. Manche Anlageklassen wie z. B. Hedgefonds haben kaum Anlagebeschränkungen und handeln mit komplexen Produkten, so dass hier keine geeigneten produktnahen Benchmarks existieren, die als vergleichbare passive Anlage herangezogen werden könnten.

---

<sup>35</sup>Der Investor sollte bei der Anlageentscheidung berücksichtigen, dass er möglicherweise auch Klumpenrisiken wie die eigene Arbeitskraft oder direkt gehaltene Immobilien in seinem Vermögen hat.

<sup>36</sup>Anleihen, die bis zur Fälligkeit gehalten werden, unterliegen keinem Zinsänderungsrisiko. Haben die Emittenten sehr gute Bonität, kann auch das Ausfallrisiko vernachlässigt werden, so dass man die Anlage in Anleihen der „risikolosen“ Anlage in Bankeinlagen gleichstellen kann.

### 3.4.3 Die Zeitskala

Zum Dritten muss man sich festlegen, auf welche Zeitspanne (Anlageperiode) sich ein Performancemaß beziehen soll. Aus Sicht des Anlegers sollte sich ein Performancemaß auf die Zeitspanne beziehen, für die die Anlage ab dem Entscheidungszeitpunkt festgelegt werden soll. Plant der Anleger, seine jetzt zu treffenden Anlageentscheidungen in einem Jahr das nächste Mal zu überprüfen und ggf. anzupassen, sollten sich alle einfließenden Größen auf ein Jahr beziehen.

Bei der Schätzung eines Performancemaßes aus der beobachteten Wertentwicklung eines Anlageobjekts benutzt man die Daten von  $n$  beobachteten Anlageperioden der Vergangenheit. Je größer die Zeitreihenlänge  $n$  ist, desto genauer kann man, Konstanz der gemessenen Größen vorausgesetzt, das Performancemaß schätzen. Werden Tagesrenditen verwendet, kommt man schnell auf eine hohe Zeitreihenlänge. Sie sind aber nicht immer verfügbar (Bsp.: Hedgefonds) und weisen oft ausgeprägte Zeitreiheneigenschaften wie temporär abhängige Volatilitäten auf. Je länger die Anlageperiode ist, desto besser sind die beobachteten Renditen mit der vereinfachenden Annahme unabhängig und identisch (normal-)verteilter Renditen vereinbar (siehe Abschnitt 2.3). Je größer  $n$  und je länger die Anlageperiode ist, desto weiter muss man jedoch in die Vergangenheit zurückgehen. Ob Renditen, die länger als drei, fünf oder zehn Jahre zurückliegen, noch relevante Informationen für die heutigen Renditeeigenschaften liefern, muss im Einzelfall beurteilt werden. Ein in Literatur und Praxis verbreiteter Kompromiss ist die Verwendung der Monatsrenditen von fünf Jahren ( $n = 5 \cdot 12 = 60$ ). Ein Investor, der seine Anlageentscheidungen monatlich überprüft, würde somit die Werte der Performancemaße für die vor ihm liegende Monatsperiode mittels der Daten aus 60 gleich langen Perioden der unmittelbaren Vergangenheit schätzen. Bevorzugt ein Investor eine andere Dauer der Anlageperiode, dann kann er die aus Monatsdaten berechneten Werte mancher Performancemaße *unter der Annahme, dass die Renditen unabhängig und identisch verteilt sind*, umrechnen. Für die Umrechnung von Performancemaßen auf Monatsbasis in Performancemaße für eine Anlage von  $k$  Monaten gilt dann:<sup>37</sup>

$$Sh_{X,k} = \frac{\mu_{X,k}}{\sigma_{X,k}} = \frac{k \cdot \mu_{X,1}}{\sqrt{k} \cdot \sigma_{X,1}} = \sqrt{k} \cdot Sh_{X,1}$$

$$Tr_{X,k} = \frac{\frac{\mu_{X,k}}{\sigma_{MX,k}}}{\sigma_{M,k}^2} = \frac{\frac{k \cdot \mu_{X,1}}{k \cdot \sigma_{MX,1}}}{k \cdot \sigma_{M,1}^2} = k \cdot Tr_{X,1}$$

$$\alpha_{X,k} = \mu_{X,k} - \frac{\sigma_{MX,k}}{\sigma_{M,k}^2} \mu_{M,k} = k \cdot \mu_{X,1} - \frac{k \cdot \sigma_{MX,1}}{k \cdot \sigma_{M,1}^2} \cdot k \cdot \mu_{M,1} = k \cdot \alpha_{X,1}$$

$$MM_{ori,X,k} = \mu_{X,k} \frac{\sigma_{X,k}}{\sigma_{M,k}} = k \cdot \mu_{X,1} \frac{\sqrt{k} \sigma_{X,1}}{\sqrt{k} \sigma_{M,1}} = k \cdot MM_{ori,X,1}$$

<sup>37</sup>Für die Umrechnung des Sharpe-Ratios bei nicht unabhängig und identisch verteilten Renditen siehe Schmid und Schmidt (2010).

$$MM_{alt,X,k} = \mu_{X,k} \frac{\sigma_{X,k}}{\sigma_{M,k}} - \mu_{M,k} = k \cdot \mu_{X,1} \frac{\sqrt{k}\sigma_{X,1}}{\sqrt{k}\sigma_{M,1}} - k\mu_{M,1} = k \cdot MM_{alt,X,1}$$

Für  $\kappa_X(\alpha, z)$  und  $ERVaR_X(\alpha)$  existieren keine derartigen Formeln, weil die Umrechnung der partiellen Momente bzw. der Quantile von der Form der Renditeverteilung abhängig ist.

### 3.5 Studien zur Performancemessung

Zum Abschluss des Kapitels folgen nun noch einige Hinweise auf Studien zu verschiedenen Eigenschaften von Wertpapierperformance und Performancemaßen. Ein wichtiger Punkt bei der Nutzung der Performancemessung für Anlageentscheidungen ist die Frage nach der Dauerhaftigkeit (Persistenz) von Performance, engl.: *performance persistence*. Zeigt ein Wertpapier, bei dem man in der Vergangenheit überdurchschnittliche Performance beobachten konnte, eine solche auch in den Folgeperioden? Brown und Goetzmann (1995) untersuchen Investmentfonds und finden echte Hinweise auf Persistenz allenfalls bei Fonds mit unterdurchschnittlicher Performance. Nachweise von Performance Persistence in älteren Studien seien darauf zurückzuführen, dass dort der Survivorship Bias nicht berücksichtigt worden sei. Weiterhin sei bei einer Trendfolgerstrategie (Investition in die Gewinner der Vorperiode) zwar eine im Vergleich zum Gesamtmarkt leicht überdurchschnittliche Rendite beobachtbar, die aber mit erhöhtem Gesamtrisiko einhergehe. Auch Carhart (1997) findet auf Basis eines Vier-Faktoren-Modells<sup>38</sup> bei Investmentfonds eine leichte kurzfristige Persistenz bei schwächeren Fonds. Der weitaus größte Teil der in den Rohdaten beobachtbaren Persistenz könne durch die Faktoren erklärt werden. Edwards und Caglayan (2001) sehen bei Hedgefonds dagegen ein etwas höheres Maß an Persistenz und werten dies als Hinweis darauf, dass zumindest einige Hedgefonds-Manager zu nachhaltig besseren Anlageergebnissen fähig seien. Eling (2009) findet Persistenz bei Hedgefonds vorwiegend für einen Zeitraum bis zu sechs Monaten. Unterschiede der Ergebnisse verschiedener Studien seien zum großen Teil auf unterschiedliche Methodologien zurückzuführen. Zudem unterscheide sich der Grad der Persistenz deutlich zwischen Hedgefonds-Strategien. Insgesamt sei angesichts von Einstiegs- und Ausstiegsfristen zweifelhaft, ob sich Anleger Persistenzeffekte nutzbar machen könnten.<sup>39</sup>

Eine zweite Frage ist die, wie stark sich die Performancemaße bei der Reihung von Portfolios unterscheiden. Ist die Wahl des Performancemaßes bei der Bildung von Rangfolgen wichtig oder reihen sie eine gegebene Anzahl von Portfolios weitgehend identisch? Eling und Schuhmacher (2007) berechnen für 2763

<sup>38</sup>Seinen Regressionsansatz mit den Faktoren Markt, Unternehmensgröße, Buchwert-Marktwert-Verhältnis und Momentum kann man als eine Verallgemeinerung von Jensens Alpha auffassen.

<sup>39</sup>Im Anhang der Artikels von Edwards und Caglayan findet man eine kommentierte Liste weiterer Veröffentlichungen zum Thema Performance Persistence. Eine aktuellere Übersicht zur Performance Persistence bei Hedgefonds steht im Anhang des Artikels von Eling.

$r_{Sp;A,B}$	Omega	Sortino	Kappa( $\alpha = 3$ )	MSR
Sharpe	0,99	0,99	0,98	0,97
Omega		0,99	0,98	0,97
Sortino			$\approx 1,00$	0,98
Kappa( $\alpha = 3$ )				0,98

Tabelle 3.2: Korrelation der Schätzer ausgewählter Performancemaße nach Eling und Schuhmacher (2007)

Hedgefonds eine größere Zahl von Performancemaßen<sup>40</sup> und ermitteln die Rangkorrelationskoeffizienten nach Spearman für Paare von Performancemaßen. Für Maße ohne Marktportfolio lag die Rangkorrelation dabei stets bei mindestens 0,92. Die Werte der Rangkorrelationskoeffizienten zwischen Sharpe-Ratio, Omega, Sortino-Ratio, Kappa( $\alpha = 3$ ) und Modified Sharpe Ratio<sup>41</sup> stehen in Tabelle 3.2.<sup>42</sup> Die Performancemaße reihen die Fonds fast identisch. Weniger stark ist die Rangkorrelation, wenn man Performancemaße auf Basis des systematischen Risiko betrachtet. Die Rangkorrelation von Sharpe-Ratio mit Jensens Alpha und Treynor-Ratio lag bei 0,69 bzw. 0,44.<sup>43</sup> Eling (2008) dehnt diese Analyse auf Aktien-, Renten-, Immobilienfonds und einige weitere Assetklassen aus, wobei er sich auf Performancemaße ohne Marktportfolio beschränkt. Die Ergebnisse sind sehr ähnlich, die Rangkorrelation zum Sharpe-Ratio liegt für alle eben genannten Maße in allen Assetklassen bei mindestens 0,97. Eling stellt daraufhin die Frage, weshalb man andere Maße einsetzen sollte, wenn sie doch mit dem einfachen und vergleichsweise gut interpretierbaren Sharpe-Ratio so eng korreliert sind und empfiehlt die durchgängige Nutzung des Sharpe-Ratios. Schuhmacher und Eling (2012) geben eine theoretische Begründung für die hohe Rangkorrelation an. Sie zeigen, dass sich eine ganze Reihe von Performancemaßen (darunter Omega, Sortino, Kappa( $\alpha = 3$ ), Normal-ERVaR und M-CVaR) unter bestimmten Bedingungen als streng monoton steigende Funktionen des Sharpe-Ratios darstellen lassen. Die wesentliche Bedingung ist dabei, dass die Renditen einer Lage-Skalen-Familie angehören. Wenn sich also die Renditeverteilungen der relevanten Anlagealternativen im Wesentlichen nur in Mittelwert und Standardabweichung unterscheiden, kann der Investor erwarten, dass jene Performancemaße sehr ähnliche Ergebnisse wie das Sharpe-Ratio liefern werden, selbst wenn die Verteilungen z. B. eine hohe Kurtosis und/oder eine deutliche Schiefe aufweisen.

<sup>40</sup>darunter Sharpe-Ratio, Omega, Sortino-Ratio, Kappa( $\alpha = 3$ ), ERVaR, Treynor-Ratio und Jensens Alpha

<sup>41</sup>Die Autoren berechnen den Excess Return on Value-at-Risk nicht über das empirische Quantil, sondern nur über Normal-VaR und Modified VaR

<sup>42</sup>entnommen aus Table 2, S. 2639, Eling und Schuhmacher (2007)

<sup>43</sup>Dabei wurde ein Mischportfolio aus 60 % Renten, 20 % Aktien und je 10 % Immobilien und Geldmarktanlage als Marktportfolio konstruiert.



## Kapitel 4

# Statistische Inferenz

Der Punktschätzung von Performancemaßen folgt nun die Intervallschätzung sowie die Bestimmung der Varianz von Performancemaß-Schätzern. Alle in Abschnitt 3.3 vorgestellten Punktschätzer beinhalten Quotienten von Zufallsvariablen, so dass die Herleitung der exakten Verteilung des Schätzers analytisch im Allgemeinen nicht möglich ist.

Zur Näherung der Verteilung der Performancemaß-Schätzer bzw. einzelner Parameter dieser Verteilung aus einer Stichprobe stehen zwei grundsätzlich verschiedene Ansätze zur Verfügung. Die *Delta-Methode* nutzt zur Bestimmung der Varianz des Performancemaß-Schätzers, welcher eine nichtlineare Funktion aus Momenten ist, eine asymptotische lineare Näherung. In die sich ergebende Formel aus theoretischen Momenten können dann empirische Momente eingesetzt werden, die aus der Stichprobe errechnet werden. Bei der Bootstrap-Methode versucht man, auf Basis der Stichprobendaten simulativ eine große Zahl neuer „Stichproben“ zu generieren und daraus jeweils die interessierenden Maße zu schätzen. Die empirische Verteilung dieser simulierten Maße wird als Näherung der Verteilung des Performancemaß-Schätzers verwendet.

In den folgenden beiden Abschnitten werden Delta-Methode und Bootstrap jeweils zuerst theoretisch dargestellt, danach werden bereits aus der Literatur vorliegende Ergebnisse referiert. Im Hauptteil werden mittels Delta-Methode Varianzformeln für weitere Performancemaß-Schätzer hergeleitet.<sup>1</sup> Schließlich werden Möglichkeiten zur Ermittlung von Konfidenzintervallen diskutiert. Die Bewertung der Ergebnisse der beiden Verfahren anhand von Simulationen erfolgt im nächsten Kapitel.

---

<sup>1</sup>Für den Bootstrap-Ansatz entfällt dieser Unterabschnitt, da aus dem Bootstrap keine geschlossenen Formeln resultieren.

## 4.1 Die Delta-Methode

### 4.1.1 Ansatz

Ausgangspunkt der Delta-Methode, auch *Cramérs Theorem* genannt, ist die Taylor-Entwicklung einer Funktion  $g(x)$  um die Stelle  $x_0$ :

$$g(x_0 + h) = g(x_0) + h \frac{dg(x)}{dx}(x_0) + \frac{h^2}{2} \frac{d^2g(x)}{d^2x}(x_0) + \dots + \frac{h^k}{k!} \frac{d^k g(x)}{d^k x}(x_0) + o(h^k)$$

mit  $h \in \mathbb{R}$  und unter der Annahme, dass  $g(x)$  an der Stelle  $x_0$  mindestens  $k$ -mal differenzierbar ist. Ein unbekannter Funktionswert  $g(x_0 + h)$  wird also mit dem Funktionswert und dessen Ableitungen an einer bekannten Stelle  $x_0$  angenähert. Je höher die Ordnung  $k$  der Approximation, desto geringer ist für  $h < 1$  der Näherungsfehler  $o(h^k)$  um den Preis steigender Komplexität.<sup>2</sup> Die lineare Näherung

$$g(x_0 + h) = g(x_0) + h \frac{dg(x)}{dx}(x_0) + o(h)$$

konvergiert für  $h \rightarrow 0$  bereits mit  $o(h)$ . Die Taylor-Approximation kann man nun dafür verwenden, die asymptotische Verteilung einer Funktion  $g(X_n)$  des empirischen Moments  $X_n = \bar{\theta}$  zu bestimmen. Nach dem Zentralen Grenzwertsatz gilt für  $\bar{X}$  aus einer einfachen Stichprobe, die Existenz von  $\sigma_X^2$  vorausgesetzt,

$$\sqrt{n}(\bar{X} - \mu) \sim N(0, \sigma_X^2),$$

wobei  $\sim$  für die Konvergenz nach Verteilung steht.

**Cramérs Theorem** lautet:<sup>3</sup>

Wenn

1.  $\sqrt{n}(\hat{\theta}_n - \theta) \sim N(0, \sigma^2)$ ,
2.  $g(\dots)$  eine Funktion  $\mathbb{R} \rightarrow \mathbb{R}$  ist, die in der Umgebung von  $\theta$  stetig ist und
3. die Ableitung  $dg(\theta)/d\theta$  an der Stelle  $\theta$  nicht gleich null ist,

dann gilt:

$$\sqrt{n}(g(\hat{\theta}_n) - g(\theta)) \sim N\left(0, \left(\frac{dg(\theta)}{d\theta}\right)^2 \sigma^2\right).$$

Wenn also  $\hat{\theta}_n$  nach dem Zentralen Grenzwertsatz gegen  $\theta$  konvergiert, so konvergiert auch die (geeignete) Funktion  $g(\hat{\theta}_n)$  gegen  $g(\theta)$ , ist asymptotisch normalverteilt und hat näherungsweise die Varianz

$$\sigma_{g(\hat{\theta}_n)}^2 \approx \frac{1}{n} \left(\frac{dg(\theta)}{d\theta}\right)^2 \sigma^2.$$

---

<sup>2</sup>Die Notation  $b_n = o(a_n)$  bedeutet, dass die Folge  $b_n$  für  $n \rightarrow \infty$  langsamer wächst als die Folge  $a_n$ , also  $\lim_{n \rightarrow \infty} \frac{b_n}{a_n} = 0$ .

<sup>3</sup>Beweis siehe Lehmann (1999, S. 86)

Dieser Satz lässt sich auf den mehrdimensionalen Fall eines Parametervektors  $\boldsymbol{\theta}$  und den Vektor der entsprechenden Schätzer  $\boldsymbol{\theta}_n$  verallgemeinern.<sup>4</sup> Dazu benötigen wir zunächst die Erweiterung des Zentralen Grenzwertsatzes auf multivariate Zufallsvariable.

Sei  $\mathbf{X}_k$ ,  $k = 1, \dots, n$ , eine Folge unabhängig und identisch verteilter  $d$ -dimensionaler Zufallsvariablen mit Erwartungswertvektor  $\boldsymbol{\mu} \in \mathbb{R}^{d \times 1}$  und Kovarianzmatrix  $\boldsymbol{\Sigma} \in \mathbb{R}^{d \times d}$ , dann gilt:

$$\sqrt{n} \left( \frac{1}{n} \sum_{k=1}^n \mathbf{X}_k - \boldsymbol{\mu} \right) \sim N(\mathbf{0}, \boldsymbol{\Sigma}).$$

Nun kann Cramérs Theorem auf den multivariaten Fall verallgemeinert werden:

Wenn

1.  $\sqrt{n} \left( \widehat{\boldsymbol{\theta}}_n(\mathbf{X}_1, \dots, \mathbf{X}_n) - \boldsymbol{\theta} \right) \sim N(\mathbf{0}, \boldsymbol{\Sigma})$ ,
2.  $g(\dots)$  eine Funktion  $\mathbb{R}^d \rightarrow \mathbb{R}$  ist, die in der Umgebung von  $\boldsymbol{\theta}$  stetig ist und
3. der Gradient  $\nabla g = (\partial g(\boldsymbol{\theta}) / \partial \boldsymbol{\theta})$  nicht gleich null ist,

dann gilt:

$$\sqrt{n} \left( g(\widehat{\boldsymbol{\theta}}_n) - g(\boldsymbol{\theta}) \right) \sim N(0, (\nabla g)' \boldsymbol{\Sigma} \nabla g).$$

#### 4.1.2 Varianzen für einzelne Maße

Jobson und Korkie (1981) haben die Delta-Methode erstmals zur Herleitung der Varianz von Performancemaßschätzern angewendet. Dementsprechend wird zunächst die Herleitung der Formeln für die dort bereits behandelten Maße Sharpe-Ratio und Treynor-Ratio gezeigt. Die Formeln für die weiteren Maße (Jensen's Alpha, Modigliani-Modigliani, die Kappa-Maße und den ERVaR) folgen im Anschluss daran und sind, soweit nicht anders angegeben, neu.

Es werden jeweils zunächst Formeln für die Varianzen von Performancemaßschätzern unter Annahme unabhängig und identisch verteilter Renditen hergeleitet. Wo sinnvoll, wurden die Formeln für den Fall unabhängig und identisch normalverteilter Renditen vereinfacht. Für das Sharpe-Ratio wird zusätzlich eine Verallgemeinerung von Schmid und Schmidt (2010) ohne Annahme unabhängig und identisch verteilter Renditen vorgestellt. Anschließend folgt jeweils die Herleitung der Formel für die Differenz des Performancemaßes zweier Portfolios.

##### 4.1.2.1 Sharpe-Ratio

Der Schätzer des Sharpe-Ratios eignet sich aufgrund seiner einfachen Struktur gut zur Demonstration der Delta-Methode.

Man kann den Schätzer

$$\widehat{Sh}_X = \frac{\widehat{\mu}_X}{\widehat{\sigma}_X}$$

---

<sup>4</sup>vgl. Spanos (1999, S. 491-493)

zerlegen in die Schätzer der Momente und die sie verbindende Funktion  $g$ :

$$\widehat{Sh}_X = g(a, b) \quad \text{mit} \quad g(x) = \frac{a}{\sqrt{b}}$$

Wir nehmen nun an, dass die Renditen  $X_k$ ,  $k = 1, \dots, n$  unabhängig und identisch verteilt seien. An die Verteilung von  $X$  stellen wir lediglich die Anforderungen, dass mindestens die ersten vier Momente existieren und dass die Varianz  $\sigma_X^2$  größer als null ist, da ansonsten der Nenner des Bruchs null wird. Letztere Annahme, die lediglich besagt, dass die Renditen risikobehaftet sind, wird für alle Herleitungen in diesem Abschnitt benötigt und im Folgenden nicht mehr eigens erwähnt. Darüber hinaus wird keine Verteilungsannahme getroffen. Es gilt der zentrale Grenzwertsatz

$$\sqrt{n} \begin{pmatrix} \bar{X} - \mu_X \\ S_X^2 - \sigma_X^2 \end{pmatrix} \sim N \left( \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \Sigma = \begin{pmatrix} \sigma_X^2 & \sigma_{X,(X-\mu_X)^2} \\ \sigma_{X,(X-\mu_X)^2} & \sigma_{(X-\mu_X)^2}^2 \end{pmatrix} \right). \quad (4.1)$$

Der Ausdruck  $\sigma_{X,(X-\mu_X)^2}$  bezeichnet die Kovarianz zwischen  $X$  und  $(X-\mu_X)^2$ , also  $E((X-\mu_X)((X-\mu_X)^2 - (E(X-\mu_X))^2)) = E(X(X-\mu_X)^2) - \mu_X \sigma_X^2$ . Der Gradient von  $g$  lautet:

$$\nabla g = \left( \frac{\partial g(a, b)}{\partial a}, \frac{\partial g(a, b)}{\partial b} \right)' = \left( \frac{1}{\sqrt{b}}, -\frac{a}{2\sqrt{b^3}} \right)'$$

Nun wird Cramérs Theorem an der Stelle  $(a = \mu_X, b = \sigma_X^2)$  angewendet:

$$\begin{aligned} \sqrt{n}(\widehat{Sh} - Sh) &\sim N(0, (\nabla g)' \Sigma \nabla g) \\ &\sim N \left( 0, \begin{pmatrix} \frac{1}{\sigma_X} & -\frac{\mu_X}{2\sigma_X^3} \end{pmatrix} \begin{pmatrix} \sigma_X^2 & \sigma_{X,(X-\mu_X)^2} \\ \sigma_{X,(X-\mu_X)^2} & \sigma_{(X-\mu_X)^2}^2 \end{pmatrix} \begin{pmatrix} \frac{1}{\sigma_X} \\ -\frac{\mu_X}{2\sigma_X^3} \end{pmatrix} \right) \end{aligned}$$

Die asymptotische Varianz von  $\widehat{Sh}_X$  ist somit:

$$\begin{aligned} n \sigma_{\widehat{Sh}}^2 &= \begin{pmatrix} \frac{1}{\sigma_X} & -\frac{\mu_X}{2\sigma_X^3} \end{pmatrix} \begin{pmatrix} \sigma_X^2 & \sigma_{X,(X-\mu_X)^2} \\ \sigma_{X,(X-\mu_X)^2} & \sigma_{(X-\mu_X)^2}^2 \end{pmatrix} \begin{pmatrix} \frac{1}{\sigma_X} \\ -\frac{\mu_X}{2\sigma_X^3} \end{pmatrix} \\ &= \sum_{i=1}^2 \sum_{j=1}^2 (\nabla g)_i \Sigma_{ij} (\nabla g)_j \\ &= \frac{1}{\sigma_X^2} \sigma_X^2 + 2 \frac{1}{\sigma_X} \left( -\frac{\mu_X}{2\sigma_X^3} \right) \sigma_{X,(X-\mu_X)^2} + \left( -\frac{\mu_X}{2\sigma_X^3} \right)^2 \sigma_{(X-\mu_X)^2}^2 \\ &= 1 - \frac{\mu_X}{\sigma_X} \frac{\sigma_{X,(X-\mu_X)^2}}{\sigma_X^3} + \frac{\mu_X^2}{\sigma_X^2} \frac{\sigma_{(X-\mu_X)^2}^2}{4\sigma_X^4} \\ &= 1 - Sh_X \frac{\sigma_{X,(X-\mu_X)^2}}{\sigma_X^3} + Sh_X^2 \frac{\sigma_{(X-\mu_X)^2}^2}{4\sigma_X^4} \end{aligned}$$

Die Kovarianz  $\sigma_{X,(X-\mu_X)^2}$  und die Varianz  $\sigma_{(X-\mu_X)^2}^2$  kann man umformen in: (siehe Formeln A.1 und A.2 in Anhang A.1)<sup>5</sup>

$$\begin{aligned}\sigma_{X,(X-\mu_X)^2} &= \sigma_X^3 \gamma_{1,X} \\ \sigma_{(X-\mu_X)^2}^2 &= \sigma_X^4 (\gamma_{2,X} - 1)\end{aligned}$$

Hierbei sind wiederum  $\gamma_{1,X}$  die Schiefe von  $X$  und  $\gamma_{2,X}$  die Kurtosis von  $X$ . Eingesetzt in die Varianzformel erhalten wir als Ergebnis:<sup>6</sup>

$$n \sigma_{\widehat{Sh}}^2 = 1 - Sh_X \gamma_{1,X} + \frac{1}{4} Sh_X^2 (\gamma_{2,X} - 1) \quad (4.2)$$

Die Varianz des Schätzers des Sharpe-Ratios beträgt also bei einem Sharpe-Ratio von null  $1/n$ . Ausgehend davon steigt die Varianz, wenn der Betrag des Sharpe-Ratios steigt, insbesondere dann, wenn die Kurtosis der Verteilung hoch ist. Außerdem steigt (sinkt) die Varianz mit steigendem Sharpe-Ratio, wenn die Schiefe der Verteilung negativ (positiv) ist.

Nimmt man zusätzlich an, dass  $X$  *normalverteilt* ist, folgt  $\gamma_{1,X} = 0$  und  $\gamma_{2,X} = 3$ , so dass sich die Formel zu

$$n \sigma_{\widehat{Sh},norm.}^2 = 1 + \frac{1}{2} Sh_X^2 \quad (4.3)$$

vereinfacht.

Analog wird das Verfahren für die **Differenz des Sharpe-Ratios zweier Portfolios**  $X$  und  $Y$  angewandt. Der Schätzer für  $\Delta_{Sh} = Sh_X - Sh_Y$  ist

$$\widehat{\Delta}_{Sh} = \widehat{Sh}_X - \widehat{Sh}_Y = \frac{\widehat{\mu}_X}{\widehat{\sigma}_X} - \frac{\widehat{\mu}_Y}{\widehat{\sigma}_Y} = \frac{\overline{X}}{S_X^*} - \frac{\overline{Y}}{S_Y^*}.$$

Die Funktion  $g$  ist nun

$$\widehat{\Delta}_{Sh} = g(a, b, c, d) \quad \text{mit} \quad g(x) = \frac{a}{\sqrt{b}} - \frac{c}{\sqrt{d}}.$$

Wir gehen nunmehr von einem bivariaten Zufallsvektor  $(X, Y)_k, k = 1, \dots, n$  der (Über-)Renditen zweier Portfolios aus. Die  $(X, Y)_k$  seien unabhängig und identisch verteilt und die (gemischten) Momente  $E(X^u Y^v)$  seien für alle  $\{(u, v) \mid u, v \in \mathbb{N}_0; u + v \leq 4\}$  existent. Der Grenzwertsatz lautet<sup>7</sup>

$$\sqrt{n} \begin{pmatrix} \overline{X} - \mu_X \\ S_X^2 - \sigma_X^2 \\ \overline{Y} - \mu_Y \\ S_Y^2 - \sigma_Y^2 \end{pmatrix} \sim N(\mathbf{0}, \mathbf{\Sigma}),$$

<sup>5</sup>Diese und weitere im Verlauf der Arbeit folgende Termumformungen sind gesammelt in Anhang A.1 zu finden.

<sup>6</sup>Diese und alle weiteren Varianzformeln aus diesem Kapitel sind in Anhang A.2 gesammelt.

<sup>7</sup>Da  $\mathbf{\Sigma}$  als Kovarianzmatrix per Definition symmetrisch ist, wird hier und in der weiteren Folge aus Gründen der besseren Darstellbarkeit darauf verzichtet, die Elemente unterhalb der Hauptdiagonalen aufzuführen.

$$\text{mit } \Sigma = \begin{pmatrix} \sigma_X^2 & \sigma_{X,(X-\mu_X)^2} & \sigma_{X,Y} & \sigma_{X,(Y-\mu_Y)^2} \\ \cdot & \sigma_{(X-\mu_X)^2}^2 & \sigma_{(X-\mu_X)^2,Y} & \sigma_{(X-\mu_X)^2,(Y-\mu_Y)^2} \\ \cdot & \cdot & \sigma_Y^2 & \sigma_{Y,(Y-\mu_Y)^2} \\ \cdot & \cdot & \cdot & \sigma_{(Y-\mu_Y)^2}^2 \end{pmatrix}.$$

Mit dem Gradienten von  $g$

$$\nabla g = \left( \frac{1}{\sqrt{b}}, -\frac{a}{2\sqrt{b^3}}, -\frac{1}{\sqrt{d}}, \frac{c}{2\sqrt{d^3}} \right)'$$

an der Stelle ( $a = \mu_X$ ,  $b = \sigma_X^2$ ,  $c = \mu_Y$ ,  $d = \sigma_Y^2$ ) erhält man die quadratische Form

$$\begin{aligned} n \sigma_{\Delta Sh}^2 &= \sum_{i=1}^4 \sum_{j=1}^4 (\nabla g)_i \Sigma_{ij} (\nabla g)_j \\ &= 2 - \frac{\mu_X \sigma_{X,(X-\mu_X)^2}}{\sigma_X^4} - \frac{2\sigma_{X,Y}}{\sigma_X \sigma_Y} + \frac{\mu_Y \sigma_{X,(Y-\mu_Y)^2}}{\sigma_X \sigma_Y^3} \\ &\quad + \frac{\mu_X^2 \sigma_{(X-\mu_X)^2}^2}{4\sigma_X^6} + \frac{\mu_X \sigma_{Y,(X-\mu_X)^2}}{\sigma_X^3 \sigma_Y} - \frac{\mu_X \mu_Y \sigma_{(X-\mu_X)^2,(Y-\mu_Y)^2}}{2\sigma_X^3 \sigma_Y^3} \\ &\quad - \frac{\mu_Y \sigma_{Y,(Y-\mu_Y)^2}}{\sigma_Y^4} + \frac{\mu_Y^2 \sigma_{(Y-\mu_Y)^2}^2}{4\sigma_Y^6}. \end{aligned}$$

Diese Formel kann man noch etwas vereinfachen

$$\begin{aligned} n \sigma_{\Delta Sh}^2 &= 2 - 2\rho_{X,Y} - \frac{\mu_X \gamma_{1,X}}{\sigma_X} - \frac{\mu_Y \gamma_{1,Y}}{\sigma_Y} + \frac{\mu_X^2 (\gamma_{2,X} - 1)}{4\sigma_X^2} + \frac{\mu_Y^2 (\gamma_{2,Y} - 1)}{4\sigma_Y^2} \\ &\quad + \frac{\mu_X \sigma_{Y,(X-\mu_X)^2}}{\sigma_X^3 \sigma_Y} + \frac{\mu_Y \sigma_{X,(Y-\mu_Y)^2}}{\sigma_X \sigma_Y^3} - \frac{\mu_X \mu_Y \sigma_{(X-\mu_X)^2,(Y-\mu_Y)^2}}{2\sigma_X^3 \sigma_Y^3} \end{aligned}$$

und mittels der Sharpe-Ratios zu

$$\begin{aligned} n \sigma_{\Delta Sh}^2 &= 2 - 2\rho_{X,Y} - Sh_X \gamma_{1,X} - Sh_Y \gamma_{1,Y} + \frac{1}{4} Sh_X^2 (\gamma_{2,X} - 1) \\ &\quad + \frac{1}{4} Sh_Y^2 (\gamma_{2,Y} - 1) + Sh_X \frac{\sigma_{Y,(X-\mu_X)^2}}{\sigma_X^2 \sigma_Y} + Sh_Y \frac{\sigma_{X,(Y-\mu_Y)^2}}{\sigma_X \sigma_Y^2} \\ &\quad - Sh_X Sh_Y \frac{\sigma_{(X-\mu_X)^2,(Y-\mu_Y)^2}}{2\sigma_X^2 \sigma_Y^2} \end{aligned}$$

umformen.

Dennoch bleibt die Formel unübersichtlich. Die Zahl der Summanden lässt sich ohne Spezifizierung der Verteilung nicht reduzieren, da die höheren Kovarianzen bestehen bleiben. In praktischen Anwendungen könnte es sinnvoller sein, die quadratische Form  $(\nabla g)' \Sigma \nabla g$  zu verwenden.

Mit der zusätzlichen Annahme, dass  $X$  und  $Y$  bivariat normalverteilt sind, kann man  $\gamma_{1,X} = \gamma_{1,Y} = 0$  und  $\gamma_{2,X} = \gamma_{2,Y} = 3$  einsetzen. Auch die Kovarianzen lassen sich in diesem Fall vereinfachen. Dazu ist ein Spezialfall des *Theorems*

von Isserlis nützlich:<sup>8</sup>

THEOREM VON ISSERLIS:

Die Zufallsvariablen  $X_1, \dots, X_n$  seien (nicht-singulär oder singulär) multivariat normalverteilt mit  $E(X_i) = 0$  für alle  $i \in \{1, \dots, n\}$  und  $Cov(X_i, X_j) = \sigma_{ij}$ . Dann gilt

$$E(X_1 \cdot X_2 \cdot \dots \cdot X_n) = \begin{cases} 0, & \text{falls } n \text{ ungerade,} \\ \sum \sigma_{i_1, i_2}, \sigma_{i_3, i_4}, \dots, \sigma_{i_{n-1}, i_n} & \text{falls } n \text{ gerade,} \end{cases}$$

wobei sich die Summe über alle Zerlegungen  $\{\{i_1, i_2\}, \{i_3, i_4\}, \dots, \{i_{n-1}, i_n\}\}$  der Indexmenge  $\{1, 2, \dots, n\}$  erstreckt.

Wir benötigen den Spezialfall

$$E(X_1 X_2 X_3 X_4) = \sigma_{1,2} \sigma_{3,4} + \sigma_{1,3} \sigma_{2,4} + \sigma_{1,4} \sigma_{2,3}.$$

Sind die Zufallsvariablen  $A, B, C, D$  multivariat normalverteilt mit Kovarianzmatrix  $\Sigma$ , so sind  $(A - \mu_A), (B - \mu_B), (C - \mu_C), (D - \mu_D)$  multivariat normalverteilt mit Erwartungswertvektor  $\mathbf{0}$  und Kovarianzmatrix  $\Sigma$ . In Anhang A.1 sind alle hier und im weiteren Verlauf dieses Kapitels verwendeten Umformungen auf Basis des Theorems von Isserlis gesammelt.

In der obigen Formel kann man ersetzen:

$$\begin{aligned} \sigma_{Y, (X - \mu_X)^2} &= 0 \\ \sigma_{X, (Y - \mu_Y)^2} &= 0 \\ \sigma_{(X - \mu_X)^2, (Y - \mu_Y)^2} &= 2\sigma_{X,Y}^2 \end{aligned}$$

Die Varianz von  $\hat{\Delta}_{Sh}$  im Fall bivariat normalverteilter Renditen beträgt also

$$\boxed{n \sigma_{\hat{\Delta}_{Sh}}^2 = 2(1 - \rho_{X,Y}) + \frac{1}{2}(Sh_X^2 + Sh_Y^2 - 2Sh_X Sh_Y \rho_{X,Y}^2)} \quad (4.4)$$

Dies ist gleich Memmels (2003) korrigierter Fassung der Formel von Jobson und Korkie (1981, S. 895).<sup>9</sup> Die Varianz des Sharpe-Differenzschätzers nimmt mit steigender Korrelation von  $X$  und  $Y$  ab. Der erste Summand sinkt mit steigender Korrelation von  $X$  und  $Y$ ; im letzten Summanden taucht  $\rho_{X,Y}$  ebenfalls auf, aber durch Quadrierung und Multiplikation mit  $Sh_X$  und  $Sh_Y$ <sup>10</sup> fällt der Term weniger ins Gewicht.

Das Ergebnis (4.4) erhalten wir ebenfalls, wenn wir die gemischten Momente in der Kovarianzmatrix  $\Sigma$  von Anfang an durch die Werte bei Annahme normalverteilter Renditen ersetzen. Die Kovarianzmatrix unter Normalvertei-

<sup>8</sup>Isserlis (1918), zitiert nach Schlittgen und Streitberg (2001, S. 506)

<sup>9</sup>Man beachte, dass zu Jobson und Korkie (1981) noch eine weitere Korrektur, Cadsby (1986), veröffentlicht wurde.

<sup>10</sup>In der Praxis sind Sharpe-Ratios betragsmäßig deutlich kleiner als 1.

lungsumannahme  $\Sigma_{NV}$  lautet dann unter Verwendung des Satzes von Isserlis

$$\Sigma_{NV} = \begin{pmatrix} \sigma_X^2 & 0 & \sigma_{X,Y} & 0 \\ \cdot & 2\sigma_X^4 & 0 & 2\sigma_{X,Y}^2 \\ \cdot & \cdot & \sigma_Y^2 & 0 \\ \cdot & \cdot & \cdot & 2\sigma_Y^4 \end{pmatrix}$$

und das Resultat (4.4) ergibt sich direkt aus  $(\nabla g)' \Sigma_{NV} (\nabla g)$ .

Für die Varianz des Sharpe-Ratios gibt es zudem einige Ansätze für den Nicht-*i.i.d.*-Fall:

Schmid und Schmidt (2010) leiten Formeln für nicht unabhängig und identisch verteilte Renditen her. Für Renditen aus einem GARCH(1,1)-Prozess mit normalverteilten Störtermen lautet die Varianzformel:<sup>11</sup>

$$n \sigma_{\widehat{Sh}, GARCH(1,1)}^2 = 1 + \frac{1}{4} Sh_X^2 (\gamma_{2,X} - 1) \left( 1 + \frac{1 - \beta_1^2 - \beta_1 \alpha_1}{1 - \beta_1^2 - 2\beta_1 \alpha_1} \cdot \frac{2\alpha_1}{1 - \alpha_1 - \beta_1} \right)$$

Schmid und Schmidt stellen zudem einen nichtparametrischen Ansatz für recht allgemeine Bedingungen vor. Sie treffen keine Verteilungsannahme, sondern verlangen lediglich, dass die Renditezeitreihe  $(X_k)_{k \in \mathbb{Z}}$  ein stark stationärer Prozess sei, bei dem die Abhängigkeit der Werte  $X_{t+s}$  vom Wert  $X_t$  für  $s \rightarrow \infty$  mit einer bestimmten Mindestgeschwindigkeit gegen null geht.<sup>12</sup>

Die Schätzvarianz des Sharpe-Ratios ist allgemein

$$n \sigma_{\widehat{Sh}}^2 = \frac{\sigma_{11}}{\sigma_X^2} - \frac{\sigma_{12}}{\sigma_X^3} Sh_X + \frac{1}{4} Sh_X^2 \frac{\sigma_{22}}{\sigma_X^4} \quad \text{mit} \quad \mathbf{\Gamma} := \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix} \quad (4.5)$$

Setzt man für  $\mathbf{\Gamma}$  die aus Formel (4.1) bekannte Kovarianzmatrix  $\Sigma$  ein, erhält man den obigen *i.i.d.*-Ansatz. Schmid und Schmidt schätzen  $\mathbf{\Gamma}$  mit

$$\begin{aligned} \widehat{\mathbf{\Gamma}} &= \sum_{j=-l}^l \widehat{\mathbf{\Gamma}}_j \\ &= \sum_{j=-l}^l \frac{1}{n - |j|} \cdot \\ &\quad \sum_{t=1-j+\max(0,j)}^{n-j+\min(0,j)} \begin{pmatrix} (X_t - \bar{X}_n)(X_{t+j} - \bar{X}_n) & (X_t - \bar{X}_n)(X_{t+j} - \bar{X}_n)^2 \\ (X_t - \bar{X}_n)(X_{t+j} - \bar{X}_n)^2 & (X_t - \bar{X}_n)^2 (X_{t+j} - \bar{X}_n)^2 - (S_n^2)^2 \end{pmatrix}. \end{aligned}$$

Für  $n, l \rightarrow \infty$ , wobei  $l = l(n)$  mit  $l = o(n^{1/3})$ , konvergiert der Schätzer  $\widehat{\mathbf{\Gamma}}$  gegen  $\mathbf{\Gamma}$ .

<sup>11</sup>Zur Gültigkeit der Konvergenzaussage müssen die Parameter des GARCH(1,1)-Prozesses gewisse Bedingungen erfüllen, die im wesentlichen die Stationarität des Prozesses und die Existenz des vierten Moments der Verteilung sichern, siehe Schmid und Schmidt (2010, S. 347f.)

<sup>12</sup>Für die genauen Bedingungen vgl. Schmid und Schmidt (2010, S. 340).



Ledoit und Wolf (2008) schlagen für die Differenz zweier Sharpe-Ratios vor,  $\Gamma$  mittels HAC-Kerndichteschätzung<sup>13</sup> nach Andrews (1991) bzw. Andrews und Monahan (1992) zu schätzen. Die  $\hat{\Gamma}_j$  werden dabei mittels Kerndichteschätzung zu einem Kovarianzschätzer  $\hat{\Gamma}$  aggregiert.

#### 4.1.2.2 Treynor-Ratio

Die Herleitung der Varianz des Schätzers für das Treynor-Ratio erfolgt methodisch analog. Wir gehen wiederum von einer bivariaten Zufallsvariable  $(X, M)_k$ ,  $k = 1, \dots, n$  aus, hier der (Über-)Renditen des Portfolios  $X$  und eines Marktportfolios  $M$ . Die  $(X, M)_k$  seien unabhängig und identisch verteilt und die Momente  $E(X^u M^v)$  seien für alle  $\{(u, v) \mid u, v \in \mathbb{N}_0; u + v \leq 4\}$  existent. Weiterhin dürfen Portfolio- und Marktrendite nicht unkorreliert sein ( $\rho_{X,M} \neq 0$ ), da sonst das  $\beta_X$  und somit der Nenner des Treynor-Ratios null wäre. Der Schätzer für das Treynor-Ratio lautet

$$\widehat{Tr}_X = \frac{\widehat{\mu}_X}{\widehat{\beta}_X} = \frac{\widehat{\mu}_X \widehat{\sigma}_M^2}{\widehat{\sigma}_{X,M}} = \frac{\overline{X} S_M^2}{S_{X,M}}.$$

Die Funktion  $g(a, b, c)$  und dessen Gradient sind

$$g = \frac{a \cdot b}{c}, \quad \nabla g = \left( \frac{b}{c}, \frac{a}{c}, -\frac{ab}{c^2} \right)'.$$

Nach dem Zentralen Grenzwertsatz gilt

$$\sqrt{n} \begin{pmatrix} \overline{X} - \mu_X \\ S_M^2 - \sigma_M^2 \\ S_{X,M} - \sigma_{X,M} \end{pmatrix} \sim N \left( \mathbf{0}, \begin{pmatrix} \sigma_X^2 & \sigma_{X,(M-\mu_M)^2} & \sigma_{X,(M-\mu_M)(X-\mu_X)} \\ \cdot & \sigma_{(M-\mu_M)^2}^2 & \sigma_{(M-\mu_M)^2,(M-\mu_M)(X-\mu_X)} \\ \cdot & \cdot & \sigma_{(M-\mu_M)(X-\mu_X)}^2 \end{pmatrix} \right).$$

Die Anwendung von Cramérs Theorem an der Stelle ( $a = \mu_X$ ,  $b = \sigma_M^2$ ,  $c = \sigma_{X,M}$ ) ergibt

$$\begin{aligned} n \sigma_{\widehat{Tr}}^2 &= \sum_{i=1}^3 \sum_{j=1}^3 (\nabla g)_i \Sigma_{ij} (\nabla g)_j \\ &= \frac{\sigma_M^4 \sigma_X^2}{\sigma_{X,M}^2} + \frac{2\sigma_M^2 \mu_X \sigma_{X,(M-\mu_M)^2}}{\sigma_{X,M}^2} - \frac{2\mu_X \sigma_M^4 \sigma_{X,(M-\mu_M)(X-\mu_X)}}{\sigma_{X,M}^3} + \frac{\mu_X^2 \sigma_{(M-\mu_M)^2}^2}{\sigma_{X,M}^2} \\ &\quad - \frac{2\mu_X^2 \sigma_M^2 \sigma_{(M-\mu_M)^2,(M-\mu_M)(X-\mu_X)}}{\sigma_{X,M}^3} + \frac{\mu_X^2 \sigma_M^4 \sigma_{(M-\mu_M)(X-\mu_X)}^2}{\sigma_{X,M}^4} \\ (A.2) \quad &= \boxed{Tr_X^2 \left[ \frac{\sigma_X^2}{\mu_X^2} + (\gamma_{2,M} - 1) + \frac{\sigma_{(X-\mu_X)(M-\mu_M)}^2}{\sigma_{X,M}^2} + 2 \frac{\sigma_{X,(M-\mu_M)^2}}{\sigma_M^2 \mu_X} \right.} \\ &\quad \left. - 2 \frac{\sigma_{X,(X-\mu_X)(M-\mu_M)}}{\sigma_{X,M} \mu_X} - 2 \frac{\sigma_{(M-\mu_M)^2,(X-\mu_X)(M-\mu_M)}}{\sigma_{X,M} \sigma_M^2} \right]}. \end{aligned}$$

<sup>13</sup>HAC = heteroskedasticity and autocorrelation consistent

In der letzten Umformung wird Formel (A.2) aus dem Anhang genutzt. Unter der Annahme unabhängig und identisch bivariat *normalverteilter Renditen* kann man die Formel mittels (A.10), (A.12) und (A.16) vereinfachen oder den Ansatz  $(\nabla g)' \Sigma_{NV} (\nabla g)$  mit

$$\Sigma_{NV} = \begin{pmatrix} \sigma_X^2 & 0 & 0 \\ \cdot & 2\sigma_M^4 & 2\sigma_M^2 \sigma_{X,M} \\ \cdot & \cdot & \sigma_X^2 \sigma_M^2 + \sigma_{X,M}^2 \end{pmatrix}$$

neu berechnen. Man erhält

$$n \sigma_{\widehat{Tr}, norm.}^2 = \frac{\sigma_M^4 \sigma_X^2}{\sigma_{X,M}^2} - \frac{\mu_X^2 \sigma_M^4}{\sigma_{X,M}^2} + \frac{\mu_X^2 \sigma_M^6 \sigma_X^2}{\sigma_{X,M}^4},$$

was sich vereinfachen lässt (mit  $VC_X =$  Variationskoeffizient von  $X$ ) zu

$$n \sigma_{\widehat{Tr}, norm.}^2 = Tr_X^2 \left( VC_X^2 + \frac{1}{\rho_{X,M}^2} - 1 \right).$$

Diese Formel weicht von der Formel in Jobson und Korkie (1981, S.893) leicht ab, wobei dort offenbar ein Druckfehler vorliegt.<sup>14</sup> Die Varianz steigt also mit steigendem Variationskoeffizienten und sinkt, je stärker Portfolio und Markt korreliert sind. Sind  $X$  und  $M$  unkorreliert, dann ist die Varianz nicht definiert, da dann der Nenner des Bruchs null wird.

Die **Differenz des Treynor-Ratios zweier Portfolios**  $\Delta_{Tr}$  wird mittels

$$\widehat{\Delta}_{Tr} = \widehat{Tr}_X - \widehat{Tr}_Y = \frac{\widehat{\mu}_X \widehat{\sigma}_M^2}{\widehat{\sigma}_{X,M}} - \frac{\widehat{\mu}_Y \widehat{\sigma}_M^2}{\widehat{\sigma}_{Y,M}} = S_M^2 \left( \frac{\overline{X}}{S_{X,M}} - \frac{\overline{Y}}{S_{Y,M}} \right)$$

geschätzt.

Wir haben nun einen trivariaten Zufallsvektor  $(M, X, Y)_k$ ,  $k = 1, \dots, n$  der (Über-)Renditen zweier Portfolios  $X$  und  $Y$  und eines Portfolios  $M$ , das für beide Portfolios gleichermaßen ein geeignetes Marktportfolio darstellt. Die  $(M, X, Y)_k$  seien unabhängig und identisch verteilt und die Momente  $E(M^u X^v Y^w)$  seien für alle  $\{(u, v, w) \mid u, v, w \in \mathbb{N}_0; u + v + w \leq 4\}$  existent. Außerdem seien  $\rho_{X,M} \neq 0$  und  $\rho_{Y,M} \neq 0$ . Der Ansatz ist nun:

$$\sqrt{n} \begin{pmatrix} \overline{X} - \mu_X \\ \overline{Y} - \mu_Y \\ S_M^2 - \sigma_M^2 \\ S_{X,M} - \sigma_{X,M} \\ S_{Y,M} - \sigma_{Y,M} \end{pmatrix} \sim N(\mathbf{0}, \Sigma) \quad \text{mit} \quad \Sigma =$$

<sup>14</sup>Im Artikel von Jobson und Korkie lautet die Varianzformel

$$\left( \sigma_i^2 + \mu_i^2 \left( 1 - \frac{1}{\rho_{im}^2} \right) \right) \frac{\sigma_m^4}{\sigma_{im}^2}, \quad \text{in unserer Notation: } Tr_X^2 \left( VC_X^2 + 1 - \frac{1}{\rho_{X,M}^2} \right).$$

Es fällt auf, dass hier der Ausdruck innerhalb der Klammer bei geringer Korrelation zwischen  $X$  und  $M$  negativ wird, wodurch der ganze Term negativ würde.

$$\begin{pmatrix} \sigma_X^2 & \sigma_{X,Y} & \sigma_{X,(M-\mu_M)}^2 & \sigma_{X,(X-\mu_X)(M-\mu_M)} & \sigma_{X,(Y-\mu_Y)(M-\mu_M)} \\ \cdot & \sigma_Y^2 & \sigma_{Y,(M-\mu_M)}^2 & \sigma_{Y,(X-\mu_X)(M-\mu_M)} & \sigma_{Y,(Y-\mu_Y)(M-\mu_M)} \\ \cdot & \cdot & \sigma_{(M-\mu_M)}^2 & \sigma_{(M-\mu_M)^2,(X-\mu_X)(M-\mu_M)} & \sigma_{(M-\mu_M)^2,(Y-\mu_Y)(M-\mu_M)} \\ \cdot & \cdot & \cdot & \sigma_{(X-\mu_X)(M-\mu_M)}^2 & \sigma_{(X-\mu_X)(M-\mu_M),(Y-\mu_Y)(M-\mu_M)} \\ \cdot & \cdot & \cdot & \cdot & \sigma_{(Y-\mu_Y)(M-\mu_M)}^2 \end{pmatrix},$$

$$\widehat{\Delta}_{Tr} = g(a, b, c, d, e) \quad \text{mit} \quad g = c \left( \frac{a}{d} - \frac{b}{e} \right) \quad \text{und}$$

$$\nabla g = \left( \frac{c}{d}, -\frac{c}{e}, \frac{a}{d} - \frac{b}{e}, -\frac{ac}{d^2}, \frac{bc}{e^2} \right)'.$$

Somit ist

$$n \sigma_{\widehat{\Delta}_{Tr}}^2 = \sum_{i=1}^5 \sum_{j=1}^5 (\nabla g)_i \Sigma_{ij} (\nabla g)_j,$$

ausgewertet an der Stelle ( $a = \mu_X$ ,  $b = \mu_Y$ ,  $c = \sigma_M^2$ ,  $d = \sigma_{X,M}$ ,  $e = \sigma_{Y,M}$ ). Für das (recht umfangreiche) ausmultiplizierte Ergebnis dieser Rechnung wird auf Formel (A.24) auf Seite 194 im Anhang verwiesen.

Nimmt man zusätzlich an, dass  $(M, X, Y)$  multivariat normalverteilt ist, vereinfacht sich  $\Sigma$  zu

$$\Sigma_{NV} = \begin{pmatrix} \sigma_X^2 & \sigma_{X,Y} & 0 & 0 & 0 \\ \cdot & \sigma_Y^2 & 0 & 0 & 0 \\ \cdot & \cdot & 2\sigma_M^4 & 2\sigma_M^2 \sigma_{X,M} & 2\sigma_M^2 \sigma_{Y,M} \\ \cdot & \cdot & \cdot & \sigma_M^2 \sigma_X^2 + \sigma_{X,M}^2 & \sigma_M^2 \sigma_{X,Y} + \sigma_{X,M} \sigma_{Y,M} \\ \cdot & \cdot & \cdot & \cdot & \sigma_M^2 \sigma_Y^2 + \sigma_{Y,M}^2 \end{pmatrix},$$

der Ansatz  $(\nabla g)' \Sigma_{NV} (\nabla g)$  wird ausmultipliziert zu

$$\begin{aligned} n \sigma_{\widehat{\Delta}_{Tr, norm.}}^2 &= \frac{\sigma_M^4 \sigma_X^2}{\sigma_{X,M}^2} - \frac{2\sigma_M^4 \sigma_{X,Y}}{\sigma_{X,M} \sigma_{Y,M}} + \frac{\sigma_M^4 \sigma_Y^2}{\sigma_{Y,M}^2} - \frac{\sigma_M^4 \mu_X^2}{\sigma_{X,M}^2} + \frac{2\mu_X \mu_Y \sigma_M^4}{\sigma_{X,M} \sigma_{Y,M}} \\ &\quad + \frac{\mu_X^2 \sigma_M^6 \sigma_X^2}{\sigma_{X,M}^4} - \frac{2\mu_X \sigma_M^6 \mu_Y \sigma_{X,Y}}{\sigma_{X,M}^2 \sigma_{Y,M}^2} - \frac{\sigma_M^4 \mu_Y^2}{\sigma_{Y,M}^2} + \frac{\mu_Y^2 \sigma_M^6 \sigma_Y^2}{\sigma_{Y,M}^4}. \end{aligned}$$

Nach der Zusammenfassung einiger Terme erhält man als Ergebnis

$$n \sigma_{\widehat{\Delta}_{Tr, norm.}}^2 = \boxed{Tr_X^2 \left( VC_X^2 + \frac{1}{\rho_{X,M}^2} - 1 \right) + Tr_Y^2 \left( VC_Y^2 + \frac{1}{\rho_{Y,M}^2} - 1 \right) - 2\rho_{X,Y} Tr_X Tr_Y \left( VC_X VC_Y + \frac{1}{\rho_{X,M} \rho_{Y,M}} - \frac{1}{\rho_{X,Y}} \right)}.$$

### 4.1.2.3 Jensens Alpha

Jensens Alpha ist das einzige hier betrachtete Performancemaß, für das schon seit seiner Definition eine Varianzformel existiert. Jensens Alpha ist die Konstante einer linearen Einfachregression der Portfolioüberrendite auf die Marktüberrendite. Für die Varianz des Schätzers der Konstanten  $\alpha$  einer linearen Regression  $X_i = \alpha + \beta M_i + \varepsilon_i$  gilt unter Annahme unabhängig und identisch normalverteilter Störterme  $\varepsilon_i$  die Formel

$$n \sigma_{\hat{\alpha}}^2 = \sigma_{\varepsilon}^2 \left( 1 + \frac{\mu_M^2}{\sigma_M^2} \right). \quad (4.6)$$

Das für Regressionen bei heteroskedastischen und autokorrelierten Residuen entwickelte Instrumentarium zur Varianzermittlung und zur Durchführung von Tests steht ebenfalls zur Verfügung. Dies hat dazu geführt, dass Tests auf überdurchschnittliche Performance zumeist anhand von Jensens Alpha oder dem Alpha von Faktormodellen durchgeführt werden.<sup>15</sup>

Zur Herleitung der Varianz des Schätzers von Jensens Alpha mittels der Delta-Methode werden ähnlich wie beim Treynor-Ratio folgende Annahmen getroffen:  $(X, M)_k$ ,  $k = 1, \dots, n$  seien unabhängig und identisch verteilt und die Momente  $E(X^u M^v)$  für alle  $\{(u, v) \mid u, v \in \mathbb{N}_0; u+v \leq 4\}$  existent. Der Schätzer für  $\alpha$  ist

$$\hat{\alpha} = \hat{\mu}_X - \hat{\beta}_{X,M} \hat{\mu}_M = \hat{\mu}_X - \frac{\hat{\mu}_M \hat{\sigma}_{X,M}}{\hat{\sigma}_M^2}.$$

Die Funktion  $g(a, b, c, d)$  und ihr Gradient sind

$$g = a - \frac{bd}{c}, \quad \nabla g = \left( 1, -\frac{d}{c}, \frac{bd}{c^2}, -\frac{b}{c} \right)'$$

Mittels der Asymptotik

$$\begin{aligned} & \sqrt{n} \begin{pmatrix} \bar{X} - \mu_X \\ \bar{M} - \mu_M \\ S_M^2 - \sigma_M^2 \\ S_{X,M} - \sigma_{X,M} \end{pmatrix} \\ & \sim N \left( \mathbf{0}, \begin{pmatrix} \sigma_X^2 & \sigma_{X,M} & \sigma_{X,(M-\mu_M)^2} & \sigma_{X,(X-\mu_X)(M-\mu_M)} \\ \cdot & \sigma_M^2 & \sigma_{M,(M-\mu_M)^2} & \sigma_{M,(X-\mu_X)(M-\mu_M)} \\ \cdot & \cdot & \sigma_{(M-\mu_M)^2}^2 & \sigma_{(M-\mu_M)^2,(X-\mu_X)(M-\mu_M)} \\ \cdot & \cdot & \cdot & \sigma_{(X-\mu_X)(M-\mu_M)}^2 \end{pmatrix} \right) \end{aligned}$$

<sup>15</sup>Beispiele: Cumby und Glen (1990), Barras et al. (2010)

wird die Varianz an der Stelle  $(a = \mu_X, b = \mu_M, c = \sigma_M^2, d = \sigma_{X,M})$  berechnet:

$$\begin{aligned}
n \sigma_{\hat{\alpha}}^2 &= \sum_{i=1}^4 \sum_{j=1}^4 (\nabla g)_i \Sigma_{ij} (\nabla g)_j \\
&= \sigma_X^2 - \frac{\sigma_{X,M}^2}{\sigma_M^2} + \frac{\mu_M \sigma_{X,M} \sigma_{X,(M-\mu_M)^2}}{\sigma_M^4} - \frac{2 \mu_M \sigma_{X,(X-\mu_X)(M-\mu_M)}}{\sigma_M^2} \\
&\quad - \frac{2 \sigma_{X,M}^2 \mu_M \sigma_{M,(M-\mu_M)^2}}{\sigma_M^6} + \frac{2 \mu_M \sigma_{X,M} \sigma_{M,(X-\mu_X)(M-\mu_M)}}{\sigma_M^4} \\
&\quad + \frac{\mu_M^2 \sigma_{X,M}^2 \sigma_{(M-\mu_M)^2}^2}{\sigma_M^8} - \frac{2 \mu_M^2 \sigma_{X,M} \sigma_{(M-\mu_M)^2, (X-\mu_X)(M-\mu_M)}}{\sigma_M^6} \\
&\quad + \frac{\mu_M^2 \sigma_{(X-\mu_X)(M-\mu_M)}}{\sigma_M^4}
\end{aligned}$$

Für bivariat normalverteilte  $(X, M)$  vereinfacht sich die Korrelationsmatrix im Grenzwertsatz zu

$$\Sigma_{NV} = \begin{pmatrix} \sigma_X^2 & \sigma_{X,M} & 0 & 0 \\ \cdot & \sigma_M^2 & 0 & 0 \\ \cdot & \cdot & 2\sigma_M^4 & 2\sigma_M^2 \sigma_{X,M} \\ \cdot & \cdot & \cdot & \sigma_M^2 \sigma_X^2 + \sigma_{X,M}^2 \end{pmatrix},$$

daraus ergibt sich die Varianz

$$\begin{aligned}
n \sigma_{\hat{\alpha}, norm.}^2 &= \sigma_X^2 - \frac{\sigma_{X,M}^2}{\sigma_M^2} + \frac{\mu_M^2 \sigma_X^2}{\sigma_M^2} - \frac{\mu_M^2 \sigma_{X,M}^2}{\sigma_M^4} \\
&= \sigma_X^2 (1 - \rho_{X,M}^2) + \frac{\mu_M^2}{\sigma_M^2} \left( \sigma_X^2 - \frac{\sigma_{X,M}^2}{\sigma_M^2} \right) \\
&= \sigma_X^2 (1 - \rho_{X,M}^2) + \sigma_X^2 \frac{\mu_M^2}{\sigma_M^2} (1 - \rho_{X,M}^2) \\
&= \boxed{\sigma_X^2 (1 - \rho_{X,M}^2) \left( 1 + \frac{\mu_M^2}{\sigma_M^2} \right)}.
\end{aligned}$$

Da  $\rho_{X,M}^2$  gleich dem  $R^2$  einer Regression von  $X$  auf  $M$  ist, ist  $\sigma_X^2 (1 - \rho_{X,M}^2)$  gleich der Störtermvarianz  $\sigma_\varepsilon^2$  und es ergibt sich Formel (4.6). Die Varianzherleitung mittels Delta-Methode liefert hier also dasselbe Ergebnis wie aus der Regressionsanalyse bekannt.

Für die **Varianz der Differenz von Jensens Alpha**  $\sigma_{\hat{\Delta}_\alpha}^2$  zwischen zwei Portfolios  $X$  und  $Y$  bezogen auf dasselbe Marktportfolio  $M$  wird wiederum ein trivariater Zufallsvektor  $(M, X, Y)_k$ ,  $k = 1, \dots, n$  angenommen,  $M$  sei ein für  $X$  und  $Y$  gleichermaßen geeignetes Marktportfolio. Die  $(M, X, Y)_k$  seien unabhängig und identisch verteilt und die Momente  $E(M^u X^v Y^w)$  für alle

$\{(u, v, w) \mid u, v, w \in \mathbb{N}_0; u + v + w \leq 4\}$  existent.

$$\begin{aligned}\widehat{\Delta}_\alpha &= \widehat{\alpha}_X - \widehat{\alpha}_Y = \widehat{\mu}_X - \frac{\widehat{\mu}_M \widehat{\sigma}_{X,M}}{\widehat{\sigma}_M^2} - \left( \widehat{\mu}_Y - \frac{\widehat{\mu}_M \widehat{\sigma}_{Y,M}}{\widehat{\sigma}_M^2} \right) \\ &= \widehat{\mu}_X - \widehat{\mu}_Y + \frac{\widehat{\mu}_M}{\widehat{\sigma}_M^2} (\widehat{\sigma}_{Y,M} - \widehat{\sigma}_{X,M})\end{aligned}$$

Zur Berechnung der Varianz  $\sigma_{\widehat{\Delta}_\alpha}^2$  wird die Funktion

$$g(a, b, c, d, e, f) = a - b + \frac{c}{d}(f - e)$$

mit Gradient

$$\nabla g = \left( 1, -1, \frac{f-e}{d}, -c \frac{f-e}{d^2}, -\frac{c}{d}, \frac{c}{d} \right)'$$

an der Stelle  $(a = \mu_X, b = \mu_Y, c = \mu_M, d = \sigma_M^2, e = \sigma_{X,M}, f = \sigma_{Y,M})$  ausgewertet. Es gilt:

$$\sqrt{n} \begin{pmatrix} \bar{X} - \mu_X \\ \bar{Y} - \mu_Y \\ \bar{M} - \mu_M \\ S_M^2 - \sigma_M^2 \\ S_{X,M} - \sigma_{X,M} \\ S_{Y,M} - \sigma_{Y,M} \end{pmatrix} \sim N(\mathbf{0}, \mathbf{\Sigma}).$$

Hierbei ist die Kovarianzmatrix  $\mathbf{\Sigma}$  völlig analog zu den obigen Verfahren:  $\mathbf{\Sigma}$  ist eine  $6 \times 6$ -Matrix mit den Elementen  $\Sigma_{ij} = \sigma_{K_i, K_j}$  mit  $\mathbf{K} = (X, Y, M, (M - \mu_M)^2, (X - \mu_X)(M - \mu_M), (Y - \mu_Y)(M - \mu_M))$ ;  $i, j = 1, 2, 3, 4, 5, 6$ . Die Varianz wird mit dem Term

$$n \sigma_{\widehat{\Delta}_\alpha}^2 = \sum_{i=1}^6 \sum_{j=1}^6 (\nabla g)_i \Sigma_{ij} (\nabla g)_j$$

berechnet, der ausmultiplizierte Term steht als Formel (A.28) auf S. 196 im Anhang. Für unabhängige und identisch multivariat normalverteilte  $(M, X, Y)_k$  vereinfacht sich die Kovarianzmatrix zu

$$\mathbf{\Sigma}_{NV} = \begin{pmatrix} \sigma_X^2 & \sigma_{X,Y} & \sigma_{X,M} & 0 & 0 & 0 \\ \cdot & \sigma_Y^2 & \sigma_{Y,M} & 0 & 0 & 0 \\ \cdot & \cdot & \sigma_M^2 & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & 2\sigma_M^4 & 2\sigma_M^2 \sigma_{X,M} & 2\sigma_M^2 \sigma_{Y,M} \\ \cdot & \cdot & \cdot & \cdot & \sigma_X^2 \sigma_M^2 + \sigma_{X,M}^2 & \sigma_M^2 \sigma_{X,Y} + \sigma_{X,M} \sigma_{Y,M} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \sigma_Y^2 \sigma_M^2 + \sigma_{Y,M}^2 \end{pmatrix}.$$

Die Varianz für normalverteilte Daten ist nach Auflösen der Formel

$$\begin{aligned}
\sigma_{\hat{\Delta}_\alpha, norm.}^2 &= \sigma_X^2 + \sigma_Y^2 - 2\sigma_{X,Y} - \frac{\sigma_{X,M}^2}{\sigma_M^2} - \frac{\sigma_{Y,M}^2}{\sigma_M^2} + \frac{2\sigma_{X,M}\sigma_{Y,M}}{\sigma_M^2} \\
&\quad + \frac{\mu_M^2\sigma_X^2}{\sigma_M^2} + \frac{\mu_M^2\sigma_Y^2}{\sigma_M^2} - \frac{2\mu_M^2\sigma_{X,Y}}{\sigma_M^2} \\
&\quad - \frac{\mu_M^2\sigma_{X,M}^2}{\sigma_M^4} - \frac{\mu_M^2\sigma_{Y,M}^2}{\sigma_M^4} + \frac{2\mu_M^2\sigma_{X,M}\sigma_{Y,M}}{\sigma_M^4} \\
&= \sigma_X^2 + \sigma_Y^2 - 2\sigma_{X,Y} - (\sigma_X^2\rho_{X,M}^2 + \sigma_Y^2\rho_{Y,M}^2 - 2\sigma_X\sigma_Y\rho_{X,Y} - 2\sigma_X\sigma_Y\rho_{X,M}\rho_{Y,M}) \\
&\quad + \frac{\mu_M^2}{\sigma_M^2} (\sigma_X^2 + \sigma_Y^2 - 2\sigma_{X,Y} - (\sigma_X^2\rho_{X,M}^2 + \sigma_Y^2\rho_{Y,M}^2 - 2\sigma_X\sigma_Y\rho_{X,Y} - 2\sigma_X\sigma_Y\rho_{X,M}\rho_{Y,M})) \\
&= \boxed{\left(1 + \frac{\mu_M^2}{\sigma_M^2}\right) (\sigma_X^2(1 - \rho_{X,M}^2) + \sigma_Y^2(1 - \rho_{Y,M}^2) - 2\sigma_X\sigma_Y(\rho_{X,Y} - \rho_{X,M}\rho_{Y,M}))}
\end{aligned}$$

#### 4.1.2.4 Modigliani-Modigliani

Das Vorgehen im Falle des Modigliani-Modigliani-Maßes ist vollkommen analog zu jenem beim Treynor-Ratio und Jensens Alpha. Wir unterscheiden wieder zwischen den beiden Versionen des Maßes

$$MM_{X,ori} = \mu_X \frac{\sigma_M}{\sigma_X} \quad \text{und} \quad MM_{X,alt} = \mu_X \frac{\sigma_M}{\sigma_X} - \mu_M$$

bzw. den entsprechenden Schätzern

$$\widehat{MM}_{X,ori} = \bar{X} \frac{\sqrt{S_M^2}}{\sqrt{S_X^2}} \quad \text{und} \quad \widehat{MM}_{X,alt} = \bar{X} \frac{\sqrt{S_M^2}}{\sqrt{S_X^2}} - \bar{M}$$

Beginnen wir mit dem Schätzer für die alternative Formulierung  $MM_{X,alt}$ , als Funktion  $g$ :

$$g(a, b, c, d) = a \frac{\sqrt{d}}{\sqrt{c}} - b, \quad \nabla g = \left( \frac{\sqrt{d}}{\sqrt{c}}, -1, -\frac{a\sqrt{d}}{2\sqrt{c}^3}, \frac{a}{2\sqrt{cd}} \right)'$$

An der Stelle  $(\mu_X, \mu_M, \sigma_X^2, \sigma_M^2)$  entspricht dies

$$\nabla g|_{a=\mu_X, b=\mu_M, c=\sigma_X^2, d=\sigma_M^2} = \left( \frac{\sigma_M}{\sigma_X}, -1, -\frac{\mu_X\sigma_M}{2\sigma_X^3}, \frac{\mu_X}{2\sigma_X\sigma_M} \right).$$

Mit den üblichen Annahmen  $(X, M)_k, k = 1, \dots, n$  i.i.d. und  $E(X^u M^v)$  existent für alle  $\{(u, v) \mid u, v \in \mathbb{N}_0; u + v \leq 4\}$  sowie dem Grenzwert

$$\sqrt{n} \begin{pmatrix} \bar{X} - \mu_X \\ \bar{M} - \mu_M \\ S_X^2 - \sigma_X^2 \\ S_M^2 - \sigma_M^2 \end{pmatrix} \sim N \left( \mathbf{0}, \begin{pmatrix} \sigma_X^2 & \sigma_{X,M} & \sigma_{X,(X-\mu_X)^2} & \sigma_{X,(M-\mu_M)^2} \\ \cdot & \sigma_M^2 & \sigma_{M,(X-\mu_X)^2} & \sigma_{M,(M-\mu_M)^2} \\ \cdot & \cdot & \sigma_{(X-\mu_X)^2}^2 & \sigma_{(X-\mu_X)^2,(M-\mu_M)^2} \\ \cdot & \cdot & \cdot & \sigma_{(M-\mu_M)^2}^2 \end{pmatrix} \right)$$

erhält man die Varianz

$$\begin{aligned}
n \widehat{\sigma_{MM}^2}_{alt} &= 2 \sigma_M^2 - \frac{2 \sigma_M \sigma_{X,M}}{\sigma_X} - \frac{\sigma_M^2 \mu_X \sigma_{X,(X-\mu_X)^2}}{\sigma_X^4} + \frac{\mu_X \sigma_{X,(M-\mu_M)^2}}{\sigma_X^2} \\
&+ \frac{\mu_X \sigma_M \sigma_{M,(X-\mu_X)^2}}{\sigma_X^3} - \frac{\mu_X \sigma_{M,(M-\mu_M)^2}}{\sigma_X \sigma_M} + \frac{\mu_X^2 \sigma_M^2 \sigma_{(X-\mu_X)^2}^2}{4 \sigma_X^6} \\
&- \frac{\mu_X^2 \sigma_{(X-\mu_X)^2,(M-\mu_M)^2}}{2 \sigma_X^4} + \frac{\mu_X^2 \sigma_{(M-\mu_M)^2}^2}{4 \sigma_X^2 \sigma_M^2} \\
&= 2 \sigma_M^2 \left( 1 - \frac{\rho_{X,M}}{\sigma_X^2} \right) + \frac{\mu_X}{\sigma_X} \left( -\sigma_M^2 \gamma_{1,X} + \frac{\sigma_{X,(M-\mu_M)^2}}{\sigma_X} \right. \\
&\quad \left. + \frac{\sigma_M \sigma_{M,(X-\mu_X)^2}}{\sigma_X^2} - \sigma_M^2 \gamma_{1,M} + \frac{\mu_X \sigma_M^2 (\gamma_{2,X} - 1)}{4 \sigma_X} \right. \\
&\quad \left. - \frac{\mu_X \sigma_{(X-\mu_X)^2,(M-\mu_M)^2}}{2 \sigma_X^3} + \frac{\mu_X \sigma_M^2 (\gamma_{2,M} - 1)}{4 \sigma_X} \right) \\
&= 2 \sigma_M^2 \left( 1 - \frac{\rho_{X,M}}{\sigma_X^2} \right) \\
&\quad + \frac{\mu_X}{\sigma_X} \left( \frac{\sigma_{X,(M-\mu_M)^2}}{\sigma_X} + \frac{\sigma_M \sigma_{M,(X-\mu_X)^2}}{\sigma_X^2} - \sigma_M^2 (\gamma_{1,X} + \gamma_{1,M}) \right) \\
&\quad + \frac{\mu_X^2}{\sigma_X^2} \left( \frac{\sigma_M^2}{4} (\gamma_{2,X} + \gamma_{2,M} - 2) - \frac{\sigma_{(X-\mu_X)^2,(M-\mu_M)^2}}{2 \sigma_X^2} \right).
\end{aligned}$$

Bei zusätzlicher Annahme einer bivariaten Normalverteilung für die  $(X, M)_k$  lautet die Kovarianzmatrix

$$\mathbf{\Sigma}_{NV} = \begin{pmatrix} \sigma_X^2 & \sigma_{X,M} & 0 & 0 \\ \cdot & \sigma_M^2 & 0 & 0 \\ \cdot & \cdot & 2 \sigma_X^4 & 2 \sigma_{X,M}^2 \\ \cdot & \cdot & \cdot & 2 \sigma_M^4 \end{pmatrix}$$

und die Varianz des Schätzers vereinfacht sich zu

$$\begin{aligned}
n \widehat{\sigma_{MM}^2}_{alt,norm.} &= 2 \sigma_M^2 - \frac{2 \sigma_M \sigma_{X,M}}{\sigma_X} + \frac{\mu_X^2 \sigma_M^2}{\sigma_X^2} - \frac{\mu_X^2 \sigma_{X,M}^2}{\sigma_X^4} \\
&= \boxed{\sigma_M^2 \left[ 2 (1 - \rho_{X,M}) + \frac{\mu_X^2}{\sigma_X^2} (1 - \rho_{X,M}^2) \right]}. \quad (4.7)
\end{aligned}$$

Die Varianz steigt also mit der Varianz der Marktrendite und sinkt mit steigender Korrelation zwischen Portfolio und Markt, sie geht für  $\sigma_M^2 \rightarrow 0$  sowie  $\rho_{X,M} \rightarrow 1$  (und natürlich  $n \rightarrow \infty$ ) gegen null. Aus dem zweiten Summanden erkennt man eine steigende Varianz, wenn die (erwartete) Rendite des Portfolios stärker vom risikolosen Zinssatz abweicht. Da das Verhältnis  $\mu_X^2 / \sigma_X^2 (= Sh_X^2)$  in der Praxis deutlich kleiner als eins sein sollte, dürfte die Varianz – mit Ausnahme von Fällen fast perfekter Korrelation zwischen  $X$  und  $M$  – in erster Linie



von  $\sigma_M^2$ ,  $\rho_{X,M}$  und  $n$  abhängen.

Im der Originalversion von Modigliani-Modigliani fehlt die Subtraktion des Erwartungswerts der Marktrendite  $\mu_M$  und die Funktion  $g$  vereinfacht sich zu

$$g(a, b, c) = a \frac{\sqrt{c}}{\sqrt{b}}, \quad \nabla g = \left( \frac{\sqrt{c}}{\sqrt{b}}, -\frac{a\sqrt{c}}{2\sqrt{b}^3}, \frac{a}{2\sqrt{bc}} \right)',$$

was an der Stelle  $(\mu_X, \sigma_X^2, \sigma_M^2)$  bedeutet:

$$\nabla g|_{a=\mu_X, b=\sigma_X^2, c=\sigma_M^2} = \left( \frac{\sigma_M}{\sigma_X}, -\frac{\mu_X \sigma_M}{2\sigma_X^3}, \frac{\mu_X}{2\sigma_X \sigma_M} \right).$$

Die Annahmen an  $(X, M)_k$  bleiben unverändert, aus dem o. a. Grenzwert werden Zeile und Spalte der erwarteten Marktrendite gestrichen:

$$\sqrt{n} \begin{pmatrix} \bar{X} - \mu_X \\ S_X^2 - \sigma_X^2 \\ S_M^2 - \sigma_M^2 \end{pmatrix} \sim N \left( \mathbf{0}, \begin{pmatrix} \sigma_X^2 & \sigma_{X,(X-\mu_X)^2} & \sigma_{X,(M-\mu_M)^2} \\ \cdot & \sigma_{(X-\mu_X)^2}^2 & \sigma_{(X-\mu_X)^2,(M-\mu_M)^2} \\ \cdot & \cdot & \sigma_{(M-\mu_M)^2}^2 \end{pmatrix} \right)$$

$$\begin{aligned} n \widehat{\sigma_{MM}^2}_{ori} &= \sigma_M^2 - \frac{\sigma_M^2 \mu_X \sigma_{X,(X-\mu_X)^2}}{\sigma_X^4} + \frac{\mu_X \sigma_{X,(M-\mu_M)^2}}{\sigma_X^2} + \frac{\mu_X^2 \sigma_M^2 \sigma_{(X-\mu_X)^2}^2}{4\sigma_X^6} \\ &\quad - \frac{\mu_X^2 \sigma_{(X-\mu_X)^2,(M-\mu_M)^2}}{2\sigma_X^4} + \frac{\mu_X^2 \sigma_{(M-\mu_M)^2}^2}{4\sigma_X^2 \sigma_M^2} \\ &= \frac{\sigma_M^2}{\sigma_X^2} \left( \sigma_X^2 - \mu_X \sigma_X \gamma_{1,X} + \frac{\mu_X \sigma_{X,(M-\mu_M)^2}}{\sigma_M^2} \right. \\ &\quad \left. - \frac{\mu_X^2 \sigma_{(X-\mu_X)^2,(M-\mu_M)^2}}{2\sigma_M^2 \sigma_X^2} + \frac{1}{4} \mu_X^2 (\gamma_{2,X} + \gamma_{2,M} - 2) \right) \end{aligned}$$

Wenn wir nun zusätzlich wieder multivariat normalverteilte  $(X, M)_k$  annehmen, lautet die Kovarianzmatrix

$$\Sigma_{NV} = \begin{pmatrix} \sigma_X^2 & 0 & 0 \\ \cdot & 2\sigma_X^4 & 2\sigma_{X,M}^2 \\ \cdot & \cdot & 2\sigma_M^4 \end{pmatrix}$$

und wir erhalten als Formel für die Schätzvarianz

$$\begin{aligned} n \widehat{\sigma_{MM}^2}_{ori, norm.} &= \sigma_M^2 + \frac{\mu_X^2 \sigma_M^2}{\sigma_X^2} - \frac{\mu_X^2 \sigma_{X,M}^2}{\sigma_X^4} \\ &= \sigma_M^2 + \mu_X^2 \frac{\sigma_M^2}{\sigma_X^2} (1 - \rho_{X,M}^2) \\ &= \sigma_M^2 [1 + Sh_X^2 (1 - \rho_{X,M}^2)] \\ &= MM_{ori}^2 \left( \frac{\sigma_X^2}{\mu_X^2} + 1 - \rho_{X,M}^2 \right) \end{aligned} \tag{4.8}$$

Die Schätzvarianz hängt also zunächst von der Streuung des Marktes ab und erhöht sich (bei gegebenem  $\rho_{X,M}$ ) mit steigendem Betrag der Überrendite von  $X$ . Durch einfache Umformungen erkennt man, dass bei normalverteilten Renditen die Varianz von  $\widehat{MM}_{ori}$  (4.8) genau dann größer ist als die Varianz von  $\widehat{MM}_{alt}$  (4.7), wenn  $\rho_{X,M} > 0,5$ . Wenn  $X$  und  $M$  also stark positiv korreliert sind, ist die Schätzvarianz bei der alternativen Formulierung von MM kleiner als die bei Originalformulierung.

Für die Herleitung der **Varianz des Schätzers der Differenz von MM** bei zwei Portfolios  $X$  und  $Y$  mit gemeinsamem Marktportfolio  $M$  treffen wir wieder die üblichen Annahmen, dass die trivariaten Zufallsvariablen  $(M, X, Y)_k$ ,  $k = 1, \dots, n$  unabhängig und identisch verteilt seien und alle Momente  $E(M^u X^v Y^w)$  mit  $\{(u, v, w) \mid u, v, w \in \mathbb{N}_0; u + v + w \leq 4\}$  existieren. Eine Unterscheidung zwischen den beiden unterschiedlichen Versionen von MM ist nicht mehr nötig, da sich deren Unterschied bei der Differenzbildung aufhebt:

$$\begin{aligned} MM_{X,alt} - MM_{Y,alt} &= \mu_X \frac{\sigma_M}{\sigma_X} - \mu_M - \left( \mu_Y \frac{\sigma_M}{\sigma_Y} - \mu_M \right) = \\ &= \mu_X \frac{\sigma_M}{\sigma_X} - \mu_Y \frac{\sigma_M}{\sigma_Y} = MM_{X,ori} - MM_{Y,ori} \end{aligned}$$

Der Schätzer für die Differenz  $\Delta_{MM} = MM_X - MM_Y$  ist

$$\begin{aligned} \widehat{\Delta}_{MM} &= \widehat{MM}_X - \widehat{MM}_Y = \widehat{\mu}_X \frac{\widehat{\sigma}_M}{\widehat{\sigma}_X} - \widehat{\mu}_M - \left( \widehat{\mu}_Y \frac{\widehat{\sigma}_M}{\widehat{\sigma}_Y} - \widehat{\mu}_M \right) \\ &= \widehat{\sigma}_M \left( \frac{\widehat{\mu}_X}{\widehat{\sigma}_X} - \frac{\widehat{\mu}_Y}{\widehat{\sigma}_Y} \right) \quad \left( = \widehat{\sigma}_M \left( \widehat{Sh}_X - \widehat{Sh}_Y \right) \right) \\ &= S_M \left( \frac{\bar{X}}{S_X} - \frac{\bar{Y}}{S_Y} \right). \end{aligned}$$

Die Differenz der MM zweier Portfolios ist also gleich der Differenz der Sharpe-Ratios der beiden Portfolios, multipliziert mit der Standardabweichung der Rendite des gemeinsamen Marktportfolios.

$$\begin{aligned} g(a, b, c, d, e) &= \sqrt{e} \left( \frac{a}{\sqrt{c}} - \frac{b}{\sqrt{d}} \right) \\ \nabla g &= \left( \frac{\sqrt{e}}{\sqrt{c}}, -\frac{\sqrt{e}}{\sqrt{d}}, -\frac{a\sqrt{e}}{2\sqrt{c^3}}, \frac{b\sqrt{e}}{2\sqrt{d^3}}, \frac{1}{2\sqrt{e}} \left( \frac{a}{\sqrt{c}} - \frac{b}{\sqrt{d}} \right) \right)' \end{aligned}$$

Der Ansatz lautet nun

$$\sqrt{n} \begin{pmatrix} \bar{X} - \mu_X \\ \bar{Y} - \mu_Y \\ S_X^2 - \sigma_X^2 \\ S_Y^2 - \sigma_Y^2 \\ S_M^2 - \sigma_M^2 \end{pmatrix} \sim N(\mathbf{0}, \Sigma) \quad \text{mit}$$

$$\Sigma = \begin{pmatrix} \sigma_X^2 & \sigma_{X,Y} & \sigma_{X,(X-\mu_X)^2} & \sigma_{X,(Y-\mu_Y)^2} & \sigma_{X,(M-\mu_M)^2} \\ \cdot & \sigma_Y^2 & \sigma_{Y,(X-\mu_X)^2} & \sigma_{Y,(Y-\mu_Y)^2} & \sigma_{Y,(M-\mu_M)^2} \\ \cdot & \cdot & \sigma_{(X-\mu_X)^2}^2 & \sigma_{(X-\mu_X)^2,(Y-\mu_Y)^2} & \sigma_{(X-\mu_X)^2,(M-\mu_M)^2} \\ \cdot & \cdot & \cdot & \sigma_{(Y-\mu_Y)^2}^2 & \sigma_{(Y-\mu_Y)^2,(M-\mu_M)^2} \\ \cdot & \cdot & \cdot & \cdot & \sigma_{(M-\mu_M)^2}^2 \end{pmatrix}.$$

An der Stelle  $\nabla g|(a = \mu_X, b = \mu_Y, c = \sigma_X^2, d = \sigma_Y^2, e = \sigma_M^2)$  erhält man den Varianzterm

$$n \sigma_{\hat{\Delta}_{MM}}^2 = \sum_{i=1}^5 \sum_{j=1}^5 (\nabla g)_i \Sigma_{ij} (\nabla g)_j$$

der wiederum umfangreich ist, so dass hier auf Formel (A.34) im Anhang auf S. 197 verwiesen wird.

Sind die  $(M, X, Y)_k$  zusätzlich multivariat normalverteilt, beträgt die Kovarianzmatrix

$$\Sigma_{NV} = \begin{pmatrix} \sigma_X^2 & \sigma_{X,Y} & 0 & 0 & 0 \\ \cdot & \sigma_Y^2 & 0 & 0 & 0 \\ \cdot & \cdot & 2\sigma_X^4 & 2\sigma_{X,Y}^2 & 2\sigma_{X,M}^2 \\ \cdot & \cdot & \cdot & 2\sigma_Y^4 & 2\sigma_{Y,M}^2 \\ \cdot & \cdot & \cdot & \cdot & 2\sigma_M^4 \end{pmatrix}$$

und die gesuchte Varianz für unabhängig und identisch normalverteilte Daten ist ( $Sh_A =$  Sharpe-Ratio von  $A$ )

$$\begin{aligned} n \sigma_{\hat{\Delta}_{MM},norm.}^2 &= 2\sigma_M^2 - \frac{2\sigma_M^2\sigma_{X,Y}}{\sigma_X\sigma_Y} + \frac{\mu_X^2\sigma_M^2}{\sigma_X^2} + \frac{\mu_Y^2\sigma_M^2}{\sigma_Y^2} - \frac{\mu_X\mu_Y\sigma_M^2}{\sigma_X\sigma_Y} \\ &\quad - \frac{\mu_X\mu_Y\sigma_M^2\sigma_{X,Y}^2}{\sigma_X^3\sigma_Y^3} - \frac{\mu_X^2\sigma_{X,M}^2}{\sigma_X^4} - \frac{\mu_Y^2\sigma_{Y,M}^2}{\sigma_Y^4} \\ &\quad + \frac{\mu_X\mu_Y\sigma_{X,M}^2}{\sigma_X^3\sigma_Y} + \frac{\mu_X\mu_Y\sigma_{Y,M}^2}{\sigma_X\sigma_Y^3} \\ &= \sigma_M^2 (2 - 2\rho_{X,Y} + Sh_X^2 + Sh_Y^2 - Sh_X Sh_Y \\ &\quad - Sh_X Sh_Y \rho_{X,Y}^2 - Sh_X^2 \rho_{X,M}^2 - Sh_Y^2 \rho_{Y,M}^2 \\ &\quad + Sh_X Sh_Y \rho_{X,M}^2 + Sh_X Sh_Y \rho_{Y,M}^2) \\ &= \boxed{\sigma_M^2 (2(1 - \rho_{X,Y}) + Sh_X^2 (1 - \rho_{X,M}^2) + Sh_Y^2 (1 - \rho_{Y,M}^2) \\ &\quad - Sh_X Sh_Y (1 + \rho_{X,Y}^2 - \rho_{X,M}^2 - \rho_{Y,M}^2))} \end{aligned}$$

In der Tat ist dieser Ausdruck<sup>16</sup> der Formel für die Varianz der Differenz zweier Sharpe-Ratios (Formel 4.4 auf S. 75) sehr ähnlich. Setzt man in  $\sigma_{\hat{\Delta}_{MM},norm.}^2$  die Werte  $\sigma_M^2 = 1$  und  $\rho_{X,M} = \rho_{Y,M} = \sqrt{0,5}$  ein, ergibt sich  $\sigma_{\hat{\Delta}_{Sh},norm.}^2$ :

<sup>16</sup>Im Anhang findet man diese Formel ohne Zusammenfassung von Termen zu Sharpe-Ratios als Formel (A.35), siehe S. 198.

### 4.1.2.5 Kappa

Die Kappa-Maße  $\kappa_{\alpha,z} = \frac{\mu_{X-z}}{\sqrt[\alpha]{E(z-X)_+^\alpha}}$  sind nur definiert, wenn der Nenner nicht null ist, also  $E(z-X)_+^\alpha > 0$ , was erfüllt ist, wenn  $P(X < z) > 0$ . Zusätzlich müssen die ersten  $\alpha$  Momente von  $X$  existieren.<sup>17</sup> Damit der Schätzer

$$\hat{\kappa}_{\alpha,z} = \frac{\bar{X} - z}{\sqrt[\alpha]{\hat{E}((z-X)_+^\alpha)}} = \frac{\bar{X} - z}{\sqrt[\alpha]{\frac{1}{n} \sum_{i=1}^n (\max(z - X_i, 0))^\alpha}}$$

ein Ergebnis liefert, muss auch dessen Nenner von null verschieden sein. Hierfür muss zumindest  $X_{(1)}$ , der kleinste Wert der Stichprobe, kleiner als  $z$  sein. Wenn  $P(X_{(1)} < z) = 1$ , dann ist der Schätzwert stets definiert.

Für den *i.i.d.*-Fall gilt:

$$P(X_{(1)} < z) = 1 - P(X_{(1)} \geq z) = 1 - (P(X \geq z))^n$$

Ist  $P(X < z) > 0$ , dann folgt der Grenzwert  $\lim_{n \rightarrow \infty} P(X_{(1)} < z) = 1$ .

Somit können die Formeln für die Varianz des Schätzers wie gewohnt hergeleitet werden. Dies erfolgt hier zunächst für  $\kappa_{\alpha,z}$  allgemein; aus der sich ergebenden Formel werden anschließend die Werte  $\alpha = 1$  bzw.  $\alpha = 2$  für Omega und Sortino-Ratio eingesetzt. Neben der Annahme, dass die Elemente der Folge  $X_k$ ,  $k = 1, \dots, n$  unabhängig und identisch verteilt sind, muss für die Herleitung der Varianz von  $\kappa_{\alpha,z}$  verlangt werden, dass mindestens die ersten  $2\alpha$  Momente existieren. Wir nutzen die Asymptotik

$$\sqrt{n} \begin{pmatrix} (\bar{X} - z) - E(X - z) \\ \hat{E}((z - X)_+^\alpha) - E((z - X)_+^\alpha) \end{pmatrix} \sim N \left( \mathbf{0}, \mathbf{\Sigma} = \begin{pmatrix} \sigma_{X-z}^2 & \sigma_{X-z, (z-X)_+^\alpha} \\ \cdot & \sigma_{(z-X)_+^\alpha}^2 \end{pmatrix} \right).$$

Die Funktion  $g(a, b) = a \cdot b^{-\frac{1}{\alpha}}$  hat den Gradienten  $\nabla g = \left( b^{-\frac{1}{\alpha}}, -\frac{1}{\alpha} a b^{-1-\frac{1}{\alpha}} \right)'$  und wird im folgenden Delta-Ansatz an der Stelle  $a = E(X - z)$  und  $b = E((z - X)_+^\alpha)$  ausgewertet.

<sup>17</sup>Da sich partielle Momente von gewöhnlichen Momenten lediglich dadurch unterscheiden, dass bei ihnen nur über Teile des Trägers von  $X$  statt über den gesamten Träger integriert wird, folgt aus der Existenz eines gewöhnlichen Moments  $k$ -ter Ordnung die Existenz der partiellen Momente  $k$ -ter Ordnung.

Die gesuchte Varianz ist entsprechend

$$\begin{aligned}
 n \sigma_{\hat{\kappa}(\alpha, z)}^2 &= \sum_{i=1}^2 \sum_{j=1}^2 (\nabla g)_i \Sigma_{ij} (\nabla g)_j \\
 &= \frac{\sigma_{X-z}^2}{(E((z-X)_+^\alpha))^{\frac{2}{\alpha}}} + \frac{1}{\alpha^2} \frac{(E(X-z))^2 \sigma_{(z-X)_+^\alpha}^2}{(E((z-X)_+^\alpha))^{2+\frac{2}{\alpha}}} \\
 &\quad - \frac{2}{\alpha} \frac{E(X-z) \sigma_{X-z, (z-X)_+^\alpha}}{(E((z-X)_+^\alpha))^{1+\frac{2}{\alpha}}} \\
 &= \kappa_{\alpha, z}^2 \left( \frac{\sigma_{X-z}^2}{(E(X-z))^2} + \frac{1}{\alpha^2} \frac{\sigma_{(z-X)_+^\alpha}^2}{(E((z-X)_+^\alpha))^2} \right. \\
 &\quad \left. - \frac{2}{\alpha} \frac{\sigma_{X-z, (z-X)_+^\alpha}}{E(X-z) E((z-X)_+^\alpha)} \right).
 \end{aligned}$$

Durch Auflösen der (Ko-)Varianzen nach dem Muster

$$\frac{\sigma_{A,B}}{c \cdot E(A)E(B)} = \frac{E(AB) - E(A)E(B)}{c \cdot E(A)E(B)} = \frac{E(AB)}{c \cdot E(A)E(B)} - \frac{1}{c},$$

das Zusammenfassen der Subtrahenden der sich daraus ergebenden Terme sowie die Beziehung

$$\begin{aligned}
 (X-z)(z-X)_+^\alpha &= [(X-z)_+ - (z-X)_+](z-X)_+^\alpha \\
 &= \underbrace{=0, \text{ wenn } X \leq z}_{(X-z)_+} \cdot \underbrace{=0, \text{ wenn } X \geq z}_{(z-X)_+^\alpha} - (z-X)_+ \cdot (z-X)_+^\alpha \\
 &= \underbrace{(X-z)_+ \cdot (z-X)_+^\alpha}_{=0} - (z-X)_+ \cdot (z-X)_+^\alpha \\
 &= -(z-X)_+^{\alpha+1}
 \end{aligned}$$

erhält man die Formel

$$\begin{aligned}
 n \sigma_{\kappa_{\alpha, z}}^2 &= \kappa_{\alpha, z}^2 \left( \frac{E((X-z)^2)}{(E(X-z))^2} + \frac{2}{\alpha} \frac{E((z-X)_+^{\alpha+1})}{E(X-z)E((z-X)_+^\alpha)} \right. \\
 &\quad \left. + \frac{1}{\alpha^2} \frac{E((z-X)_+^{2\alpha})}{(E((z-X)_+^\alpha))^2} - \left(1 - \frac{1}{\alpha}\right)^2 \right).
 \end{aligned}$$

Durch Einsetzen von  $\alpha = 1$  folgt die Formel für die **Varianz des Schätzers**

von  $\kappa_{1,z}$

$$\begin{aligned}
n\sigma_{\kappa_{1,z}}^2 &= \kappa_{1,z}^2 \left( \frac{E((X-z)^2)}{(E(X-z))^2} + \frac{2E((z-X)_+^2)}{E(X-z)E((z-X)_+)} \right. \\
&\quad \left. + \frac{E((z-X)_+^2)}{(E((z-X)_+))^2} - 0 \right) \\
&= \frac{(E(X-z))^2}{(E((z-X)_+))^2} \left( \dots \right) \\
&= \frac{E((X-z)^2)}{(E((z-X)_+))^2} + \frac{2E(X-z)E((z-X)_+^2)}{(E((z-X)_+))^3} \\
&\quad + \frac{(E(X-z))^2 E((z-X)_+^2)}{(E((z-X)_+))^4}. \tag{4.9}
\end{aligned}$$

Da  $\kappa_{1,z} = \Omega_z - 1$  ist, gilt diese Varianzformel auch für  $\Omega_z$ . Speziell für **Omega** haben Schmid und Schmidt (2008) eine kompakte Varianzformel hergeleitet:

$$n\sigma_{\Omega_z}^2 = \Omega_z^2 \left( \frac{E((X-z)_+^2)}{(E((X-z)_+))^2} + \frac{E((z-X)_+^2)}{(E((z-X)_+))^2} \right) \tag{4.10}$$

Formel (4.9) lässt sich durch eine Reihe einfacher Umformungen in diese Formel überführen, siehe den Abschnitt A.3.1 im Anhang.

Für  $\alpha = 2$  ergibt sich die Formel für die **Varianz des Schätzers des Sortino-Ratios** (Sor= $\kappa_2$ )

$$\begin{aligned}
n\sigma_{\text{Sor}_z}^2 &= \text{Sor}_z \left( \frac{E((X-z)^2)}{(E(X-z))^2} + \frac{2}{2} \frac{E((z-X)_+^3)}{E(X-z)E((z-X)_+^2)} \right. \\
&\quad \left. + \frac{1}{2^2} \frac{E((z-X)_+^4)}{(E((z-X)_+^2))^2} - \left(1 - \frac{1}{2}\right)^2 \right) \\
&= \frac{(E(X-z))^2}{E((z-X)_+^2)} \left( \frac{E((X-z)^2)}{(E(X-z))^2} + \frac{E((z-X)_+^3)}{E(X-z)E((z-X)_+^2)} \right. \\
&\quad \left. + \frac{E((z-X)_+^4)}{4(E((z-X)_+^2))^2} - \frac{1}{4} \right) \\
&= \frac{E((X-z)^2)}{E((z-X)_+^2)} + \frac{E(X-z)E((z-X)_+^3)}{(E((z-X)_+^2))^2} \\
&\quad + \frac{(E(X-z))^2 E((z-X)_+^4)}{4(E((z-X)_+^2))^3} - \frac{(E(X-z))^2}{4E((z-X)_+^2)}
\end{aligned}$$

**Vergleich zweier Kappas:** Das Performancemaß Kappa hat anders als die vorher untersuchten Maße zwei Parameter. Beim Vergleich zweier Portfolios anhand von Kappa ist man nicht darauf beschränkt, für beide Portfolios dieselben Parameter zu wählen, prinzipiell könnte man für die Portfolios  $X$  und  $Y$  auch

die Maße  $\kappa_{\alpha_X, z_X}$  und  $\kappa_{\alpha_Y, z_Y}$  mit jeweils angepassten Parametern vergleichen. Dabei sollte man von der Verwendung unterschiedlicher  $\alpha$  absehen, da Kappas mit unterschiedlichen  $\alpha$  letztlich unterschiedliche Performancemaße sind. Für den Vergleich zweier Portfolios mit unterschiedlichen Mindestrenditen  $z_X, z_Y$  sind dagegen sinnvolle Anwendungen zumindest denkbar.<sup>18</sup>

Für die **Varianz von**

$$\begin{aligned} \hat{\Delta}_{\kappa_{\alpha, z_X, z_Y}} &= \hat{\kappa}_{\alpha, z_X}(X) - \hat{\kappa}_{\alpha, z_Y}(Y) \\ &= \frac{\bar{X} - z_X}{\sqrt{\frac{1}{n} \sum_{i=1}^n (\max(z_X - X_i, 0))^\alpha}} - \frac{\bar{Y} - z_Y}{\sqrt{\frac{1}{n} \sum_{i=1}^n (\max(z_Y - Y_i, 0))^\alpha}}, \end{aligned}$$

der geschätzten Differenz zweier Kappas, sei hier nur der Ansatz gezeigt, da sich die ausmultiplizierte quadratische Form nur unzureichend vereinfachen lässt. Für die  $(X, Y)_k, k = 1, \dots, n$  Renditepaare der Portfolios  $X$  und  $Y$  wird wiederum angenommen, sie seien unabhängig und identisch verteilt und die (gemischten) Momente  $E(X^u Y^v)$  für alle  $\{(u, v) \mid u, v \in \mathbb{N}_0; u + v \leq 4\}$  existent. Zusätzlich seien  $P(X < z_X) > 0$  und  $P(Y < z_Y) > 0$ .

Dann gilt

$$\begin{aligned} &\sqrt{n} \begin{pmatrix} (\bar{X} - z_X) - E(X - z_X) \\ \hat{E}((z_X - X)_+^\alpha) - E((z_X - X)_+^\alpha) \\ (\bar{Y} - z_Y) - E(Y - z_Y) \\ \hat{E}((z_Y - Y)_+^\alpha) - E((z_Y - Y)_+^\alpha) \end{pmatrix} \sim \\ N \left( \mathbf{0}, \Sigma = \begin{pmatrix} \sigma_{X-z_X}^2 & \sigma_{X-z_X, (z_X-X)_+^\alpha} & \sigma_{X-z_X, Y-z_Y} & \sigma_{X-z_X, (z_Y-Y)_+^\alpha} \\ \cdot & \sigma_{(z_X-X)_+^\alpha}^2 & \sigma_{(z_X-X)_+^\alpha, Y-z_Y} & \sigma_{(z_X-X)_+^\alpha, (z_Y-Y)_+^\alpha} \\ \cdot & \cdot & \sigma_{Y-z_Y}^2 & \sigma_{Y-z_Y, (z_Y-Y)_+^\alpha} \\ \cdot & \cdot & \cdot & \sigma_{(z_Y-Y)_+^\alpha}^2 \end{pmatrix} \right). \end{aligned}$$

Die Funktion  $g(a, b, c, d) = \frac{a}{b} - \frac{c}{d}$  hat den Gradienten  $\nabla g = (\frac{1}{b}, -\frac{a}{b^2}, -\frac{1}{d}, \frac{c}{d^2})$ , der an der Stelle

$$\nabla g|_{(a = \mu_X - z_X, b = E((z_X - X)_+^\alpha), c = \mu_Y - z_Y, d = E((z_Y - Y)_+^\alpha))}$$

ausgewertet wird. Die gesuchte Varianz kann dann mittels

$$n \sigma_{\hat{\Delta}_{\kappa_{\alpha, z_X, z_Y}}}^2 = \sum_{i=1}^4 \sum_{j=1}^4 (\nabla g)_i \Sigma_{ij} (\nabla g)_j \quad (4.11)$$

berechnet werden.

<sup>18</sup>Man stelle sich vor, die Wertpapiere  $X$  und  $Y$  seien in verschiedenen Währungen notiert. Rechnet man zwecks Vergleich die Renditen von  $Y$  zu den jeweiligen Wechselkursen in die Währung von  $X$  um, so lastet das Währungsrisiko einseitig auf Portfolio  $Y$ . Statt dessen könnte man die Renditen der Ausgangswährungen verwenden und als  $z_X$  und  $z_Y$  die sicheren Renditen der beiden Währungen wählen.

#### 4.1.2.6 Excess Return on Value-at-Risk

Die Delta-Methode lässt sich auch zur Herleitung der asymptotischen Varianz des ERVaR nutzen. Der Schätzer des ERVaR $_{\alpha}$  ist

$$\widehat{ERVaR}_{\alpha} = \frac{\bar{X}}{-\hat{x}_{\alpha}},$$

also haben die Funktion  $g(a, b) = \frac{a}{-b}$  und deren Gradient  $\nabla g = \left(-\frac{1}{b}, \frac{a}{b^2}\right)'$  eine vergleichsweise einfache Form. Der Grenzwert zur Herleitung der Varianz ist

$$\sqrt{n} \begin{pmatrix} \bar{X} - \mu_X \\ \hat{x}_{\alpha} - x_{\alpha} \end{pmatrix} \sim N \left( \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \Sigma = \begin{pmatrix} \sigma_X^2 & \sigma_{X, \hat{x}_{\alpha}} \\ \sigma_{X, \hat{x}_{\alpha}} & \sigma_{\hat{x}_{\alpha}}^2 \end{pmatrix} \right) \quad (4.12)$$

Ferguson (1999) leitet diesen Grenzwert und eine Formel für die asymptotische Kovarianz von Mittelwert und Quantil  $\sigma_{X, \hat{x}_{\alpha}}$  her:

$$\sigma_{X, \hat{x}_{\alpha}} = \frac{\tau(\alpha)}{f(x_{\alpha})}$$

Im Nenner steht mit  $f(x_{\alpha})$  die Dichte der Verteilung von  $X$  an der Stelle des Quantils  $x_{\alpha}$ . Der Ausdruck  $\tau(\alpha)$  ist nach Ferguson

$$\tau(\alpha) = E(L_{\alpha}(X - x_{\alpha}))$$

mit

$$L_{\alpha}(X - x_{\alpha}) = \begin{cases} \alpha(x - x_{\alpha}) & \text{wenn } x > x_{\alpha}, \\ (1 - \alpha)(x_{\alpha} - x) & \text{wenn } x \leq x_{\alpha}. \end{cases}$$

Zwecks besserer Handhabbarkeit soll die Fallunterscheidung in diesem Term auf einfache Beträge reduziert werden:

$$\begin{aligned} L_{\alpha}(X - x_{\alpha}) &= \max(\alpha(x - x_{\alpha}), 0) + \max((1 - \alpha)(x_{\alpha} - x), 0) \\ &= [0, 5 \alpha(x - x_{\alpha}) + 0, 5 \alpha|x - x_{\alpha}|] \\ &\quad + [0, 5(1 - \alpha)(x_{\alpha} - x) + 0, 5(1 - \alpha)|x_{\alpha} - x|] \\ &= (0, 5 \alpha - 0, 5(1 - \alpha))(x - x_{\alpha}) + (0, 5 \alpha + 0, 5(1 - \alpha))|x - x_{\alpha}| \\ &= (\alpha - 0, 5)(x - x_{\alpha}) + 0, 5|x - x_{\alpha}| \end{aligned}$$

Der Erwartungswert  $\tau(\alpha) = E(L_{\alpha}(X - x_{\alpha}))$  beträgt somit

$$\begin{aligned} \tau(\alpha) &= E[(\alpha - 0, 5)(X - x_{\alpha}) + 0, 5|X - x_{\alpha}|] \\ &= 0, 5 E(|X - x_{\alpha}|) + (\alpha - 0, 5)E(X - x_{\alpha} + \mu_X - \mu_X) \\ &= 0, 5 E(|X - x_{\alpha}|) + (\alpha - 0, 5)[\underbrace{E(X - \mu_X)}_{=0} + E(\mu_X - x_{\alpha})] \\ &= 0, 5 E(|X - x_{\alpha}|) + (\alpha - 0, 5)(\mu_X - x_{\alpha}). \end{aligned} \quad (4.13)$$



Die Varianz  $\sigma_{\hat{x}_\alpha}^2$  des Quantilschätzers (und dessen asymptotische Normalität) ist schon länger bekannt<sup>19</sup>:

$$\sigma_{\hat{x}_\alpha}^2 = \frac{\alpha(1-\alpha)}{(f(x_\alpha))^2} \quad (4.14)$$

Zunächst wird mittels des Grenzwerts (4.12) und des Gradienten  $\nabla g(a, b)$  an der Stelle  $(a = \mu_X, b = x_\alpha)$  die Varianzformel hergeleitet:

$$\begin{aligned} n \sigma_{\widehat{ERVaR}_X}^2 &= \begin{pmatrix} -\frac{1}{x_\alpha} & \frac{\mu_X}{x_\alpha^2} \end{pmatrix} \begin{pmatrix} \sigma_X^2 & \sigma_{X, \hat{x}_\alpha} \\ \sigma_{X, \hat{x}_\alpha} & \sigma_{\hat{x}_\alpha}^2 \end{pmatrix} \begin{pmatrix} -\frac{1}{x_\alpha} \\ \frac{\mu_X}{x_\alpha^2} \end{pmatrix} \\ &= \frac{1}{x_\alpha^2} \sigma_X^2 - 2 \frac{\mu_X}{x_\alpha^3} \sigma_{X, \hat{x}_\alpha} + \frac{\mu_X^2}{x_\alpha^4} \sigma_{\hat{x}_\alpha}^2 \end{aligned}$$

Hierin (4.13) und (4.14) eingesetzt ergibt:

$$\begin{aligned} n \sigma_{\widehat{ERVaR}_X}^2 &= \frac{\sigma_X^2}{x_\alpha^2} - \frac{2 \mu_X [0,5 E(|X - x_\alpha|) + (\alpha - 0,5)(\mu_X - x_\alpha)]}{x_\alpha^3 f(x_\alpha)} \\ &\quad + \frac{\mu_X^2 \alpha(1-\alpha)}{x_\alpha^4 (f(x_\alpha))^2} \\ &= ERVaR_X^2 \left( \frac{\sigma_X^2}{\mu_X^2} - \frac{E(|X - x_\alpha|) + (2\alpha - 1)(\mu_X - x_\alpha)}{\mu_X x_\alpha f(x_\alpha)} \right. \\ &\quad \left. + \frac{\alpha(1-\alpha)}{(x_\alpha f(x_\alpha))^2} \right) \end{aligned}$$

Soll diese Varianz aus einer Stichprobe geschätzt werden, kann man wiederum die einzelnen Terme durch konsistente Schätzer derselben ersetzen. Für  $\mu_X$ ,  $\sigma_X^2$  und  $x_\alpha$  wurde dies bereits in Abschnitt 3.3 beschrieben und für  $E(|X - x_\alpha|)$  bietet sich der Schätzer  $n^{-1} \sum_{i=1}^n |X - \hat{x}_\alpha|$  an. Es bleibt die Frage nach einer geeigneten Schätzmethode für  $f(x_\alpha)$ .

Lässt sich die Verteilung von  $X$  gut und zuverlässig durch einen bestimmten Verteilungstyp beschreiben, bietet sich eine parametrische Schätzung der Dichte an. Im Fall von Finanzmarktdaten ist die  $t$ -Verteilung ein guter Kandidat, wenn die Verteilung annähernd symmetrisch ist. Folgt  $X$  einer Verteilung aus der Lage-Skalen-Familie einer Standard- $t_\nu$ -Verteilung, dann gilt:

$$\frac{X - \mu_X}{\sigma_X \sqrt{\frac{\nu-2}{\nu}}} \sim t_\nu$$

<sup>19</sup>vgl. Stuart und Ord (1987, S. 330f.)

Wir bestimmen zunächst das Quantil  $x_\alpha$  unter der Annahme der  $t_\nu$ -Verteilung:

$$\begin{aligned} F_{t_\nu} \left( \frac{x_\alpha - \mu_X}{\sigma_X \sqrt{\frac{\nu-2}{\nu}}} \right) &= \alpha \\ \frac{x_\alpha - \mu_X}{\sigma_X \sqrt{\frac{\nu-2}{\nu}}} &= F_{t_\nu}^{-1}(\alpha) \\ x_\alpha &= F_{t_\nu}^{-1}(\alpha) \sigma_X \sqrt{\frac{\nu-2}{\nu}} + \mu_X \end{aligned} \quad (4.15)$$

Für Dichten von Lage-Skalen-Familien  $y = a + bx$  gilt die Beziehung

$$f_Y(y) = \frac{1}{b} f_X \left( \frac{y-a}{b} \right).$$

Eingesetzt  $a = \mu_X$ ,  $b = \sigma_X \sqrt{(\nu-2)/\nu}$ ,  $y = x_\alpha$  und der Dichtefunktion  $f_{t_\nu}$  erhält man unter Einsetzen von (4.15) für  $x_\alpha$ :

$$\begin{aligned} f(x_\alpha) &= \frac{1}{\sigma_X \sqrt{\frac{\nu-2}{\nu}}} f_{t_\nu} \left( \frac{x_\alpha - \mu_X}{\sigma_X \sqrt{\frac{\nu-2}{\nu}}} \right) \\ &= \frac{1}{\sigma_X \sqrt{\frac{\nu-2}{\nu}}} f_{t_\nu} \left( \frac{F_{t_\nu}^{-1}(\alpha) \sigma_X \sqrt{\frac{\nu-2}{\nu}} + \mu_X - \mu_X}{\sigma_X \sqrt{\frac{\nu-2}{\nu}}} \right) \\ &= \frac{1}{\sigma_X \sqrt{\frac{\nu-2}{\nu}}} f_{t_\nu} (F_{t_\nu}^{-1}(\alpha)) \end{aligned}$$

Die Standardabweichung  $\sigma_X$  wird wie üblich durch  $S_X^*$  geschätzt. Mit der Beziehung

$$\gamma_2(X \sim t(\nu)) = \frac{6}{\nu-4} + 3 \Leftrightarrow \nu = \frac{6}{\hat{\gamma}_{2,X} - 3} + 4 \Rightarrow \hat{\nu} = \frac{6}{\hat{\gamma}_{2,X} - 3} + 4$$

kann man aus der empirischen Kurtosis die Zahl der Freiheitsgrade einer als  $t$ -verteilt angenommenen Zufallsvariablen schätzen.<sup>20</sup> Also lautet der Schätzer für die Dichte an der Stelle des Quantils  $x_\alpha$ :

$$\hat{f}(x_\alpha) = \frac{1}{\sqrt{(\hat{\nu}-2)/\hat{\nu}} S_X} f_{t_{\hat{\nu}}}(F_{t_{\hat{\nu}}}^{-1}(\alpha))$$

Falls die  $t$ -Verteilung ungeeignet zur Modellierung von  $f(x_\alpha)$  erscheint oder man bewusst auf jede parametrische Annahme verzichten möchte, kann man

<sup>20</sup>Da diese Formel nur für  $\hat{\gamma}_{2,X} > 3$  gültig ist, kann für Werte  $\hat{\gamma}_{2,X} \leq 3$  die Normalverteilung eingesetzt werden, gegen die die  $t$ -Verteilung für  $\nu \rightarrow \infty$  konvergiert. Eine andere Möglichkeit, die hier zur Vermeidung von Fallunterscheidungen bevorzugt wird, ist die, eine Untergrenze  $3 + \delta$ ,  $\delta > 0$  zu setzen, unterhalb derer  $\hat{\nu}$  auf  $\frac{6}{\delta} + 4$  festgelegt wird. Bei  $\delta = 0,01$  führt dies zu einer Obergrenze von 604 Freiheitsgraden, einer fast exakten Näherung der Normalverteilung.

auch nichtparametrische Verfahren zur Dichteschätzung einsetzen. Das (neben dem Histogramm) wohl am weitesten verbreitete Verfahren in diesem Bereich ist die Kerndichteschätzung.<sup>21</sup> Hierbei werden jedem Datenpunkt und seiner Umgebung gemäß einer Bandbreite und einer Kernfunktion eine Wahrscheinlichkeitsmasse (meist  $1/n$ ) zugewiesen. Erstere bestimmt die Form, in der Wahrscheinlichkeitsmasse auf die Umgebung eines Datenpunkts verteilt wird, letztere die Breite der lokalen Umgebung. Schließlich werden die einzelnen Wahrscheinlichkeitsmassen aufaddiert und ergeben bei vielen Verfahren eine Funktion, die die Eigenschaften einer Dichtefunktion aufweist. Die Form dieser Funktion wird maßgeblich durch die Bandbreite bestimmt. Kleine Bandbreiten bedeuten starke Ballung der Wahrscheinlichkeitsmasse in der engeren Umgebung von Datenpunkten. Im Kontext der Anwendung auf die Varianz des ERVaR muss beachtet werden, dass im Falle, dass  $n \cdot \alpha$  ganzzahlig ist, immer ein Datenpunkt auf  $x_\alpha = x_{n\alpha}$  fällt und die Kernfunktion an diesem Punkt in der Regel ihr Maximum erreicht. Ist  $n \cdot \alpha$  nicht ganzzahlig, fällt bei der angewendeten Interpolationsmethode für  $x_\alpha$  kein Datenpunkt auf das geschätzte Quantil. Somit muss man damit rechnen, dass  $E(\hat{f}(x_\alpha))$  von der Ganzzahligkeit von  $n \cdot \alpha$  abhängt. Die Stärke dieses Effekts ist stark von der extern (ggf. über einen Anpassungsalgorithmus) vorzugebenden Bandbreite abhängig.

Aus diesen Gründen wird hier ein anderes Verfahren eingesetzt, das eine lokale Linearisierung der empirischen Verteilungsfunktion vornimmt. Bei stetigen Zufallsvariablen gilt  $f(x) = dF(x)/dx$ . Die theoretische Verteilungsfunktion  $F(x)$  lässt sich mit steigendem  $n$  nach dem Satz von Glivenko-Cantelli beliebig genau durch die empirische Verteilungsfunktion  $\hat{F}(x)$  schätzen. Allerdings lässt sich die Dichte nicht direkt mittels der Ableitung bestimmen, da  $\hat{F}(x)$  eine Treppenfunktion mit bis zu  $n$  Sprungstellen ist, so dass  $d\hat{F}(x)/dx$  an allen Stellen gleich null ist – außer den Sprungstellen, an denen  $\hat{F}(x)$  nicht differenzierbar ist.

Diesem Problem kann man dadurch begegnen, dass man größere Abschnitte der empirischen Verteilungsfunktion betrachtet. Dabei kann man entweder eine feste Bandbreite definieren und zählen, wie viele Sprungstellen in diesen Bereich fallen<sup>22</sup>, oder man gibt die Anzahl der Sprungstellen vor und betrachtet, wie weit die beiden äußeren Sprungstellen voneinander entfernt sind. Hier wird letztere Möglichkeit angewandt.

Konkret werden zur Schätzung von  $f(x_0)$  der  $k$ -größte Stichprobenwert kleiner als  $x_0$ ,  $x_{-k}$ , und der  $k$ -kleinste Stichprobenwert größer als  $x_0$ ,  $x_{+k}$ , in Betracht gezogen. Gibt es keine doppelten Daten und taucht  $x_0$  in der Stichprobe auf, weist die empirische Verteilungsfunktion auf dem Abschnitt  $x_{-k}$  bis  $x_{+k}$  genau  $2k$  Sprungstellen der Höhe  $1/n$  auf. Dementsprechend ist die durchschnittliche Steigung auf diesem Abschnitt

$$\frac{\hat{F}(x_{+k}) - \hat{F}(x_{-k})}{x_{+k} - x_{-k}} = \frac{2k/n}{x_{+k} - x_{-k}}$$

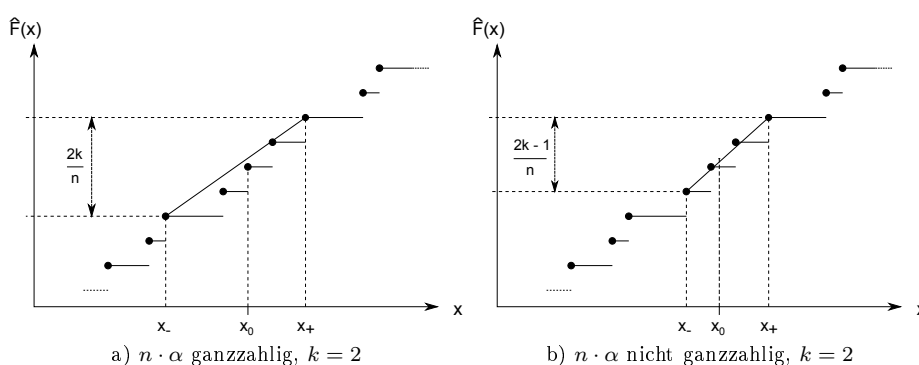
<sup>21</sup>Siehe hierzu z.B. Wand und Jones (1995) oder Silverman (1998).

<sup>22</sup>Dies entspricht einer Kerndichteschätzung mit der Gleichverteilung als Kernfunktion.

ein Schätzer für  $f(x_0)$ .<sup>23</sup> Dieses Prinzip soll nun auf die Schätzung von  $f(x_\alpha)$  angewendet werden. Schätzen wir  $x_\alpha$  mit Interpolation, so gilt bei aufsteigender Sortierung der Werte  $x_1, \dots, x_n$  die Eingrenzung  $x_{\lfloor n\alpha \rfloor} \leq \hat{x}_\alpha \leq x_{\lceil n\alpha \rceil}$ . Ist das Produkt  $n \cdot \alpha$  ganzzahlig, gilt  $x_{\lfloor n\alpha \rfloor} = x_{\lceil n\alpha \rceil}$  und der Schätzer für  $f(x_\alpha)$  ist:

$$\hat{f}(x_\alpha) = \frac{2k}{n(x_{n\alpha+k} - x_{n\alpha-k})}$$

Zwischen den Werten  $x_{n\alpha-k}$  und  $x_{n\alpha+k}$  hat die empirische Verteilungsfunktion  $2k$  Sprungstellen der Höhe  $1/n$ , wobei eine Sprungstelle, auf die mehrere Werte fallen, als mehrere Sprungstellen gezählt wird. Der Schätzwert für die Dichte ist die Steigung der in der Abbildung 4.1a eingezeichneten Gerade zwischen den entsprechenden Punkten auf der Verteilungsfunktion.



Schätzung der Dichte an der Stelle  $x_0$  durch die Steigung einer Gerade zwischen Sprungstellen der empirischen Verteilungsfunktion

Abbildung 4.1: Nichtparametrischer Dichteschätzer

Für nicht ganzzahlige Werte von  $n \cdot \alpha$  gilt dasselbe Prinzip, nur fällt auf  $\hat{x}_\alpha$  kein Wert und die Funktion überschreitet nur  $2k - 1$  Sprungstellen (Abbildung 4.1b):

$$\hat{f}(x_\alpha) = \frac{2k - 1}{n(x_{\lfloor n\alpha \rfloor + k} - x_{\lceil n\alpha \rceil - k})}$$

Wie bei anderen Methoden der Dichteschätzung hat auch hier die Wahl des Glättungsparameters  $k$  großen Einfluss auf die statistischen Eigenschaften des Dichteschätzers. Ist  $k$  klein, im Extremfall  $k = 1$ , ist die Varianz des Schätzers  $\hat{f}(x_0)$  hoch - man stelle sich vor, dass die beiden nächstliegenden Werte in der Stichprobe zufällig sehr nahe bei  $x_0$  liegen. Dann ist der Nenner des Bruchs

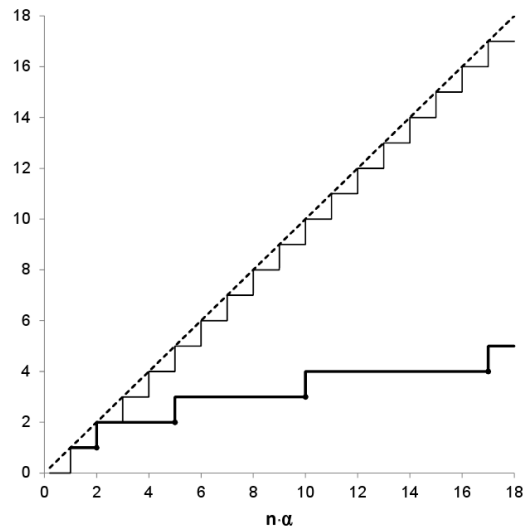
<sup>23</sup> Dieses Verfahren hat Ähnlichkeit mit dem Nächste-Nachbarn-Verfahren (siehe z.B. Silverman (1998, S. 96 ff.)), bei dem zur Schätzung von  $f(x_0)$  ebenfalls die Abstände zu den  $x_0$  nächstliegenden Werten herangezogen werden. Der Hauptunterschied des hier vorgeschlagenen Verfahrens ist, dass hier jeweils gleich viele Nachbarn größer und kleiner  $x_0$  gesucht werden. Beim klassischen Nächste-Nachbarn-Verfahren spielt die Richtung der Nachbarschaft keine Rolle, es können sogar alle Nachbarn von  $x_0$  aus gesehen in derselben Richtung liegen.

entsprechend klein und der Wert des Schätzers sehr hoch. Ist  $k$  groß, liegen die Werte  $x_{-k}$  und  $x_{+k}$  weit auseinander. Wenn die wahre Dichte zwischen  $x_{-k}$  und  $x_{+k}$  nicht konstant ist, ist der Schätzer verzerrt.

Bei der Anwendung der Schätzmethode auf die Dichte an der Stelle des Value-at-Risk  $\hat{f}(\hat{x}_\alpha)$  kommt hinzu, dass linksseitig von  $\hat{x}_\alpha$  nur  $(\lceil n\alpha \rceil - 1)$  Datenpunkte liegen, und dies sind bei üblichen Größen für  $\alpha$  und  $n$  nur sehr wenige. Dies begrenzt den Parameter  $k$ ; für  $n\alpha \leq 1$  ist der Schätzer gar nicht anwendbar. Ist die Untergrenze  $x_{-k}$  des Schätzbereichs das Minimum ( $x_1$ ) der Stichprobe, ist der Schätzer entsprechend ausreißeranfällig. Aus diesem Grund ist es sinnvoll, die Wahl des Parameters  $k$  nicht von  $n$ , sondern von  $n\alpha$  abhängig zu machen. Hier wird  $k$  nach folgender Formel festgelegt:

$$k = \lceil \sqrt{\lceil n\alpha - 1 \rceil} \rceil$$

In Abbildung 4.2 wird die Abhängigkeit von  $k$  vom Produkt  $n \cdot \alpha$  verdeutlicht. Grundsätzlich sollte man für kleine Werte von  $n\alpha$  überlegen, ob man nicht an Stelle dieses nichtparametrischen Verfahrens die oben genannte parametrische Methode oder ein spezielles Modell zur Anpassung der Ränder von Verteilungen vorziehen sollte.



Gestrichelte Linie:  $n \cdot \alpha$   
 Dünne Linie: Anzahl Werte kleiner als das  $\alpha$ -Quantil in der Stichprobe  
 Breite Linie: Parameterwert  $k$  bei gegebenem  $n \cdot \alpha$

Abbildung 4.2: Abhängigkeit des Parameters  $k$  von  $n \cdot \alpha$

### 4.1.3 Konfidenzintervalle auf Basis der Varianzformeln

Konfidenzintervalle  $[\hat{\theta}_-(\alpha), \hat{\theta}_+(\alpha)]$  sind Paare von Schätzern, die einen unbekannt Parameterwert  $\theta$  mit einer Wahrscheinlichkeit  $1 - \alpha$  einschließen. Eine zusätzliche wünschenswerte Eigenschaft liegt darin, dass der Wert  $\theta$  gleich oft über- wie unterschätzt werden sollte, d.h.  $P(\hat{\theta}_+(\alpha) < \theta) = P(\hat{\theta}_-(\alpha) > \theta) = \alpha/2$  (Überdeckungssymmetrie<sup>24</sup>).

Da die Punktschätzer der Performancemaße  $\hat{\theta}$  asymptotisch normalverteilt sind, kann man auf gewohnte Weise Konfidenzintervalle auf Basis der Quantile der (Standard-)Normalverteilung berechnen. Demnach sind die Grenzen eines asymptotischen Konfidenzintervalls zum Niveau  $1 - \alpha$  gegeben durch

$$\hat{\theta} \pm u_{1-\frac{\alpha}{2}} \cdot \hat{\sigma}_{\hat{\theta}} .$$

Dies ist eine asymptotische Aussage – für  $(n \rightarrow \infty)$  sind die oben genannten wünschenswerten Eigenschaften erfüllt. Für konkrete Anwendungen stellt sich die Frage, ob und inwieweit diese Asymptotik im Einzelfall eine geeignete Näherung darstellt (siehe hierzu Kapitel 5). Dabei wird deutlich werden, dass die Antwort stark von der jeweiligen Ausgangslage (untersuchtes Performancemaß, Stichprobengröße, Parameter) abhängt.

## 4.2 Die Bootstrap-Methode

Die Bootstrapping-Verfahren (kurz: der Bootstrap) stellen eine grundsätzlich andere inferenzstatistische Herangehensweise dar. Das Ziel liegt hier nicht in der Gewinnung theoretischer Größen, die ggf. anschließend empirisch schätzbar sind. Statt dessen sollen Eigenschaften von Schätzern in einer konkreten Anwendungssituation simulativ auf Basis derselben Stichprobenwerte gewonnen werden, aus denen die Schätzwerte berechnet wurden. Der Bootstrap gehört zur Klasse der Resampling-Verfahren (= Verfahren, die auf erneutem Ziehen aus Stichprobendaten beruhen) und wurde in Efron (1979) erstmals zusammenhängend dargestellt. Vorteile des Verfahrens sind ein sehr breites Anwendungsspektrum, einfache Durchführung und die Möglichkeit zum Verzicht auf Verteilungsannahmen. Heute werden eine größere Zahl verschiedener Ansätze zu den Bootstrapverfahren gezählt. In dieser Arbeit wird nur eines dieser Verfahren, der nichtparametrische Bootstrap, verwendet. Auf dessen Basis können mit mehreren unterschiedlichen Methoden Konfidenzintervalle für interessierende Größen berechnet werden. Fünf davon werden im zweiten Teil dieses Abschnitts erläutert. Anschließend werden einige Anwendungen des Bootstrap auf Performancemaße vorgestellt. Die Ausführungen sind recht kurz gehalten, auf Herleitungen und Beweise wird weitgehend verzichtet. Ausführliche Darstellungen der Bootstrapverfahren bieten Efron und Tibshirani (1993) und Davison und Hinkley (1997).

---

<sup>24</sup>Dieser Ausdruck wurde gewählt, um Verwechslungen mit der Eigenschaft auszuschließen, dass die Intervallgrenzen gleichen Abstand zum Punktschätzer haben (Symmetrie um  $t_n$ ).

### 4.2.1 Der nichtparametrische Bootstrap

Gesetzt,  $\theta$  sei eine theoretische Größe, die durch die Verteilung eines Vektors von Zufallsvariablen  $\mathbf{X}$  festgelegt ist. Es existiere nun ein Schätzer

$$\hat{\theta} = g(\hat{F}(\mathbf{X})),$$

dessen unbekannte Verteilungseigenschaften uns interessieren.

Wir verfügen über eine Stichprobe  $\mathbf{x}_N = (\mathbf{x}_1, \dots, \mathbf{x}_n)$  von  $n$  unabhängigen Realisationen aus  $\mathbf{X}$ , aus der wir die empirische Verteilungsfunktion  $F_n$  und den Schätzwert  $t_n = \hat{\theta}_{\mathbf{x}_N} = g(F_n)$  berechnen können. Die Verteilungsfunktion von  $\mathbf{X}$ ,  $F$ , und der wahre Wert  $\theta$  seien unbekannt. Es wird – von Regularitätsbedingungen abgesehen – keine Annahme über die Form von  $F$  getroffen.<sup>25</sup> Wie können nun Aussagen über die Verzerrung, die Varianz oder – allgemeiner betrachtet – über die Verteilung von  $\hat{\theta}$  gewonnen werden?

Würden wir  $F$  kennen, könnten wir aus  $F$  per Simulation beliebig viele Datenreihen der Länge  $n$  generieren, jeweils  $t_n$  berechnen und so die Verteilung von  $\hat{\theta}$  für eine bestimmte Stichprobenlänge  $n$  näherungsweise bestimmen. Die Grundidee des nichtparametrischen Bootstrap ist, genau dies zu tun, indem die empirische Verteilungsfunktion  $F_n$  als Schätzung für die unbekannte Verteilung  $F$  verwendet wird. Dies entspricht dem Ziehen mit Zurücklegen aus der vorliegenden Stichprobe  $\mathbf{x}_N$ . Eine Pseudo-Stichprobe der Länge  $n$  durch Ziehen mit Zurücklegen aus  $n$  Ausgangsdaten wird als Bootstrap-Stichprobe  $\mathbf{x}^*$  bezeichnet. Aus dieser Bootstrap-Stichprobe wird auf dieselbe Weise wie aus der Ausgangsstichprobe die Größe  $\theta$  geschätzt, der Schätzwert wird als Bootstrap-Replikation  $t_n^*$  bezeichnet. Dieses Verfahren wird nun  $B$ -mal wiederholt. Die aus  $(\mathbf{x}_1^*, \dots, \mathbf{x}_B^*)$  berechneten Werte für  $t_n^*$ , nämlich  $(t_{n1}^*, \dots, t_{nB}^*)$ , ergeben die Bootstrapverteilung des Schätzers  $\hat{\theta}$ , dargestellt durch die empirische Verteilungsfunktion  $F_{\hat{\theta}_n}^*$ . Aus dieser Bootstrapverteilung kann man Schätzer für einige interessierende Größen berechnen.

- Der Bootstrap-Schätzer für den *Standardfehler* von  $\hat{\theta}$  ist die empirische Standardabweichung der Bootstrap-Replikationen

$$SE_{\hat{\theta}}^* = \sqrt{\frac{1}{B-1} \sum_{b=1}^B (t_{nb}^* - \bar{t}_n^*)^2}.$$

- Die *Verzerrung* des Schätzers  $\hat{\theta}$  kann als Abweichung des Mittelwerts der Bootstrap-Verteilung  $\bar{t}_n^*$  vom Schätzwert aus der Stichprobe  $t_n$  geschätzt werden.

$$\text{bias}_{\hat{\theta}}^* = \bar{t}_n^* - t_n$$

<sup>25</sup>Beim hier nicht dargestellten parametrischen Bootstrap gehen wir hingegen davon aus, dass wir die Verteilung von  $\mathbf{X}$  mit Ausnahme einer oder mehrerer Verteilungsparameter kennen.

Die Frage nach der Wahl von  $B$  kann zwecks Minimierung des Simulationsfehlers prinzipiell mit „so groß wie möglich“ beantwortet werden, wenn als praktische Begrenzung von  $B$  der hinnehmbare Aufwand an Rechenzeit gesehen wird. Ein sinnvoller Mindestwert für  $B$  ist situationsabhängig; Efron und Tibshirani (1993, S. 52) sehen  $B = 200$  für die Schätzung von Standardfehlern als ausreichend an. Für Konfidenzintervalle seien erheblich höhere Replikationszahlen notwendig, Efron und Tibshirani rechnen für 90%- und 95%-Konfidenzintervalle in der Regel mit  $B = 2000$ .

Bei der Wahl von  $B$  sollte man auch beachten, dass es beim nichtparametrischen Bootstrap, einer Simulation auf Basis von Stichprobendaten, im Vergleich zu einer exakten Lösung auf Basis einer bekannten theoretischen Verteilung (mindestens) zwei Fehlerquellen gibt: Erstens ist der Ausgangspunkt nicht die theoretische Verteilung von  $\mathbf{X}$ , sondern die empirische Verteilung einer einzelnen Stichprobe. Ob die konkrete empirische Verteilung eine gute Näherung der wahren Verteilung ist, ist im Einzelfall unbekannt. Tendenziell wird dieser „Stichprobenfehler“ mit steigender Stichprobengröße  $n$  kleiner. Zweitens erfolgt die Berechnung der Ergebnisse nicht über eine exakte mathematische Herleitung, sondern über eine Simulation. Der „Simulationsfehler“ sinkt mit steigender Anzahl von Bootstrapreplikationen  $B$ . Wenn die Anzahl vorhandener Daten  $n$  vorgegeben ist, kann man durch Erhöhung von  $B$  nur den Simulationsfehler reduzieren, aber nicht den Stichprobenfehler. Mängel aus einer sehr schmalen Datenbasis können also nicht durch eine besonders hohe Anzahl von Bootstrapreplikationen ausgeglichen werden.

Zudem ist es nicht immer klar, ob sich ein Bootstrap-Ansatz für einen konkreten Anwendungsfall eignet. Der Bootstrap lässt sich zwar auf sehr vielfältige Problemstellungen anwenden, aber es sind Situationen bekannt, für die der Bootstrap ungeeignet ist. Wenn beispielsweise  $\hat{\theta}$  das Stichprobenminimum ist und man dessen Eigenschaften wie oben beschrieben mit dem nichtparametrischen Bootstrap untersuchen will, werden die Bootstrapreplikationen stets Werte annehmen, die in der Ausgangsstichprobe auftauchen; kleinster und häufigster Wert wird das Stichprobenminimum der Ausgangsstichprobe sein.<sup>26</sup> Einen Einstieg in die Problematik findet man in Davison und Hinkley (1997, S. 37 ff.), eine ausführliche Darstellung in Mammen (1992).

Da wir in dieser Arbeit hauptsächlich mit simulierten Daten arbeiten, die wahren Werte  $\theta$  kennen und mit der Delta-Methode einen Vergleichsansatz zur Verfügung haben, sollten wir ggf. in der Lage sein zu erkennen, wann Bootstrap-Ansätze schlechte Ergebnisse liefern.

**Beispiel DAX/BMW:** Als Beispiel für die Berechnungen in diesem Abschnitt betrachten wir monatliche Renditen des DAX und der BMW-Aktie aus fünf Jahren<sup>27</sup>, für die wir das Treynor-Ratio schätzen. Wir haben sechzig Datenpaare  $(M_{DAX}, X_{BMW})$ , kurz  $(M, X)$  zur Verfügung. Unter Vernachlässigung

<sup>26</sup>Für  $n \rightarrow \infty$  mit der Wahrscheinlichkeit  $\lim_{n \rightarrow \infty} [1 - (1 - 1/n)^n] = 1 - e^{-1} \approx 0,6321$ .

<sup>27</sup>Stetige Monatsrenditen der Jahre 2007 - 2011 von DAX 30 (ISIN DE0008469008) und BMW (ISIN DE0005190003) auf Basis von bereinigten Schlusskursen zum Monatsersten aus



eines sicheren Zinssatzes ( $R_f = 0$ ) erhalten wir aus der Stichprobe folgende Werte:

$$\begin{aligned}\bar{M} &= -0,00160 & \bar{X} &= 0,00334 \\ S_M^* &= 0,06705 & S_X^* &= 0,09627 \\ S_{M,X}^* &= 0,00486 & r_{M,X} &= 0,75254 \\ \widehat{Tr}_X &= \frac{\bar{X}S_M^*}{S_X^*r_{M,X}} = 0,003091 = t_n\end{aligned}$$

Bei der Anwendung des einfachen nichtparametrischen Bootstrapverfahrens nehmen wir ganz analog zu Abschnitt 4.1 an, die Renditepaare seien unabhängig und identisch verteilt. Es werden nun aus den  $n = 60$  Datenpaaren  $((M, X)_1, \dots, (M, X)_{60})$  mit Zurücklegen  $B = 2000$  Pseudo-Stichproben zu je 60 Werten gezogen und daraus die Werte  $t_{601}^*, \dots, t_{602000}^*$  berechnet. Der Durchschnitt dieser Werte war bei diesem Simulationslauf  $t_{60}^* = 0,003446$ . Demnach beträgt der geschätzte Bias

$$\text{bias}_{\widehat{Tr}_X}^* = \bar{t}_{60}^* - \widehat{Tr}_X = 0,003446 - 0,003091 = 0,000355.$$

Der Standardfehler als empirische Standardabweichung der  $t_{60}^*$  beträgt

$$SE_{\widehat{Tr}_X}^* = 0,01235.$$

## 4.2.2 Bootstrap-Konfidenzintervalle

Für Bootstrap-Konfidenzintervalle sind die Werte  $t_n^*$  Ausgangspunkt der Berechnungen. Zwei Berechnungsansätze sind intuitiv verständlich, erstens die Verwendung des geschätzten Standardfehlers zur Konstruktion eines Konfidenzintervalls um den Schätzwert herum und zweitens die Ermittlung von Quantilen der Bootstrapverteilung. Beide Methoden lassen sich noch verfeinern.

### 4.2.2.1 Bootstrap-Konfidenzintervalle über Standardfehler

Kann man davon ausgehen, dass ein Schätzer  $\hat{\theta}$  approximativ normalverteilt ist, kann man das Konfidenzintervall

$$\hat{\theta} \pm \widehat{SE}_{\hat{\theta}} \cdot F_{t_{n-1}}^{-1}(1 - \alpha/2)$$

verwenden, ein Intervall um den Schätzwert herum, dessen Breite vom geschätzten Standardfehler des Schätzers und dem  $(1 - \alpha/2)$ -Quantil der  $t$ -Verteilung mit  $n - 1$  Freiheitsgraden determiniert wird. Es liegt nun nahe, den Standardfehler des Schätzers wie oben gezeigt per Bootstrap zu schätzen. Das Konfidenzintervall lautet dann:<sup>28</sup>

$$t_n \pm SE_{\hat{\theta}}^* \cdot F_{t_{n-1}}^{-1}(1 - \alpha/2),$$

und wenn man zusätzlich noch die Verzerrung berücksichtigt:

$$t_n - \text{bias}_{\hat{\theta}}^* \pm SE_{\hat{\theta}}^* \cdot F_{t_{n-1}}^{-1}(1 - \alpha/2).$$

dem XETRA-Handel, Daten von finance.yahoo.com.

<sup>28</sup>vgl. Efron und Tibshirani (1993, Formel (12.19) auf S. 159)

Dieses Konfidenzintervall wird im Folgenden als *Bootstrap-t-Konfidenzintervall* bezeichnet.

Oft wird die Verteilung von  $\hat{\theta}$  bei gegebener Stichprobengröße nur unzureichend durch eine Normalverteilung approximiert oder man möchte auf diese parametrische Annahme von vornherein verzichten. Dazu lässt sich obiger Ansatz erweitern, denn man kann auch die Prüfverteilung aus den Daten heraus per Bootstrap ermitteln. Wir ermitteln dazu für jede Bootstrap-Stichprobe die studentisierte Größe

$$Z_b^* = \frac{t_{nb}^* - t_n}{\widehat{SE}(\hat{\theta}|F_{\mathbf{x}_b^*})} = \frac{t_{nb}^* - t_n}{SE^*(t_{nb}^*)}$$

Im Nenner des ersten Bruchs steht der Standardfehler des Schätzers  $\hat{\theta}$ , wenn  $\hat{\theta}$  aus der empirischen Verteilungsfunktion *der b-ten Bootstrap-Stichprobe* berechnet wird. Wenn keine Formel für  $\widehat{SE}(\hat{\theta}|F_{\mathbf{x}_b^*})$  zur Verfügung steht, kann man diesen Standardfehler wiederum per Bootstrap schätzen. Dazu wird die zugehörige Bootstrap-Stichprobe  $\mathbf{x}_b^*$  herangezogen, aus der  $C$  neue Bootstrap-Stichproben (zweiter Ordnung)  $\mathbf{x}_{bc}^{**}$  gezogen werden. Aus jeder Bootstrap-Stichprobe zweiter Ordnung wird ein Parameterschätzwert  $t_{bc}^{**}$  ermittelt. Der Standardfehler  $SE(\hat{\theta}|F_{\mathbf{x}_b^*})$  lässt sich dann als empirische Standardabweichung

$$SE^*(t_{nb}^*) = \sqrt{\frac{\sum_{c=1}^C (t_{bc}^{**} - \bar{t}_{nb}^{**})^2}{C - 1}}$$

schätzen. Nun lassen sich die Werte  $Z_b^*$  berechnen. Die empirische Verteilung der  $B$  Werte von  $Z_b^*$  dient als Prüfverteilung für  $t_n$ , die nur für die gegebene Stichprobe  $\mathbf{x}_N$  gültig ist und für ein Konfidenzintervall desselben Schätzers bei einer anderen Stichprobe vollständig neu berechnet werden muss. Das *studentisierte Bootstrap-Konfidenzintervall*

$$\left[ t_n + SE_{\hat{\theta}}^* \cdot F_{Z^*}^{-1}(\alpha/2), t_n + SE_{\hat{\theta}}^* \cdot F_{Z^*}^{-1}(1 - \alpha/2) \right]$$

liegt nun nicht mehr symmetrisch um  $t_n$ , da im Normalfall  $F_{Z^*}^{-1}(\alpha/2) \neq -F_{Z^*}^{-1}(1 - \alpha/2)$  ist.

Diese Methode hat den Vorteil, sich automatisch an die Form der Bootstrap-Verteilung anzupassen und konvergiert schneller gegen das exakte Konfidenzintervall als das einfache t-Konfidenzintervall.<sup>29</sup> Efron und Tibshirani (1993, S. 161 f.) kritisieren dagegen, das studentisierte Bootstrap-Intervall sei zwar in vielen Fällen, z.B. bei Lagemaßen, gut anwendbar, neige aber in anderen Fällen zu erraticem Verhalten. Zudem erhöht das „doppelte“ Bootstrappen, das für die Ermittlung der Prüfverteilung nötig ist, den Rechenaufwand stark. Auch wenn man sich nach dem von Efron und Tibshirani (1993, S. 161 f.) vorgeschlagenen Mindestwert von  $C = 25$  (bei  $B \geq 1000$ ) richtet, vervielfacht sich der Rechenaufwand.

<sup>29</sup>siehe Davison und Hinkley (1997, S. 211 ff.)

Der doppelte Bootstrap ist, da wir die Varianzformeln aus Abschnitt 4.1 kennen, nicht die einzige zur Verfügung stehende Methode zur Schätzung der Standardfehler  $SE(\hat{\theta}|F_{\mathbf{x}_b^*})$ . Man kann auch auf Basis jeder Bootstrap-Replikation  $\mathbf{x}_b^*$  die Varianz  $\sigma_{\hat{\theta}}^2$  des jeweiligen Performancemaßschätzers schätzen und (dessen Quadratwurzel) in den Studentisierungsterm einsetzen:

$$Z_b^* = \frac{t_{nb}^* - t_n}{\widehat{SE}(\hat{\theta}|F_{\mathbf{x}_b^*})} = \frac{t_{nb}^* - t_n}{S_{\hat{\theta}}(\mathbf{x}_b^*)}$$

Diese Methode spart gegenüber dem doppelten Bootstrap in erheblichem Maße Rechenzeit und wird in dieser Arbeit für das studentisierte Bootstrap-Konfidenzintervall durchgehend verwendet. Dabei wird stets die Varianzformel ohne Verteilungsannahme für die Renditen verwendet.

#### 4.2.2.2 Bootstrap-Konfidenzintervalle über Perzentile

Eine zweite einfache Möglichkeit zur Bestimmung von Konfidenzintervallen besteht darin, die Bootstrap-Verteilung  $F_{\hat{\theta}_n}^*$  unmittelbar als Näherung der Verteilung des Schätzers  $F_{\hat{\theta}_n}$  zu verwenden. Entsprechend besteht das *Perzentil-Konfidenzintervall*

$$\left[ t_n^*(\lceil B \cdot \frac{\alpha}{2} \rceil), t_n^*(\lceil B \cdot (1 - \frac{\alpha}{2}) \rceil) \right]$$

aus den empirischen  $(\alpha/2)$ - und  $(1 - \alpha/2)$ -Quantilen der Bootstrapverteilung  $F_{\hat{\theta}_n}^*$ .

Dieses Konfidenzintervall eignet sich gut als erster Referenzpunkt, da es keinerlei Verteilungsmodell impliziert, leicht zu berechnen ist und gegebenenfalls vorhandene Grenzen des Definitionsbereichs von  $\theta$  respektiert.<sup>30</sup> Die Präzision dieser Intervalle ist dagegen häufig unbefriedigend.<sup>31</sup> Die Konvergenz in Richtung der exakten, überdeckungssymmetrischen Konfidenzintervalle ist für  $n \rightarrow \infty$  relativ langsam und die Häufigkeit der Überdeckung des wahren Parameters kann in praktischen Anwendungen deutlich vom theoretischen Niveau  $(1 - \alpha)$  abweichen.

Eine Fortentwicklung des Perzentil-Konfidenzintervalls stellt das  $BC_a$ -Konfidenzintervall ( $BC_a = \text{Bias-corrected, accelerated}$ ) dar.<sup>32</sup> Es besteht ebenfalls aus zwei Quantilen der Bootstrap-Verteilung  $F_{\hat{\theta}_n}^*$ :

$$\left[ t_n^*(\lceil B\alpha_1 \rceil), t_n^*(\lceil B\alpha_2 \rceil) \right]$$

Anstelle der  $\alpha$ - und  $(1 - \alpha)$ -Quantile sind die (empirischen)  $\alpha_1$ - und  $\alpha_2$ -Quantile der Bootstrap-Verteilung die Grenzen des Konfidenzintervalls. Die Werte  $\alpha_1$  und

<sup>30</sup>Vgl. Efron und Tibshirani (1993, S. 171 ff.). Soll zum Beispiel ein Konfidenzintervall für das Bestimmtheitsmaß  $R^2$  einer Regression ermittelt werden, liegen die Grenzen eines Perzentil-Konfidenzintervalls stets im Bereich zwischen 0 und 1. Bei den Konfidenzintervallen aus Abschnitt 4.2.2.1 können die Grenzen auch darüber hinausgehen.

<sup>31</sup>vgl. Davison und Hinkley (1997, S. 203, S. 211ff.), Efron und Tibshirani (1993, S. 174f.)

<sup>32</sup>vgl. Efron und Tibshirani (1993, S. 184 ff.); ausführliche Darstellung in Efron (1987)

$\alpha_2$  sind

$$\alpha_1 = \Phi \left( \widehat{z}_0 + \frac{\widehat{z}_0 + u_{\frac{\alpha}{2}}}{1 - \widehat{a}(\widehat{z}_0 + u_{\frac{\alpha}{2}})} \right) \quad \text{und}$$

$$\alpha_2 = \Phi \left( \widehat{z}_0 + \frac{\widehat{z}_0 + u_{1-\frac{\alpha}{2}}}{1 - \widehat{a}(\widehat{z}_0 + u_{1-\frac{\alpha}{2}})} \right),$$

wobei  $\Phi$  die Verteilungsfunktion und  $u_z$  das  $z$ -Quantil der Standardnormalverteilung sind. Sind der Verzerrungskorrekturterm  $\widehat{z}_0$  und der Akzelerationsterm  $\widehat{a}$  beide null, erhält man das gewöhnliche Perzentil-Konfidenzintervall. Zur Berechnung der beiden Terme:

$\widehat{z}_0$  ist das Standardnormal-Quantil des Anteils der gebootstrapteten Werte, die kleiner oder gleich  $t_n$  sind:

$$\widehat{z}_0 = \Phi^{-1} \left( F_{\widehat{\theta}_n}^*(t_n) \right)$$

Der Term wird null, wenn genau die Hälfte der Bootstrapreplikationen kleiner oder gleich dem Schätzwert aus der Stichprobe sind.

$\widehat{a}$  kann mit Hilfe der Jackknife-Methode<sup>33</sup> ermittelt werden:

$$\widehat{a} = \frac{\sum_{i=1}^n \left( \widehat{\theta}_{(\cdot)}^{jack} - \widehat{\theta}_{(i)}^{jack} \right)^3}{6 \left[ \sum_{i=1}^n \left( \widehat{\theta}_{(\cdot)}^{jack} - \widehat{\theta}_{(i)}^{jack} \right)^2 \right]^{1,5}}$$

$BC_\alpha$ -Konfidenzintervalle teilen die positiven Eigenschaften der Perzentil-Intervalle, konvergieren aber vergleichbar schnell wie die studentisierten Bootstrap-Intervalle.<sup>34</sup> Solange die Stichprobenlänge  $n$  nicht zu groß ist, wird nur wenig zusätzliche Rechenzeit benötigt.

**Beispiel DAX/BMW (Fortsetzung):** Für die oben aus Renditedaten des DAX und der BMW-Aktie generierte Bootstrap-Verteilung  $F_{Tr_X, n=60, B=2000}^*$  erhalten wir folgende 95%-Konfidenzintervalle:

- *Bootstrap-t-Konfidenzintervall ohne Korrektur des Bias:*

$$\begin{aligned} & \widehat{Tr}_X \pm SE_{\widehat{Tr}_X}^* \cdot F_{t_{59}}^{-1}(0, 975) \\ \Rightarrow & 0,00309 \pm 0,01235 \cdot 2,00099 \\ \Rightarrow & [-0,02162, 0,02780] \end{aligned}$$

<sup>33</sup>Der Jackknife ist ebenfalls ein Resampling-Verfahren, bei dem in der hier verwendeten einfachsten Version die interessierende Statistik unter Wegfall von jeweils genau einem Wert wiederholt neu berechnet wird:  $\widehat{\theta}_{(i)}^{jack}(\mathbf{x}) = \widehat{\theta}(\mathbf{x} \setminus x_i)$ . Der Jackknife-Schätzer  $\widehat{\theta}_{(\cdot)}^{jack}$  ist der Mittelwert aller möglichen Jackknife-Replikationen:

$$\widehat{\theta}_{(\cdot)}^{jack} = \frac{1}{n} \sum_{i=1}^n \widehat{\theta}_{(i)}^{jack}$$

Eine kurze Einführung in die Jackknife-Methode bietet Efron und Tibshirani (1993, Kap. 11, S. 141 - 152)

<sup>34</sup>vgl. Efron und Tibshirani (1993, S.187)

- *Bootstrap-t-Konfidenzintervall mit Bias-Korrektur*: Das Intervall wird um den Bias von 0,000355 nach unten verschoben:

$$[-0,02198, 0,02744]$$

- *Studentisiertes Bootstrap-Konfidenzintervall*

$$[-0,01887, 0,02805]$$

- *Perzentil-Konfidenzintervall*

$$\left[ t_{60}^* (\lceil 2000 \cdot 0,025 \rceil = 50), t_{60}^* (\lceil 2000 \cdot 0,975 \rceil = 1950) \right] = [-0,01967, 0,02881]$$

- *BC<sub>a</sub>-Konfidenzintervall*

$$\left[ t_{60}^* (37), t_{60}^* (1935) \right] = [-0,0201, 0,0281]$$

### 4.2.3 Bootstrapping von Performancemaßen

Bootstrap-Verfahren lassen sich direkt auf die verschiedensten Performancemaße anwenden. Die am Beispiel des Treynor-Ratios gezeigte Vorgehensweise kann man analog bei allen anderen hier gezeigten Performancemaßen einsetzen.<sup>35</sup>

Bootstrap-Verfahren wurden bereits von einer Reihe von Autoren zur Untersuchung der statistischen Eigenschaften von Performancemaßen angewandt. Vinod und Morey (2000) wenden den Bootstrap zur Schätzung von Konfidenzintervallen auf Sharpe- und Treynor-Ratio an, Scherer (2004) thematisiert den Bootstrap bei Sharpe- und Sortino-Ratio und schlägt eine einfache AR(1)-Korrektur für autokorrelierte Daten vor. Kosowski et al. (2006), Kosowski et al. (2007) und Fama und French (2010) setzen den Bootstrap zur Analyse von Multifaktor-Alphas bei Hedge- bzw. Publikumsfonds ein. Keiner dieser Autoren untersucht die Qualität der benutzten Varianz- und Intervallschätzer. Dies leisten Ledoit und Wolf (2008), die bei Gültigkeit von  $H_0 : Sh_X = Sh_Y$  die Ablehnungshäufigkeiten von fünf verschiedenen Tests vergleichen: drei Tests unter Annahme eines normalverteilten Schätzers – mit Jobson-Korkie-Varianzformel (nach Memmel-Korrektur) bzw. zwei verschiedenen HAC-Varianzschätzern – und zwei Bootstrap-Tests: einem auf Basis eines symmetrischen studentisierten Bootstrap-Konfidenzintervalls und einem auf Basis eines Circular-Block-Bootstrap-Konfidenzintervalls. Demnach hält nur das Block-Bootstrap-Intervall den vorgegebenen  $\alpha$ -Fehler bei allen datenerzeugenden Prozessen (i.i.d., GARCH und VAR) ein. Alle anderen Tests neigen dazu, zu liberal

<sup>35</sup>Zweifel sind hier allenfalls beim ERVaR angebracht, da dort Quantile geschätzt werden. Bei der Schätzung von Quantilen aus Bootstrap-Stichproben einer Ausgangsstichprobe kann stets nur eine begrenzte Zahl an Schätzwerten auftauchen. Ist das Produkt  $n \cdot \alpha$  ganzzahlig, beschränkt sich die Menge möglicher Bootstrap-Schätzwerte auf die Menge der in der Ausgangsstichprobe realisierten Werte. Inwiefern Bootstrap-Konfidenzintervalle für den ERVaR brauchbar sind, wird im Rahmen der Simulationen in Abschnitt 5.4 erörtert.

zu sein, wobei es aber nur bei autokorrelierten Daten bzw. bei Verwendung des NV-Tests mit Jobson-Korkie-Varianzen (bei nicht i.i.d.-normalverteilten Renditen) zu erheblichen Überschreitungen des vorgegebenen Niveaus  $\alpha$  (1 %, 5 % bzw. 10 %) kommt.

### 4.3 Weitere Ansätze

Morey und Morey (2000) leiten folgende Grenzen für ein  $(1-\alpha)$ -Konfidenzintervall für das Treynor-Ratio her:<sup>36</sup>

$$[KI_{1-\alpha}^-, KI_{1-\alpha}^+] = \frac{\bar{X}\hat{\beta} \pm t\sqrt{\bar{X}^2 S_{\hat{\beta}}^2 + \frac{S_{\bar{X}}^2}{n} (\hat{\beta}^2 - t^2 S_{\hat{\beta}}^2)}}{\hat{\beta}^2 - t^2 S_{\hat{\beta}}^2},$$

wobei  $t = t_{n-1}(1-\frac{\alpha}{2})$  das  $1-\frac{\alpha}{2}$ -Quantil der  $t$ -Verteilung mit  $n-1$  Freiheitsgraden ist. Dieses Intervall ist nur dann anwendbar, wenn der Schätzwert von  $\hat{\beta}$  zum Niveau  $\alpha$  statistisch signifikant von null verschieden ist, da sonst der Nenner des Bruchs negativ wird.

Pedersen und Satchell (2000) entwickeln Konfidenzintervalle für Jensens Alpha in einem asymmetrischen Verteilungsmodell.

---

<sup>36</sup>Formel (20) in Morey und Morey (2000) unter Korrektur eines Druckfehlers; vgl. ebd., Formel (22)

# Kapitel 5

## Simulationen

Die Varianzformeln aus dem letzten Kapitel sind Ergebnisse asymptotischer Herleitungen. Nun sollen die Eigenschaften der einzelnen Varianzschätzer mit Hilfe von Monte-Carlo-Simulationen untersucht werden. Außerdem werden die Methoden zur Herleitung von Konfidenzintervallen aus dem letzten Kapitel geprüft.

Monte-Carlo-Simulationen haben im Vergleich zur Arbeit mit realen Daten den Vorteil, dass man die Ausgangsparameter des Simulationsprozesses vollständig kennt. Dies gilt erstens für die Verteilung und die temporäre Abhängigkeitsstruktur innerhalb der Zeitreihen, so dass kontrolliert werden kann, ob die Annahmen für die Herleitung der Varianzen erfüllt sind. Zweitens kann man in der Regel die theoretischen Werte der Performancemaße über die Verteilungsparameter vorgeben.

Im ersten Abschnitt dieses Kapitels werden zunächst Methodik und Notation erläutert. In den weiteren Abschnitten werden die Simulationsergebnisse dargestellt, wobei zuerst gezeigt wird, inwieweit die Schätzer bei relativ großen Stichprobenlängen konvergieren. Anschließend werden die Eigenschaften der Schätzer der einzelnen Performancemaße bei realistischen Stichprobenlängen genauer untersucht. Den Abschluss des Kapitels bildet ein Vergleich der verschiedenen Ansätze zur Berechnung von Konfidenzintervallen.

### 5.1 Methodik

Zunächst werden ein- bis dreidimensionale Zeitreihen erzeugt. Anschließend werden daraus Performancemaße und deren Varianzen berechnet sowie weitere Analysen durchgeführt.

#### 5.1.1 Datenerzeugende Prozesse

Folgende Parameter werden bei der Erzeugung von Zufallszeitreihen vorgegeben (siehe auch Tabelle 5.1):

- *Dimension der Zeitreihe:* Für die einzelnen Maße werden folgende Zeitreihen benötigt:
  - Univariate Zeitreihe ( $d = 1$ ):  $Sh, \kappa_{\alpha,z}, ERVaR$
  - Bivariate Zeitreihe ( $d = 2$ ):  $Tr, \text{Jensens } \alpha, MM, \Delta_{Sh}, \Delta_{\kappa_{\alpha,z}}$
  - Trivariate Zeitreihe ( $d = 3$ ):  $\Delta_{Tr}, \Delta_{\alpha}, \Delta_{MM}$
- *Länge der Zeitreihe:* In der Mehrzahl der Fälle werden  $n = 60$  Werte angesetzt, entsprechend den Monatsrenditen aus fünf Jahren.
- *Zeitreihenstruktur:* In der Regel wird mit *i.i.d.*-Zeitreihen gearbeitet. Multivariate GARCH-Renditen werden mit dem Modell (2.1) (siehe S. 20) simuliert. AR-Renditen werden nur für univariate Zeitreihen simuliert, um die Varianzformel für autokorrelierte Daten von Schmid und Schmidt (2010) zu testen.
- *Verteilungstyp:* Es werden Normalverteilung und *t*-Verteilung verwendet, letztere mit zehn oder fünf Freiheitsgraden. Für die Konvergenz der meisten Varianzschätzer sind mindestens fünf Freiheitsgrade erforderlich, da dann das vierte Moment der Verteilung endlich ist.
- *Erwartete Überrenditen und Standardabweichungen:* Dies sind jeweils die  $d$ -dimensionalen Vektoren  $\mu$  und  $\sigma$ . Die weiteren Parameter in datenerzeugenden Prozessen (z.B. der Parameter  $\alpha_0$  in GARCH-Prozessen) werden derart gewählt, dass die erzeugten Zeitreihen im Mittel tatsächlich diese Vorgaben erfüllen.
- *Korrelationsmatrix:* Bei mehrdimensionalen Zeitreihen muss eine Korrelationsmatrix vorgegeben werden. Bei bivariaten Zeitreihen bedeutet dies einen variierbaren Parameter ( $\rho_{M,X}$  bzw.  $\rho_{X,Y}$ ), bei trivariaten Zeitreihen drei ( $\rho_{M,X}, \rho_{M,Y}, \rho_{X,Y}$ ).
- *Anzahl der Läufe pro Simulation  $R$*
- weitere, technische Parameter wie z.B. die Anzahl vorlaufender Perioden zum „Einschwingen“ von Nicht-i.i.d.-Zeitreihen.

Je Parameterkombination und Dimensionalität<sup>1</sup> wird jeweils nur eine Matrix von  $d \times n \times R$  Portfoliorenditen simuliert. Alle zugehörigen Maße und Varianzformeln werden also aus denselben Daten berechnet, so dass Vergleiche zwischen Performancemaßen und Varianzformeln möglich sind. Beispielsweise werden aus einmal simulierten  $R$  eindimensionalen Zeitreihen für eine bestimmte Parameterkombination alle Auswertungen für Sharpe-Ratio, Kappa und ERVaR (mit ggf. unterschiedlichen Parametern der Maße) und mit den ggf. unterschiedlichen Varianzformeln vorgenommen.

<sup>1</sup>Hiermit sind die Ausprägungen *ein Portfolio* ("X"), *Markt und ein Portfolio* ("MX"), *zwei Portfolios* ("XY") sowie *Markt und zwei Portfolios* ("MXY") gemeint.



Parameter	Kürzel/Benennung/Ausprägungen
Anzahl Dimensionen der Zeitreihe	$d = 1$ (ein Portfolio) $d = 2$ (Markt und ein Portfolio) oder (zwei Portfolios) $d = 3$ (Markt und zwei Portfolios)
Länge der Zeitreihen	$n$
Zeitreihenstruktur	i.i.d., AR(1) oder GARCH(1,1)
ggf. AR-/GARCH-Parameter	$\alpha_1$ bzw. $\alpha_1$ und $\beta_1$ <sup>a</sup>
Verteilungstyp	Normal- oder $t$ -Verteilung mit Vorgabe der Anzahl Freiheitsgrade
Vektor der erwarteten Überrenditen	$d$ Parameter $\in \mathbb{R}$
Vektor der Standardabweichungen	$d$ Parameter $\in \mathbb{R}^+$
Korrelationsmatrix	$\frac{d(d-1)}{2}$ Parameter $\in [-1; 1]$
Anzahl Simulationsläufe	$R$

<sup>a</sup>Der verbleibende Parameter  $\alpha_0$  ergibt sich aus den anderen Parametern und der gewünschten Varianz.

Tabelle 5.1: Parameter bei der Simulation von Zeitreihen

Die absolute Größe der Renditeparameter ist in der Simulation ohne eigentliche Bedeutung. Ob man eine Zeitreihenschar mit der Parameterkombination  $(\mu_X = 0,01; \sigma_X = 0,1)$  oder  $(\mu_X = 1; \sigma_X = 10)$  simuliert und analysiert, führt zu keinen qualitativen Änderungen der Ergebnisse. Anstatt einer möglichen Normierung eines Parameters (z.B.  $\sigma_X = 1$ ) wurden die Parameter so gewählt, dass eine direkte ökonomische Interpretation der Simulationsergebnisse prinzipiell möglich bleibt. Im gesamten Kapitel wird davon ausgegangen, dass Monatsrenditen vorliegen und auch die Performancemaße auf Monatsbasis berechnet werden, so dass Probleme bei der Umrechnung von Performancemaßen auf andere Zeiteinheiten (z.B. Annualisierung) nicht auftreten.

### 5.1.2 Auswertung

Aus den Parametern und den simulierten Zeitreihen wurden für jedes Performancemaß, wo möglich, folgende Größen errechnet:

- Aus den vorgegebenen Parametern kann entsprechend dem jeweiligen Verteilungsmodell ein **theoretischer Wert des Performancemaßes**  $\theta$  berechnet werden. Dazu stehen verschiedene Methoden zur Verfügung:
  - direkte analytische Berechnung aus den vorgegebenen Simulationsparametern,
  - numerische Berechnung durch Auswertung eines Integrals über eine bestimmte Verteilungsfunktion,
  - Näherung durch Simulation: Hierzu werden die einzelnen simulierten Zeitreihen zu einer einzigen Zeitreihe (der Länge  $n \cdot R$ ) zusammen-

gefasst. Da stets  $n \cdot R \geq 600\,000$  ist, kann von einer ausreichend genauen Schätzung des theoretischen Werts ausgegangen werden, wenn die Schätzer konsistent sind. Durch Simulationen angenäherte theoretische Werte sind in den Tabellen dieses Kapitels entsprechend gekennzeichnet.

Die Ermittlung des theoretischen Wertes eines Performancemaßes ist in den meisten Fällen direkt analytisch möglich. Bei dem Performancemaß Kappa werden die theoretischen partiellen Momente numerisch ermittelt.

- Aus den Renditen jeder simulierten Zeitreihe wird jeweils ein **Schätzwert des Performancemaßes**  $t_i$  berechnet. Berichtet wird der Mittelwert der Schätzungen  $\bar{t}$ .
- Die **theoretische Standardabweichung**  $\sigma_0(\hat{\theta})$  des Performancemaßes gemäß den asymptotischen Berechnungen aus Kapitel 4 kann bei normalverteilten Daten zumeist direkt aus den Parametern berechnet werden. In den meisten anderen Fällen wird sie aufgrund der dann in den Formeln enthaltenen höheren gemischten Momente numerisch berechnet oder simuliert.
- Aus den Zeitreihen kann mittels der Varianzformeln aus Kapitel 4 jeweils die **Standardabweichung des Performancemaßes**  $s_i(\hat{\theta})$  geschätzt werden.
- Aus der Gesamtheit der  $R$  Schätzwerte des Performancemaßes kann die **empirische Standardabweichung der Performancemaß-Schätzwerte**  $s(t_i)$  berechnet werden.
- Schließlich wird aus den geschätzten Performancemaß-Varianzen der einzelnen Zeitreihen eine **durchschnittliche geschätzte Varianz**  $\overline{s^2}(\hat{\theta}_i)$  errechnet. Deren Wurzel stellt eine gemittelte Standardabweichung dar, welche im Folgenden verkürzt als  $\overline{s}(\hat{\theta}_i)$  notiert wird.

Auf Basis dieser Werte können weitere Auswertungen vorgenommen werden. Zunächst einmal ist von Interesse, ob der Schätzer eines bestimmten Performancemaßes für eine bestimmte Parameterkombination annähernd normalverteilt ist oder nicht. Dies kann mit dem bereits auf S. 29 vorgestellten Anderson-Darling-Test getestet werden. Er erhält den Vorzug gegenüber dem Jarque-Bera-Test, da letzterer besonders sensibel auf Extremwerte reagiert, was bei der großen Anzahl an Schätzwerten ( $R \geq 10\,000$ ) nicht erwünscht ist. Es wird darauf getestet, ob die empirische Verteilung der Performancemaß-Schätzwerte mit einer entsprechend angepassten Normalverteilung vereinbar ist; berichtet wird jeweils der  $p$ -Wert des Anderson-Darling-Tests.

Angesichts der großen Anzahl von Datenpunkten der empirischen Verteilung reicht bereits eine kleinere Abweichung der tatsächlichen Verteilung des Schätzers von der Normalverteilung aus, um den  $p$ -Wert des Anderson-Darling-Tests

sehr klein werden zu lassen, so dass der Test auf jedem üblichem Niveau zur Ablehnung der Hypothese normalverteilter Daten führt. Dadurch ist keine Aussage möglich, wie stark die Form der empirischen Verteilung von der einer Normalverteilung abweicht. Entsprechend wird eine Maßzahl benötigt, die angibt, wie stark sich zwei Verteilungen unterscheiden, hier die empirische Verteilung von einer angepassten Normalverteilung. Eine solche Maßzahl wird im folgenden Exkurs entwickelt.

---

**Exkurs: Eine Maßzahl für den Unterschied zwischen zwei Verteilungen** Die Maßzahl wird (für stetige Zufallsvariablen  $X_1$  und  $X_2$ ) über die Größe des Bereiches  $p$  hergeleitet, der von der Abszisse und beiden Dichtefunktionen gemeinsam eingeschlossen wird, illustriert in Abbildung 5.1. Als Maß für die Differenz der beiden Funktionen dient  $\psi_{f_1, f_2} = 1 - p$ , das zudem der Hälfte der Fläche entspricht, die nur unterhalb einer der beiden Dichtefunktionen liegt. Die letztere Angabe kann man in folgende Formel fassen:<sup>2</sup>

$$\psi_{f_1, f_2} = \frac{1}{2} \int_{-\infty}^{\infty} |f_1(x) - f_2(x)| dx$$

Es gilt  $0 \leq \psi_{f_1, f_2} \leq 1$  mit der Interpretation der Extremwerte

- $\psi_{f_1, f_2} = 0$ :  $X_1$  und  $X_2$  sind identisch verteilt, so dass  $f_1(x) = f_2(x)$
- $\psi_{f_1, f_2} = 1$ : Die Träger von  $X_1$  und  $X_2$  sind disjunkt.

Will man ermitteln, wie gut eine Verteilung  $f_1(x)$  einer Normalverteilung entspricht, kann man diese standardisieren und mit der Dichtefunktion einer Standardnormalverteilung  $\phi(x)$  anhand obiger Formel vergleichen. Für diesen Spezialfall gilt:

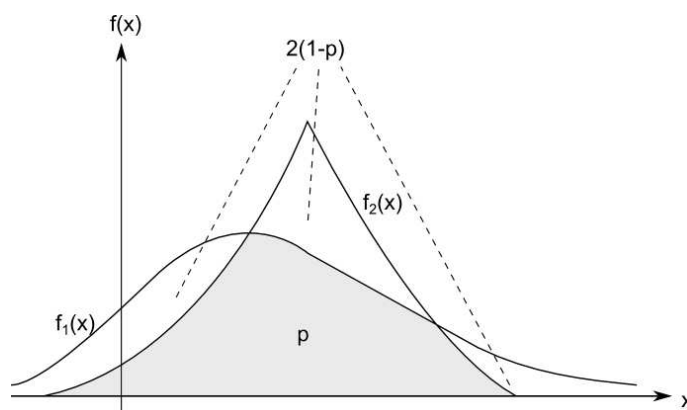
$$\psi_{f, \phi} = \frac{1}{2} \int_{-\infty}^{\infty} \left| f_{\frac{x-\mu_x}{\sigma_x}}(y) - \phi(y) \right| dy.$$

Da hier empirische Verteilungen mit theoretischen Verteilungen wie der Normalverteilung verglichen werden sollen, brauchen wir einen Schätzer für  $\psi$ . Hierfür werden Histogramme von empirischer und theoretischer Verteilung verglichen. Es werden  $K$  Klassen gebildet, für die die absolute Differenz zwischen der Anzahl der in eine Klasse  $k_j$  fallenden Beobachtungen  $n_{k_j}$  und der erwarteten Häufigkeit für eine Klasse unter Gültigkeit von  $f_0$ ,  $E(n_{k_j} | f_0) = n \cdot P(X \in k_j | f_0)$ , ermittelt wird. Die Einteilung von  $\mathbb{R}$  in  $K$  Klassen erfolgt entsprechend den  $\frac{k}{K}$ -Quantilen der theoretischen Verteilung, wodurch bei Gültigkeit von  $f_0$  in jeder Klasse aus

$$(k_1, k_2, \dots, k_K) = (]-\infty, F_0^{-1}(\frac{1}{K})], ]F_0^{-1}(\frac{1}{K}), F_0^{-1}(\frac{2}{K})], \dots, ]F_0^{-1}(\frac{K-1}{K}), +\infty[)$$

---

<sup>2</sup>Hierbei handelt es sich um die Hälfte der  $L_1$ -Distanz  $\|f - g\|_1 = \int_{-\infty}^{\infty} |f(x) - g(x)| dx$ , vgl. Karian und Dudewicz (2000, S. 197), für die in der Folge ein geeigneter Schätzer entwickelt wird. Eng verwandt ist der Überlappungskoeffizient  $OV_L(X, Y) = 1 - \psi_{f_1, f_2}$ , vgl. Schmid und Schmidt (2006), die diesen ebenfalls nichtparametrisch, aber mittels Kerndichteschätzung schätzen.



Berechnung des Maßes  $\psi_{f_1, f_2}$  für den Unterschied zwischen zwei Verteilungen, repräsentiert durch die Dichtefunktionen  $f_1(x)$  und  $f_2(x)$ :  
 $\psi_{f_1, f_2} = 1 - p$  ( $p$ : Getönte Fläche)

Abbildung 5.1: Grafische Erläuterung zu  $\psi_{f_1, f_2}$

dieselbe Anzahl  $n/K$  Beobachtungen zu erwarten ist. Die Abweichungen der Zahl der in jeweils eine Klasse fallenden Beobachtungen von  $n/K$  werden aufsummiert und durch  $2n$  geteilt<sup>3</sup>:

$$\begin{aligned}\hat{\psi}_{f, f_0} &= \frac{1}{2n} \sum_{k=1}^K \left| \sum_{i=1}^n I \left( F_0^{-1} \left( \frac{k-1}{K} \right) < x_i \leq F_0^{-1} \left( \frac{k}{K} \right) \right) - n/K \right| \\ &= \frac{1}{2n} \sum_{k=1}^K |n_k - n/K|\end{aligned}$$

Für den Vergleich einer empirischen Verteilung mit einer angepassten Normalverteilung lautet die Berechnungsvorschrift

$$\hat{\psi}_{f, \phi} = \frac{1}{2n} \sum_{k=1}^K \left| \sum_{i=1}^n I \left( u_{\frac{k-1}{K}} < \frac{x_i - \bar{x}}{s_x} \leq u_{\frac{k}{K}} \right) - n/K \right|$$

mit  $u_a$  als  $a$ -Quantil der Standardnormalverteilung.

Der Wertebereich von  $\hat{\psi}$  ist bei theoretischen Verteilungen mit Träger  $= \mathbb{R}$  das Intervall  $[0; 1 - \frac{1}{K}]$ , weil im Extremfall alle Beobachtungen in einer Klasse liegen und die Summe der Abstände dann  $1 \cdot (n - \frac{n}{K}) + (K - 1) \cdot \frac{n}{K} = 2n(1 - \frac{1}{K})$  beträgt, was  $\psi = 1 - \frac{1}{K}$  ergibt. Die hier vorgeschlagene Klassenaufteilung nach Quantilen der theoretischen Verteilung hat mehrere Vorteile: Erstens sind alle Klassen der theoretischen Verteilung gleich stark besetzt, so dass die Verteilung gleichmäßig geprüft wird. Zweitens braucht die Klassenaufteilung nicht extern vorgegeben zu werden, sondern ergibt sich aus  $f_0$  und  $K$ . Drittens verbleibt mit

<sup>3</sup> $I(A)$  ist die Indikatorfunktion, die den Wert 1 annimmt, wenn Aussage  $A$  zutrifft und ansonsten gleich 0 ist.

$K$  dann nur noch ein einziger Parameter für die Methode.

Bei gegebener Beobachtungsanzahl  $n$  tendiert  $\hat{\psi}$  dazu, mit steigendem  $K$  ebenfalls anzusteigen, da sich dann zufällige Abweichungen weniger gut innerhalb der Klassen ausgleichen können. Deshalb ist es sinnvoll, die Anzahl der Klassen  $K$  zu einer Funktion der Beobachtungsanzahl  $n$  zu machen, hier

$$K = \lfloor \sqrt{n} \rfloor.$$

Durch diese Maßnahme beträgt die Obergrenze der Maßzahl

$$\max \hat{\psi} = 1 - \frac{1}{K} = 1 - \frac{1}{\lfloor \sqrt{n} \rfloor} \leq 1 - \frac{1}{\sqrt{n}}.$$

□

Lässt sich aus den Simulationsparametern mittels der Varianzformeln eine asymptotische Varianz  $\sigma_0(\hat{\theta})$  ermitteln, kann man darauf testen, ob dieser Wert mit der empirischen Varianz der Performancemaß-Schätzungen vereinbar ist. Ist eine Zufallsvariable  $Z$  normalverteilt und ist  $\mathbf{Z}_R = \{Z_1, \dots, Z_R\}$  eine einfache Stichprobe aus  $Z$ , kann folgender **Chi-Quadrat-Test** durchgeführt werden<sup>4</sup>:

$$H_0 : \sigma^2 = \sigma_0^2 \quad \text{gegen} \quad H_1 : \sigma^2 \neq \sigma_0^2$$

Die Teststatistik lautet:

$$T = \frac{\frac{R-1}{R} \sum_{i=1}^R (Z_i - \bar{Z})^2}{\sigma_0^2}$$

$H_0$  wird abgelehnt, wenn  $T \notin [F_{\chi_{R-1}^2}^{-1}(\frac{\alpha}{2}), F_{\chi_{R-1}^2}^{-1}(1 - \frac{\alpha}{2})]$ . Berichtet wird der  $p$ -Wert des (zweiseitigen) Tests:<sup>5</sup>

$$p\text{-Wert} = 1 - 2 \cdot \left| 0,5 - F_{\chi_{R-1}^2}(T) \right| \quad (5.1)$$

Außerdem werden **theoretische  $u_{1-\alpha/2}$ - $\sigma$ -Bereiche**

$$\left[ \theta \pm u_{1-\frac{\alpha}{2}} \sigma_0(\hat{\theta}) \right]$$

berechnet und geprüft, welcher Anteil der Werte  $\hat{\theta}_i$  in dieses Intervall hineinfällt. Eine gute Näherung dieses Anteils an das vorgegebene Niveau  $1 - \alpha$  bedeutet, dass man bei bekannten Verteilungsparametern gut jenen Bereich abgrenzen kann, in den eine Realisierung mit einer bestimmten Wahrscheinlichkeit hineinfallen wird. Den  $u_{1-\alpha/2}$ - $\sigma$ -Bereich kann man außerdem als Annahmehereich

<sup>4</sup>vgl. z. B. Schlittgen (1996, S. 342)

<sup>5</sup>Diese Formel folgt direkt aus der Definition des  $p$ -Werts, siehe Erläuterung in Anhang A.3.2

Kürzel	Erläuterung
$\theta$	theoretischer Wert des Performancemaßes bei gegebener Parameterkombination (exakt oder genähert)
$\bar{t}$	Mittelwert der Schätzwerte $t_i$ von $\theta$ aus den simulierten Zeitreihen
$\sigma_0(\hat{\theta})$	theoretische asymptotische Standardabweichung des Performancemaßschätzers $\hat{\theta}$
$s(t_i)$	empirische Standardabweichung der Schätzwerte $t_i$
$\bar{s}(\hat{\theta}_i)$	$= \sqrt{\frac{1}{n} \sum_{i=1}^R s_i^2(\hat{\theta})}$ , gemittelter Wert der aus den einzelnen Zeitreihen geschätzten Standardabweichungen von $\hat{\theta}$
$p_{AD}$	$p$ -Wert des Anderson-Darling-Tests auf Normalverteilung der $t_i$
$\hat{\psi}$	Wert des oben definierten Maßes des Abstands der empirischen Verteilung von $t_i$ zu einer angepassten Normalverteilung
$p_{\sigma^2=\sigma_0^2}$	$p$ -Wert des $\chi^2$ -Tests von $H_0 : \sigma^2(\hat{\theta}) = \sigma_0^2(\hat{\theta})$ anhand von $s(t_i)$
$[u_\alpha]$	$u_{1-\frac{\alpha}{2}}$ - $\sigma$ -Bereich: Ober- und Untergrenze des Bereiches von $u_{1-\frac{\alpha}{2}}$ theoretischen Standardabweichungen um den theoretischen Wert des Performancemaßes
% $t_i$ in $u$	Anteil der Schätzwerte $t_i$ , die innerhalb des $u_{1-\frac{\alpha}{2}}$ - $\sigma$ -Bereichs liegen
% KI um $\theta$	Anteil der aus den Zeitreihen geschätzten Konfidenzintervalle, die den theoretischen Wert $\theta$ einschließen

Tabelle 5.2: Übersicht über die in den folgenden Abschnitten verwendeten Simulationsstatistiken

eines approximativen Gauß-Mittelwert-Tests der Nullhypothese  $H_0 : \theta = \theta_0$  interpretieren. In diesem Sinne gibt die Zahl einen Hinweis darauf, ob ein solcher Test sein Niveau einhalten würde.

In umgekehrter Weise wird für jede Zeitreihe  $i$  ein **Gauß-Konfidenzintervall**

$$\left[ \hat{\theta}_i \pm u_{1-\frac{\alpha}{2}} s_{\hat{\theta},i} \right]$$

geschätzt und geprüft, ob der Anteil der Intervalle, die den wahren Wert  $\theta$  einschließen, dem theoretischen Niveau  $1 - \alpha$  nahe kommt. Hiermit wird die Zuverlässigkeit der Varianzschätzung  $s_{\hat{\theta}}^2$  geprüft. Eine Aussage über die Qualität der Varianzschätzung ist nur dann möglich, wenn der Performancemaß-Schätzer in ausreichender Näherung normalverteilt ist.

Tabelle 5.2 zeigt eine Übersicht über alle im folgenden Abschnitt verwendeten Statistiken.

## 5.2 Konvergenz der Schätzer

Es soll zunächst gezeigt werden, wie sich die Schätzer bei großen Stichproben unabhängig und identisch verteilter Renditen verhalten. Sind sie in hinreichender Näherung normalverteilt? Entspricht die empirische Varianz der Schätzungen

der asymptotischen Varianz, wie sie aus den Verteilungsparametern gemäß den Formeln aus Abschnitt 4.1 berechnet wurde? Diese Ergebnisse liefern auch Hinweise darauf, ob sich Probleme, die bei kleinen Stichprobengrößen auftreten, auf die geringe Stichprobengröße zurückführen lassen, oder ob der Schätzer nicht gegen die erwarteten Werte konvergiert.

Dazu werden Zeitreihen der Länge  $n = 500$  simuliert. Die in diesem Abschnitt verwendete Parameterkombination steht in der folgenden Tabelle, wobei je nach Dimension der Zeitreihe nicht alle Werte benötigt werden.

Variable	$\mu$	$\sigma$	Korrelationsmatrix		
$X$	0,002	0,1	1	0,75	0,75
$Y$	0,003	0,1	0,75	1	0,75
$M$	0,002	0,1	0,75	0,75	1
$n = 500$ , i.i.d., Normal- bzw. $t_5$ -Verteilung					
$R = 20\,000$ ( $d = 1$ ), $R = 10\,000$ ( $d = 2, d = 3$ )					

Tabelle 5.3: Simulationsparameter in Abschnitt 5.2

Eine stetige Monatsüberrendite von 0,002 entspricht einer diskreten Jahresüberrendite von  $e^{12 \cdot 0,002} - 1 = 2,43\%$ , was - isoliert betrachtet - einen durchaus vorzeigbaren ökonomischen Anlageerfolg darstellen würde. Das untersuchte Portfolio  $X$  hat dieselbe Überrendite wie das Marktportfolio  $M$  und eine geringere Überrendite als das Konkurrenzportfolio  $Y$ . Wir beginnen mit den Ergebnissen für normalverteilte Daten und Performancemaße ohne Benchmark, für die also eine univariate Zeitreihe ausreicht, siehe Tabelle 5.4.<sup>6</sup>

Die erste Spalte der Tabelle enthält die Namen der Performancemaße, darunter jeweils eine bis vier Kurzbezeichnungen für Varianten von Varianzformeln. Für jedes Performancemaß stehen in der ersten Zeile Statistiken, die unabhängig von der verwendeten Varianzformel sind. Die Spaltenbezeichnungen für die Daten zu den einzelnen Varianzformeln sind in der Kopfzeile der Tabelle angegeben.<sup>7</sup>

Das erste betrachtete Maß ist das *Sharpe-Ratio*. Aus den Parametern ( $\mu_X = 0,002$ ,  $\sigma_X = 0,1$ ) ergibt sich der wahre Wert  $\theta = \frac{0,002}{0,1} = 0,02$ . Dieser Wert ist, wie alle anderen Werte in dieser Zeile, unabhängig davon, welche Varianzformel im weiteren Verlauf angewendet wird. Der Mittelwert der  $R = 20\,000$  Schätzwerte des Sharpe-Ratios beträgt  $\bar{t} = 0,02024$ , die empirische Standardabweichung  $s(t_i) = 0,04458$ . Die Hypothese, dass die Schätzwerte einer Normalverteilung

<sup>6</sup> *Hinweis:* Dieses Tabellenlayout berücksichtigt, dass für ein Performancemaß meist mehrere Formelvarianten zur Berechnung der Schätzvarianz zur Verfügung stehen. Entsprechend hängen einige Ergebnisse, wie die Daten zur empirischen Verteilung der Schätzwerte, nicht von der verwendeten Varianzformel ab. Andere, wie die Überdeckungshäufigkeiten von Konfidenzintervallen, sind für jede Formelvariante anders. Deshalb werden erstere Daten für jede Simulationsreihe in einer ersten Zeile angegeben. Die restlichen Daten werden für jede Formelvariante in den folgenden Zeilen aufgeführt.

<sup>7</sup> Die Bedeutung aller Kürzel kann man in Tabelle 5.2 auf S. 114 nachschlagen.

Varianzformel	$\sigma_0(\hat{\theta})$	$\bar{s}(\hat{\theta}_i)$	$p_{\sigma^2=\sigma_0^2}$	$[u_{0,05}]$	$t_i$ in $u$	$KI_{95\%}$ um $\theta$
<b>Sharpe<sub>X</sub></b>	$\theta = 0,02$ ; $\bar{t} = 0,02024$ ; $s(t_i) = 0,04458$ ; $p_{AD} = 0,1859$ ; $\hat{\psi} = 0,0324$					
i.i.d., NV	0,04473	0,04475	0,5091	[-0,0677;0,1077]	94,975%	95,065%
i.i.d.	0,04473	0,04475	0,5091	[-0,0677;0,1077]	94,975%	95,055%
allg., $j=1$	0,04473	0,04464	0,5091	[-0,0677;0,1077]	94,975%	94,855%
allg., $j=5$	0,04473	0,04428	0,5091	[-0,0677;0,1077]	94,975%	94,280%
<b>Kappa<sub>X</sub><sup><math>\alpha=1</math></sup><sub><math>z=0</math></sub></b>	$\theta = 0,05141$ ; $\bar{t} = 0,05863$ ; $s(t_i) = 0,11887$ ; $p_{AD} < 10^{-64}$ ; $\hat{\psi} = 0,0567$					
i.i.d.	0,1122	0,1196	$< 10^{-99}$	[-0,1685;0,2713]	93,560%	95,030%
<b>Kappa<sub>X</sub><sup><math>\alpha=2</math></sup><sub><math>z=0</math></sub></b>	$\theta = 0,02847$ ; $\bar{t} = 0,03144$ ; $s(t_i) = 0,06549$ ; $p_{AD} < 10^{-11}$ ; $\hat{\psi} = 0,0463$					
i.i.d.	0,06327	0,06583	$< 10^{-30}$	[-0,0955;0,1524]	94,115%	95,200%
<b>Kappa<sub>X</sub><sup><math>\alpha=3</math></sup><sub><math>z=0</math></sub></b>	$\theta = 0,02183$ ; $\bar{t} = 0,02353$ ; $s(t_i) = 0,04957$ ; $p_{AD} < 10^{-7}$ ; $\hat{\psi} = 0,0450$					
i.i.d.	0,04823	0,04983	$< 10^{-20}$	[-0,0727;0,1163]	94,255%	95,165%
<b>ERVaR<sub>X5%</sub></b>	$\theta = 0,01231$ ; $\bar{t} = 0,01332$ ; $s(t_i) = 0,02810$ ; $p_{AD} < 10^{-21}$ ; $\hat{\psi} = 0,0441$					
i.i.d., param.	0,02787	0,02828	0,0929	[-0,0423;0,0669]	94,685%	95,295%
i.i.d., nonp.	0,02787	0,02828	0,0929	[-0,0423;0,0669]	94,685%	95,250%
Erläuterung der Kürzel für die Statistiken in Tab. 5.2 auf S. 114.						

Tabelle 5.4: Simulationsergebnisse für die benchmarkfreien Maße bei i.i.d.-normalverteilten Daten und Zeitreihenlänge  $n = 500$

entstammen können, kann anhand des Anderson-Darling-Tests ( $p$ -Wert: 0,1859) nicht abgelehnt werden, was bei einem Stichprobenumfang von 20 000 auf eine sehr gute Übereinstimmung zwischen Normalverteilung und empirischer Verteilung hindeutet. Auch die Abstandsmaßzahl nimmt mit  $\hat{\psi} = 0,0324$  einen niedrigen Wert an.

Es werden nun drei verschiedene Berechnungsformeln zur Schätzung der Varianz von  $\widehat{Sh}_X$  untersucht: Die Formel für unabhängig und identisch normalverteilte Werte (4.3, im Anhang A.18), jene für unabhängig und identisch verteilte Werte ohne Normalverteilungsannahme (4.2, im Anhang A.17) und die Formel (4.5, im Anhang A.19), die auch autokorrelierte Werte zulässt. Letztere Formel besitzt einen Parameter  $j$ , die Anzahl der Perioden, deren Autokorrelation mit der aktuellen Periode mit einbezogen wird. Beispielhaft werden hier die Ergebnisse für  $j = 1$  bzw.  $j = 5$  einbezogene Perioden dargestellt.

Nach der Formel für i.i.d.-normalverteilte Werte beträgt die Standardabweichung des Schätzers  $\widehat{Sh}_X$  gemäß den Parameterwerten

$$\sigma_0(\hat{\theta}) = \sqrt{n^{-1} \left( 1 + \frac{1}{2} Sh_X^2 \right)} = \sqrt{\frac{1}{500} \left( 1 + \frac{1}{2} \cdot 0,02^2 \right)} = 0,04473.$$

Die in der Zeile darüber angegebene empirische Standardabweichung  $s(t_i) = 0,04458$  kommt dem sehr nahe und mit dem  $\chi^2$ -Varianztest erhält man einen  $p$ -Wert von 0,5091, so dass die Nullhypothese, dass der über die Asymptotik berechnete Wert  $\sigma_0(\hat{\theta}) = 0,04473$  der wahre Wert ist, nicht abgelehnt werden kann. Die Schätzungen der Varianzen mittels der i.i.d.-NV-Formel aus den einzelnen Zeitreihen ergeben eine mittlere geschätzte Standardabweichung von 0,04475; dieser Wert liegt sehr nahe bei den anderen beiden.

Im zweiten Teil der Zeile folgt zunächst der asymptotische 95%-Bereich: Für



$n \rightarrow \infty$  sollte der Anteil der Schätzwerte, die in das Intervall  $\theta \pm 1,96\sigma_0$  fallen, 95% betragen. Für die gegebenen Parameter reicht dieser Bereich von -0,0677 bis 0,1077, er schließt also trotz der deutlichen Überrendite auch bei 500 Monatsrenditen den Wert 0 mit ein. Von den 20 000 Schätzwerten fielen 18 995, entsprechend einem Anteil von 94,975%, in diesen Bereich. Der theoretische Anteil wird also fast genau erreicht. Schließlich wurden mittels der Schätzwerte für  $Sh_X$  und  $\sigma_{\widehat{Sh}_X}$  aus den 20 000 einzelnen Zeitreihen jeweils 95%-Konfidenzintervalle berechnet. Von diesen enthielten 19 013, also 95,065%, den wahren Wert 0,02. Auch hier ist die Näherung sehr gut – bei einem allerdings sehr „gutartigen“ Fall, bei dem zutreffend Normalverteilung angenommen wird und lange Zeitreihen vorliegen.

In der nächsten Zeile folgen die Ergebnisse für die verteilungsfreie i.i.d.-Formel. In diese Formel fließen Schiefe und Kurtosis von  $X$  ein, die aus den Daten geschätzt werden. Im Falle normalverteilter Daten sind ihre Werte bekannt ( $\gamma_{1,X} = 0, \gamma_{2,X} = 3$ ). In die Formel eingesetzt erhält man denselben theoretischen Wert  $\sigma_0(\hat{\theta})$  wie im NV-Fall und entsprechend sind auch die aus diesem Wert und den empirischen Schätzwerten  $t_i$  berechneten Statistiken ( $p_{\sigma^2=\sigma_0^2}$ ,  $[u_{0,05}]$ ,  $t_i$  in  $u$ ) identisch. Bei der verteilungsfreien Schätzung der Varianz von  $\widehat{Sh}_X$  aus den normalverteilten Zeitreihendaten werden jedoch  $\gamma_{1,X}$  und  $\gamma_{2,X}$  unnötigerweise mitgeschätzt und in die Varianzformel einbezogen. Dadurch wird die geschätzte Varianz „verrauscht“ und die Resultate der weiteren Statistiken könnten schlechter sein. Das ist hier aber nicht der Fall. Die mittlere Standardabweichung weicht erst ab der sechsten (nicht dargestellten) Nachkommastelle ab und die Überdeckungshäufigkeit der Konfidenzintervalle ist mit 95,055% (zufällig) sogar noch etwas besser.

Bei der generalisierten Varianzformel nach Schmid und Schmidt (siehe Formel 4.5, S. 76), deren Ergebnisse in den nächsten beiden Zeilen folgen, werden zusätzlich Autokorrelationseffekte bis zur einer bzw. fünf Perioden mit einbezogen. Wie im vorigen Fall resultiert aus der allgemeineren Formel, in die die Werte des Spezialfalls i.i.d.-NV eingesetzt werden, die i.i.d.-NV-Formel, so dass die theoretischen Werte gleich bleiben. Das zusätzliche „Rauschen“ durch die unnötige Einbeziehung von geschätzten höheren Momenten und Autokorrelationen verschlechtert die Ergebnisse im Fall  $j = 1$  kaum und erst bei  $j = 5$  merklich. Insgesamt ist die asymptotische Näherung im beschriebenen Fall sehr gut, auch für die verallgemeinerten Formeln.

Auf die *Kappa*-Maße trifft dies nicht im selben Maße zu. Bei Kappa-1 ist der Mittelwert der Schätzungen von 0,05863 deutlich vom theoretischen Wert 0,05141 verschieden und der AD-Test lehnt die Normalverteilungshypothese eindeutig ab, wenn auch der Wert von  $\hat{\psi}$  andeutet, dass die Abweichung von der Normalverteilung moderat ist. Die asymptotische theoretische Standardabweichung des Schätzers (Formel A.37) liegt mit 0,1122 unter der empirischen Standardabweichung und die beiden Werte sind nach dem  $\chi^2$ -Test nicht miteinander vereinbar. Entsprechend fallen mit 93,56 % auch weniger als 95 % der Schätzwerte in den  $\theta \pm u_{0,975}\sigma_0$ -Bereich. Wenn die Standardabweichung über die einzelnen Zeitreihen geschätzt wird, erhält man im Mittel mit  $\bar{s}(\hat{\theta}_i) = 0,1196$  in

Varianzformel	$\sigma_0(\hat{\theta})$	$\bar{s}(\hat{\theta}_i)$	$p_{\sigma^2=\sigma_0^2}$	$[u_{0,05}]$	$t_i$ in $u$	$KI_{95\%}$ um $\theta$
<b>Treynor<sub>X</sub></b>	$\theta = 0,00267$ ; $\bar{t} = 0,00259$ ; $s(t_i) = 0,00592$ ; $p_{AD} = 0,8114$ ; $\hat{\psi} = 0,0346$					
i.i.d., NV	0,00596	0,00597	0,2511	[-0,0090;0,0144]	95,27%	95,42%
i.i.d.	0,00596	0,00596	0,2511	[-0,0090;0,0144]	95,27%	95,23%
<b>Jensen<sub>X</sub></b>	$\theta = 0,0005$ ; $\bar{t} = 0,000455$ ; $s(t_i) = 0,00293$ ; $p_{AD} = 0,1950$ ; $\hat{\psi} = 0,0404$					
i.i.d., NV	0,00296	0,00296	0,2121	[-0,0053;0,0063]	95,17%	95,09%
i.i.d.	0,00296	0,00295	0,2121	[-0,0053;0,0063]	95,17%	95,09%
<b>MM<sub>X,ori</sub></b>	$\theta = 0,002$ ; $\bar{t} = 0,00201$ ; $s(t_i) = 0,00444$ ; $p_{AD} = 0,1563$ ; $\hat{\psi} = 0,0391$					
i.i.d., NV	0,00447	0,00447	0,3579	[-0,0068;0,0108]	95,16%	95,21%
i.i.d.	0,00447	0,00450	0,3579	[-0,0068;0,0108]	95,16%	95,57%
<b>MM<sub>X,alt</sub></b>	$\theta = 0$ ; $\bar{t} = -0,000001$ ; $s(t_i) = 0,00315$ ; $p_{AD} = 0,8376$ ; $\hat{\psi} = 0,0381$					
i.i.d., NV	0,00316	0,00316	0,6326	[-0,0062;0,0062]	95,06%	95,13%
i.i.d.	0,00316	0,00317	0,6326	[-0,0062;0,0062]	95,06%	95,30%
Erläuterung der Kürzel für die Statistiken in Tab. 5.2 auf S. 114.						

Tabelle 5.5: Simulationsergebnisse für die Maße mit Benchmark bei i.i.d.-normalverteilten Daten und Zeitreihenlänge  $n = 500$

etwa den Wert der empirischen Standardabweichung und die Überdeckungshäufigkeit ( $KI_{95\%}$  um  $\theta$ ) des theoretischen Werts 0,05141 durch die aus den Zeitreihen geschätzten Konfidenzintervalle beträgt nahezu exakt 95 %. Für Kappa-2 (Sortino-Ratio, Formel A.38) und Kappa-3 (Formel A.39) ergibt sich ein ähnlicher Befund, allerdings mit schwächer werdender Abweichung der Verteilung des Schätzers von der Normalverteilung. Für die Kappa-Maße gibt es also auch bei langen Zeitreihen Abweichungen von den erwünschten Idealbedingungen. Darauf wird im Unterabschnitt zu diesem Performancemaß ab S. 156 genauer eingegangen.

Für den Schätzer des *Excess Return on Value-at-Risk* sind die Ergebnisse gemischt: Der Schätzer ist deutlich verzerrt ( $\bar{t} = 0,01332$  bei einem theoretischen Wert von 0,01231) und nicht normalverteilt. Dagegen wird die Hypothese, dass die tatsächliche Schätzvarianz gleich dem durch die asymptotische Varianzformel berechneten Wert von 0,02787 (nach Formel A.40) ist, auf dem 5 % -Niveau nicht abgelehnt. Die Schätzung der Schätzvarianz aus den Daten führt zu unauffälligen Ergebnissen: Die durchschnittliche Standardabweichung liegt nahe an der tatsächlich beobachteten und die Überdeckungshäufigkeit der Konfidenzintervalle liegt nahe am vorgegebenen Niveau von 95 %. Es zeigen sich kaum Unterschiede zwischen parametrischer und nichtparametrischer Dichteschätzung innerhalb der Varianzschätzung.

Als nächstes folgen in Tabelle 5.5 die Ergebnisse für Performancemaße mit Benchmark bei der oben angegebenen Parameterkombination (vgl. Tabelle 5.3).

Bei allen vier Maßen scheint eine ausreichende Konvergenz erreicht zu sein: Die Abweichungen der Schätzwertdurchschnitte  $\bar{t}$  von den theoretischen Werten  $\theta$  sind jeweils gering und eine Normalverteilung der Schätzwerte kann für keines der Maße abgelehnt werden. Auch die Tests auf  $\sigma^2 = \sigma_0^2$  führen nie zur

Varianzformel	$\sigma_0(\hat{\theta})$	$\bar{s}(\hat{\theta}_i)$	$p_{\sigma^2=\sigma_0^2}$	$[u_{0,05}]$	$t_i$ in $u$	$KI_{95\%}$ um $\theta$
$\Delta Sh_{X,Y}$	$\theta = -0,01$ ; $\bar{t} = -0,01060$ ; $s(t_i) = 0,03170$ ; $p_{AD} = 0,935$ ; $\hat{\psi} = 0,036$					
i.i.d., NV	0,03239	0,03168	0,0022	[-0,0735;0,0535]	95,52%	95,17%
i.i.d.	0,03239	0,03163	0,0022	[-0,0735;0,0535]	95,52%	95,10%
$\Delta\kappa(X,Y)_{z=0}^{\alpha=1}$	$\theta = -0,0267$ ; $\bar{t} = -0,0285$ ; $s(t_i) = 0,08561$ ; $p_{AD} = 0,077$ ; $\hat{\psi} = 0,036$					
i.i.d.	-	0,08610	-	-	-	95,74%
$\Delta\kappa(X,Y)_{z=0}^{\alpha=2}$	$\theta = -0,0147$ ; $\bar{t} = -0,0157$ ; $s(t_i) = 0,04693$ ; $p_{AD} = 0,561$ ; $\hat{\psi} = 0,040$					
i.i.d.	-	0,04704	-	-	-	95,39%
$\Delta Tr_{X,Y}$	$\theta = -0,00133$ ; $\bar{t} = -0,00137$ ; $s(t_i) = 0,00429$ ; $p_{AD} = 0,345$ ; $\hat{\psi} = 0,036$					
i.i.d., NV	0,00422	0,00423	0,0146	[-0,0096;0,0069]	94,23%	94,37%
i.i.d.	0,00422	0,00423	0,0146	[-0,0096;0,0069]	94,23%	94,29%
$\Delta Je_{X,Y}$	$\theta = -0,001$ ; $\bar{t} = -0,00102$ ; $s(t_i) = 0,00321$ ; $p_{AD} = 0,230$ ; $\hat{\psi} = 0,0404$					
i.i.d., NV	0,00316	0,00316	0,2303	[-0,0072;0,0052]	94,38%	94,37%
i.i.d.	0,00316	0,00316	0,2303	[-0,0072;0,0052]	94,38%	94,37%
$\Delta MM_{X,Y}$	$\theta = -0,001$ ; $\bar{t} = -0,00102$ ; $s(t_i) = 0,00321$ ; $p_{AD} = 0,259$ ; $\hat{\psi} = 0,0399$					
i.i.d., NV	0,00316	0,00316	0,0419	[-0,0072, 0,0052]	94,25%	94,31%
i.i.d.	0,00316	0,00317	0,0419	[-0,0072, 0,0052]	94,25%	94,40%

Erläuterung der Kürzel für die Statistiken in Tab. 5.2 auf S. 114.

Tabelle 5.6: Simulationsergebnisse für Differenzen von Performancemaßen bei i.i.d.-normalverteilten Daten und Zeitreihenlänge  $n = 500$

Ablehnung der Nullhypothese. Die Überdeckungsanteile der Konfidenzintervalle liegen stets leicht über 95 %. Unterschiede zwischen der verteilungsfreien Varianzberechnung<sup>8</sup> und der (hier genau passenden) Formel mit NV-Annahme<sup>9</sup> sind kaum zu erkennen; allenfalls ist der Überdeckungsanteil der Konfidenzintervalle bei  $MM_{X,ori}$  geringfügig erhöht.

Tabelle 5.6 zeigt die Ergebnisse für die Performancemaß-Differenz zweier Portfolios. Bei allen sechs Maßen ist die Anpassung der Verteilung des Schätzers an die Normalverteilung sehr gut, der Anderson-Darling-Test lehnt zum 5%-Niveau nie ab. Der Test auf die Varianz lehnt jedoch in drei von vier Fällen ab. Hinsichtlich der Erfüllung des 95%-Niveaus bei Konfidenzintervallen und 1,96 $\sigma$ -Bereichen gibt es Abweichungen in der Größenordnung von ca. einem halben bis einem dreiviertel Prozentpunkt.<sup>10</sup>

In Tabelle 5.7 auf S. 121 folgen nun die Ergebnisse für Zeitreihen mit unabhängig und identisch verteilten Renditen aus einer  $t$ -Verteilung mit fünf Freiheitsgraden. Bei fünf Freiheitsgraden sind die (gemischten) Momente bis ein-

<sup>8</sup>Formeln für i.i.d.-Renditen ohne Verteilungsannahme, siehe Anhang: Treynor (A.22), Jensen (A.26),  $MM_{ori}$  (A.32),  $MM_{alt}$  (A.30)

<sup>9</sup>Formeln für i.i.d.-normalverteilte Renditen, siehe Anhang: Treynor (A.23), Jensen (A.27),  $MM_{ori}$  (A.33),  $MM_{alt}$  (A.31)

<sup>10</sup>Nummern der verwendeten Varianzformeln, siehe Anhang:

$\Delta Sh_{X,Y}$ : i.i.d.: A.20; i.i.d.-NV: A.21       $\Delta Tr_{X,Y}$ : i.i.d.: A.24 ; i.i.d.-NV: A.25

$\Delta Je_{X,Y}$ : i.i.d.: A.28 ; i.i.d.-NV: A.29       $\Delta MM_{X,Y}$ : i.i.d.: A.34 ; i.i.d.-NV: A.35

$\Delta\kappa(X,Y)_{z=0}^{\alpha}$ : nicht dargestellt, vgl. Erläuterung zu Formel 4.11, S. 91

schließlich vierten Grades endlich, so dass die Bedingungen für die Konvergenz zu den per Taylor-Entwicklung gewonnenen Formeln gerade erfüllt sind. In dieser Tabelle unterscheiden sich die theoretischen Varianzen  $\sigma_0(\hat{\theta})$  zwischen den Ansätzen mit und ohne Normalverteilungsannahme. Im ersteren Fall wird  $\sigma_0(\hat{\theta})$  weiterhin mit der jeweiligen Formel für normalverteilte Daten bestimmt, im zweiten Fall mit der allgemeineren, verteilungsfreien Formel. Da diese bei mehrdimensionalen Zeitreihen nicht vorgegebene höhere gemischte Momente enthält, wurden die theoretischen Werte hier simuliert. Per Simulation genäherte theoretische Werte werden in der Tabelle mit einem Sternchen gekennzeichnet.

Auch bei *i.i.d.*- $t_5$ -verteilten Zeitreihenwerten ist die Konvergenz zu Normalverteilung und asymptotischen Varianzen zumeist gut. Außer bei den Kappa-Maßen und ERVaR lehnt der Test auf Normalverteilung der Performancemaß-Schätzer nicht ab. Auch die jeweiligen Hypothesen  $\sigma^2 = \sigma_0^2$  werden nur in letzteren Fällen zum 5%-Niveau abgelehnt, zudem die asymptotische Varianz für die Differenz zweier Sharpe-Ratios, wenn (fälschlicherweise) von Normalverteilung ausgegangen wird. Ansonsten sind die Unterschiede zwischen den verteilungsfreien Ansätzen und jenen mit Normalverteilung ziemlich gering. Auffällig ist wiederum, dass der theoretische  $1,96\sigma$ -Bereich bei den Kappa-Maßen zwar sehr viel weniger als 95 % der Schätzwerte umfasst, dass aber die Überdeckungshäufigkeit der  $[t_i \pm 1,96\hat{\sigma}_i]$ -Konfidenzintervalle bei fast exakt 95 % liegt.

Insgesamt kann man bei der genannten Parameterkombination und der Zeitreihenlänge  $n = 500$  für die meisten Maße von guten Konvergenzeigenschaften sprechen. Es wurde allerdings nur eine einzige Parameterkombination simuliert und es wird sich in den kommenden Abschnitten zeigen, dass die Leistung mancher Performancemaßschätzer deutlich von den Eigenschaften der ausgewerteten Zeitreihen abhängt.

## 5.3 Ergebnisse für einzelne Maße

In diesem Teil werden die Ergebnisse für die einzelnen Performancemaße gesondert untersucht. Ziel ist es, die Eigenschaften der Performancemaß-Schätzer und die Qualität der Varianzermittlung bei kleinen Zeitreihenlängen und Nicht-*i.i.d.*-Zeitreihen zu prüfen.

### 5.3.1 Sharpe-Ratio

Für das Sharpe-Ratio wird eine univariate Zeitreihe  $X_t$  benötigt, die variierbaren Parameter sind Zeitreihenstruktur (*i.i.d.*, GARCH(1,1) oder AR(1)), Verteilungstyp, Zeitreihenlänge, Mittelwert und Standardabweichung. Für die letzten beiden Parameter kommt es nur auf das Verhältnis zueinander an, da das Sharpe-Ratio und alle Varianzformeln nur vom Verhältnis  $\mu_X/\sigma_X$  abhängen, nicht von den exakten Werten für  $\mu_X$  und  $\sigma_X$ . Deshalb fixieren wir  $\sigma_X$  in diesem Abschnitt auf den Wert 0,1.

Wir schauen uns zunächst (siehe Tabelle 5.8) die Ausgangsparameterkombination  $\mu_X = 0,002$ ,  $\sigma_X = 0,1$  mit  $n = 60$  und  $X_t \stackrel{i.i.d.}{\sim} N(\mu_X, \sigma_X)$  an, anschließend

Varianzformel	$\sigma_0(\hat{\theta})$	$\bar{s}(\hat{\theta}_i)$	$p_{\sigma^2=\sigma_0^2}$	$[u_{0,05}]$	$t_i$ in $u$	$KI_{95\%}$ um $\theta$
<b>Sharpe<sub>X</sub></b>	$\theta = 0,02$ ; $\bar{t} = 0,02045$ ; $s(t_i) = 0,04466$ ; $p_{AD} = 0,9128$ ; $\hat{\psi} = 0,0309$					
i.i.d., NV	0,04473	0,04475	0,7717	[-0,0677;0,1077]	94,90%	95,02%
i.i.d.	0,04474	0,04462	0,7188	[-0,0677;0,1077]	94,91%	94,69%
allg., $j = 1$	0,04474	0,04456	0,7280	[-0,0677;0,1077]	94,91%	94,51%
allg., $j = 5$	0,04474	0,04423	0,7280	[-0,0677;0,1077]	94,91%	93,93%
<b>Kappa<sub>X</sub><sup><math>\alpha=1</math></sup><sub><math>\bar{z}=0</math></sub></b>	$\theta = 0,04305$ ; $\bar{t} = 0,06489$ ; $s(t_i) = 0,12971$ ; $p_{AD} < 10^{-61}$ ; $\hat{\psi} = 0,055$					
i.i.d.	0,09292	0,13047	$< 10^{-99}$	[-0,1391;0,2251]	84,41%	95,07%
<b>Kappa<sub>X</sub><sup><math>\alpha=2</math></sup><sub><math>\bar{z}=0</math></sub></b>	$\theta = 0,02216$ ; $\bar{t} = 0,03265$ ; $s(t_i) = 0,06630$ ; $p_{AD} < 10^{-41}$ ; $\hat{\psi} = 0,0475$					
i.i.d.	0,04817	0,06647	$< 10^{-99}$	[-0,0723;0,1166]	84,76%	95,04%
<b>ERVaR<sub>X</sub>5%</b>	$\theta = 0,01298$ ; $\bar{t} = 0,01453$ ; $s(t_i) = 0,02994$ ; $p_{AD} < 10^{-29}$ ; $\hat{\psi} = 0,0478$					
i.i.d., param.	0,02955	0,03009	0,0070	[-0,0457;0,0717]	94,52%	95,07%
i.i.d., nonp.	0,02955	0,03011	0,0070	[-0,0457;0,0717]	94,52%	94,94%
<b>Treynor<sub>X</sub></b>	$\theta = 0,00267$ ; $\bar{t} = 0,00275$ ; $s(t_i) = 0,005970$ ; $p_{AD} = 0,9018$ ; $\hat{\psi} = 0,0417$					
i.i.d., NV	0,005964	0,005975	0,8926	[-0,0090;0,0144]	95,09%	95,38%
i.i.d.	0,005965*	0,005964	0,9134	[-0,0090;0,0144]	95,09%	95,09%
<b>Jensen<sub>X</sub></b>	$\theta = 0,0005$ ; $\bar{t} = 0,000535$ ; $s(t_i) = 0,002946$ ; $p_{AD} = 0,1417$ ; $\hat{\psi} = 0,0375$					
i.i.d., NV	0,002959	0,002950	0,5362	[-0,0053;0,0063]	94,72%	94,76%
i.i.d.	0,002958*	0,002943	0,5455	[-0,0053;0,0063]	94,72%	94,73%
<b>MM<sub>X,ori</sub></b>	$\theta = 0,002$ ; $\bar{t} = 0,00206$ ; $s(t_i) = 0,004465$ ; $p_{AD} = 0,8225$ ; $\hat{\psi} = 0,0383$					
i.i.d., NV	0,004473	0,004475	0,7996	[-0,0068;0,0108]	95,11%	95,40%
i.i.d.	0,004486*	0,004556	0,5033	[-0,0068;0,0108]	95,16%	95,36%
<b>MM<sub>X,alt</sub></b>	$\theta = 0$ ; $\bar{t} = 0,000031$ ; $s(t_i) = 0,003154$ ; $p_{AD} = 0,2393$ ; $\hat{\psi} = 0,044$					
i.i.d., NV	0,003163	0,003157	0,6806	[-0,0062;0,0062]	94,69%	94,87%
i.i.d.	0,003172*	0,003204	0,4160	[-0,0062;0,0062]	94,77%	95,26%
<b><math>\Delta Sh_{X,Y}</math></b>	$\theta = -0,01$ ; $\bar{t} = -0,0099$ ; $s(t_i) = 0,03176$ ; $p_{AD} = 0,2815$ ; $\hat{\psi} = 0,0404$					
i.i.d., NV	0,03239	0,03173	0,0052	[-0,0735;0,0535]	95,41%	95,04%
i.i.d.	0,03176*	0,03162	0,9823	[-0,0722;0,0522]	94,96%	94,95%
<b><math>\Delta \kappa(X,Y)_{\bar{z}=0}^{\alpha=1}</math></b>	$\theta = -0,02221$ ; $\bar{t} = -0,02903$ ; $s(t_i) = 0,0939$ ; $p_{AD} = 0,0002$ ; $\hat{\psi} = 0,049$					
i.i.d.	-	0,09428	-	-	-	95,71%
<b><math>\Delta \kappa(X,Y)_{\bar{z}=0}^{\alpha=2}</math></b>	$\theta = -0,01127$ ; $\bar{t} = -0,01469$ ; $s(t_i) = 0,04785$ ; $p_{AD} = 0,0015$ ; $\hat{\psi} = 0,0434$					
i.i.d.	-	0,04780	-	-	-	95,36%
<b><math>\Delta Tr_{X,Y}</math></b>	$\theta = -0,00133$ ; $\bar{t} = -0,00138$ ; $s(t_i) = 0,00425$ ; $p_{AD} = 0,6119$ ; $\hat{\psi} = 0,0392$					
i.i.d., NV	0,00422	0,00425	0,3061	[-0,0096;0,0069]	94,55%	94,81%
i.i.d.	0,00422*	0,00424	0,4055	[-0,0096;0,0069]	94,56%	94,82%
<b><math>\Delta Je_{X,Y}</math></b>	$\theta = -0,001$ ; $\bar{t} = -0,001027$ ; $s(t_i) = 0,00314$ ; $p_{AD} = 0,5682$ ; $\hat{\psi} = 0,4745$					
i.i.d., NV	0,00316	0,00316	0,5512	[-0,0072;0,0052]	95,50%	95,42%
i.i.d.	0,00317*	0,00315	0,4617	[-0,0072;0,0052]	95,54%	95,40%
<b><math>\Delta MM_{X,Y}</math></b>	$\theta = -0,001$ ; $\bar{t} = -0,001028$ ; $s(t_i) = 0,00315$ ; $p_{AD} = 0,6756$ ; $\hat{\psi} = 0,0454$					
i.i.d., NV	0,00316	0,00316	0,6930	[-0,0072;0,0052]	95,46%	95,49%
i.i.d.	0,00317*	0,00321	0,6391	[-0,0072;0,0052]	95,48%	95,79%
* aus den Daten simulierte Werte						
Erläuterung der Kürzel für die Statistiken in Tab. 5.2 auf S. 114.						

Tabelle 5.7: Simulationsergebnisse bei i.i.d.- $t_5$ -verteilten Daten und Zeitreihelänge  $n = 500$

Varianzformel	$\sigma_0(\hat{\theta})$	$\bar{s}(\hat{\theta}_i)$	$p_{\sigma^2=\sigma_0^2}$	$[u_{0,05}]$	$t_i$ in $u$	$KI_{95\%}$ um $\theta$
<b>Sharpe<sub>X</sub></b>	$\theta = 0,02$ ; $\bar{t} = 0,0205$ ; $s(t_i) = 0,1327$ ; $p_{AD} < 10^{-5}$ ; $\psi = 0,0232$					
i.i.d., NV	0,1291	0,1297	$< 10^{-99}$	[-0,2330;0,2730]	94,23%	95,15%
i.i.d.	0,1291	0,1297	$< 10^{-99}$	[-0,2330;0,2730]	94,23%	95,08%
allg., $j = 1$	0,1291	0,1274	$< 10^{-99}$	[-0,2330;0,2730]	94,23%	93,53%
allg., $j = 5$	0,1291	0,1209	$< 10^{-99}$	[-0,2330;0,2730]	94,23%	87,22%
Erläuterung der Kürzel für die Statistiken in Tab. 5.2 auf S. 114.						

Tabelle 5.8: Ergebnisse für das Sharpe-Ratio bei i.i.d.-normalverteilten Daten und Zeitreihenlänge  $n = 60$ , Ausgangsparameter

werden einzelne Parameter variiert. Die Anzahl der simulierten eindimensionalen Zeitreihen beträgt  $R = 100\,000$ .

Einzigiger Unterschied zu den Werten in Tabelle 5.4 ist die kleinere Zeitreihenlänge, dadurch ist die Varianz von  $\widehat{Sh}_X$  größer. Der Bias  $\bar{t} - \theta$  des Schätzers fällt nun etwas stärker ins Gewicht. Der Anderson-Darling-Test lehnt die Hypothese der Normalverteilung des Schätzers zwar ab, aber die Abweichung von der Normalverteilung scheint gering zu sein, was ein Blick auf das Histogramm der Verteilung (Abb. 5.2a) bestätigt.

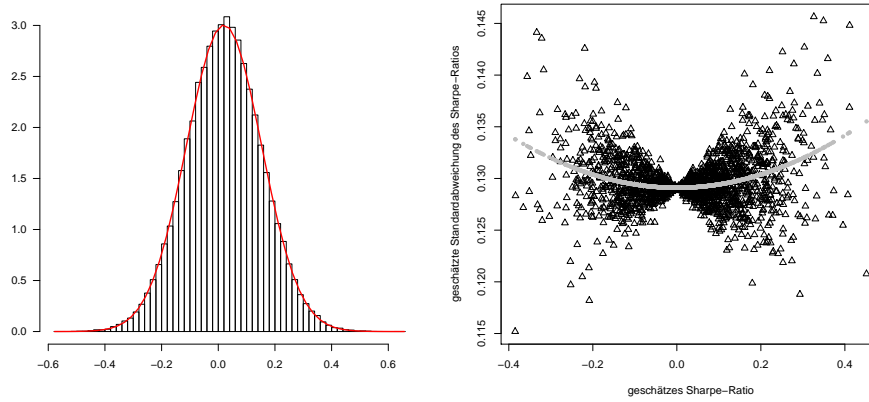
Die aus den Verteilungsparametern berechnete asymptotische Standardabweichung  $\sigma_0(\hat{\theta})$  beträgt 0,1291; die aus den Daten geschätzten Werte sind im Mittel mit 0,1297 für die beiden i.i.d.-Formeln sehr ähnlich. Die Stichprobenstandardabweichung der Schätzer liegt mit 0,1327 jedoch soweit höher, dass  $H_0 : \sigma = 0,1291$  zu jedem beliebigen Niveau abgelehnt wird. Da der Test bei der Zeitreihenlänge  $n = 500$  nicht abgelehnt hat (siehe Tabelle 5.4) und das Mittel der geschätzten Standardabweichungen  $\bar{s}(\hat{\theta}_i)$  mit der aus den Parametern berechneten Standardabweichung  $\sigma_0(\hat{\theta})$  konform geht, ist die Konvergenz bei  $n = 60$  offenbar noch nicht in ausreichendem Maße erreicht.

Da  $\sigma_0(\widehat{Sh})$  etwas zu niedrig ist, ist auch der Anteil der Schätzwerte, die in das  $1,96\sigma$ -Intervall fallen, trotz fast perfekter Normalverteilung mit 94,23 % etwas niedriger als 95 %. Dagegen liegt der Anteil der Konfidenzintervalle, die den wahren Wert  $Sh = 0,02$  einschließen, bei fast exakt 95 %. Dies kann man dadurch erklären, dass die Schätzformeln

$$n \hat{\sigma}_{\widehat{Sh},norm.}^2 = 1 + \frac{1}{2} \widehat{Sh}_X^2 \quad \text{und} \quad n \hat{\sigma}_{\widehat{Sh}}^2 = 1 - \widehat{Sh}_X \hat{\gamma}_{1,X} + \frac{1}{4} \widehat{Sh}_X^2 (\hat{\gamma}_{2,X} - 1)$$

jeweils den Schätzwert des Sharpe-Ratios enthalten. Die geschätzte Varianz unter Annahme normalverteilter Renditen ist eine quadratische Funktion des Schätzwerts des Sharpe-Ratios. Je größer der Betrag von  $\widehat{Sh}$ , desto größer ist die geschätzte Varianz bzw. Standardabweichung. In der verteilungsfreien Formel kommen Schätzwerte für Schiefe und Kurtosis der Verteilung hinzu, die für zusätzliche Streuung sorgen. Diese Verhältnisse sind in Abb. 5.2b dargestellt, in der für jede Zeitreihe<sup>11</sup> der Schätzwert des Sharpe-Ratios  $\widehat{Sh}$  gegen

<sup>11</sup>Aus Gründen der besseren Darstellung sind nur die Daten von 2000 statt  $R = 100\,000$

(a) Histogramm  $\widehat{Sh}$ (b)  $\widehat{Sh}$ - $\widehat{\sigma}_{\widehat{Sh}}$ -Diagramm

Links: Histogramm von  $\widehat{Sh}$  mit angepasster Dichtefunktion der Normalverteilung

Rechts: Punktdiagramm für 2000 Zeitreihen mit  $\widehat{Sh}$  (Abszisse) und zugehörigem  $\widehat{\sigma}_{\widehat{Sh}}$  (Ordinate) für die verteilungsfreie Varianzformel ( $\Delta$ ) und die Varianzformel mit Normalverteilungsannahme ( $\bullet$ )

Abbildung 5.2: Grafiken zur Verteilung von  $\widehat{Sh}$  für i.i.d.-normalverteilte Daten mit  $Sh = 0,02$  und Zeitreihenlänge  $n = 60$ .

den Schätzwert für die Standardabweichung des Schätzers,  $\widehat{\sigma}_{\widehat{Sh}}$ , abgetragen ist. Für die Normalverteilungsformel bilden die Punkte eine konvexe Kurve; für die verteilungsfreie Formel erkennt man eine um die Kurve herum liegende Punktwolke, die sich für  $|\widehat{Sh}| \rightarrow 0$  in Richtung  $\widehat{\sigma}_{\widehat{Sh}} = 1/\sqrt{n} = 1/\sqrt{60} \approx 0,1291$  zusammenzieht. Besonders extreme Schätzwerte gehen also mit erhöhten Varianzschätzungen einher. Wie am Vergleich der Skalen von Abszisse und Ordinate erkennbar ist, ist der Effekt nicht sehr stark. Während  $\widehat{Sh}$  etwa zwischen  $-0,4$  und  $0,4$  schwankt, bewegen sich die Schätzwerte der Standardabweichung zwischen  $0,115$  und  $0,145$ . Der Quartilsabstand von  $\widehat{Sh}$  beträgt  $0,1783$ , jener für  $\widehat{\sigma}_{\widehat{Sh}}$  nur  $0,0007$  (NV-Formel) bzw.  $0,0019$  (verteilungsfreie Formel). Die Varianzschätzung ist hier also im Vergleich zur Performancemaßschätzung ausgesprochen stabil.

Etwas anders sind die Ergebnisse der verallgemeinerten Formel für Nicht-*i.i.d.*-Daten. Für Tabelle 5.8 wurde die Standardabweichung wiederum für mögliche Autokorrelationen zu einer bzw. fünf Nachbarperioden berechnet. Hierbei steigt die Anzahl der geschätzten und anschließend aggregierten Terme deutlich, wodurch die Streuung der Schätzung nochmals steigt. In den Simulationsdaten für  $j = 1$  wurden Werte zwischen  $0,0227$  und  $0,1777$  geschätzt, mit Quartilsabstand  $0,0231$ . Bei  $j = 5$  resultierte aus der Addition der Komponenten in

---

simulierten Zeitreihen eingetragen.

Verteilung	$\sigma_0(\hat{\theta})$	$\bar{s}(\hat{\theta}_i)$	$p_{\sigma^2=\sigma_0^2}$	$[u_{0,05}]$	$t_i$ in $u$	$KI_{95\%}$ um $\theta$
<b>t<sub>10</sub></b>	$\theta = 0,02$ ; $\bar{t} = 0,02012$ ; $s(t_i) = 0,1327$ ; $p_{AD} < 10^{-5}$ ; $\hat{\psi} = 0,02533$					
i.i.d., NV	0,12911	0,12968	$< 10^{-99}$	[-0,2331;0,2731]	94,26%	95,11%
i.i.d.	0,12912	0,12901	$< 10^{-99}$	[-0,2331;0,2731]	94,26%	94,54%
allg., $j = 1$	0,12911	0,12686	$< 10^{-99}$	[-0,2331;0,2731]	94,26%	93,06%
allg., $j = 5$	0,12911	0,12040	$< 10^{-99}$	[-0,2331;0,2731]	94,26%	87,03%
<b>GARCH(1,1)</b>	$\theta = 0,02$ ; $\bar{t} = 0,03123$ ; $s(t_i) = 0,1331$ ; $p_{AD} = 0,3548$ ; $\hat{\psi} = 0,0230$					
i.i.d., NV	0,12911	0,12970	$< 10^{-99}$	[-0,2331;0,2731]	94,19%	95,10%
i.i.d.	0,12912	0,12745	$< 10^{-99}$	[-0,2331;0,2731]	94,13%	93,25%
allg., $j = 1$	0,13468*	0,12328	$< 10^{-7}$	[-0,2440;0,2840]	95,21%	90,47%
allg., $j = 5$	0,13468*	0,11534	$< 10^{-7}$	[-0,2440;0,2840]	95,21%	84,52%
<b>AR(1)</b>	$\theta = 0,02$ ; $\bar{t} = 0,0205$ ; $s(t_i) = 0,1711$ ; $p_{AD} < 10^{-6}$ ; $\hat{\psi} = 0,0214$					
i.i.d., NV	0,12911	0,13005	$< 10^{-99}$	[-0,2331;0,2731]	86,19%	87,60%
i.i.d.	0,12912	0,13000	$< 10^{-99}$	[-0,2331;0,2731]	86,19%	87,53%
allg., $j = 1$	0,17216	0,15574	0,0048	[-0,3174;0,3574]	95,04%	92,59%
allg., $j = 5$	0,17216	0,15311	0,0048	[-0,3174;0,3574]	95,04%	87,33%
* aus den Daten simulierte Werte						
Erläuterung der Kürzel für die Statistiken in Tab. 5.2 auf S. 114.						

Tabelle 5.9: Ergebnisse für das Sharpe-Ratio bei nicht normalverteilten Daten und Zeitreihenlänge  $n = 60$ , Ausgangsparameter

etwa 4 % der Fälle sogar ein negativer Wert als Schätzwert für die Varianz<sup>12</sup>, das Maximum der Werte lag bei 0,2592 und der Quartilsabstand war 0,048. Bei der kleineren Stichprobengröße  $n = 60$  werden die Ergebnisse der generalisierten Formel also schnell instabil, so dass man die Formel nicht anwenden sollte, wenn kein Verdacht auf zeitlich abhängige Renditen vorliegt – und wenn, dann mit kleinem Parameter  $j$ . Die Überdeckungshäufigkeiten der mittels dieser geschätzten Varianzen berechneten Konfidenzintervalle (letzte beiden Zeilen der Tabelle 5.8) weichen deutlich vom vorgegebenen Niveau ab.

Kommen wir nun zu Zeitreihen mit anderen Verteilungseigenschaften:

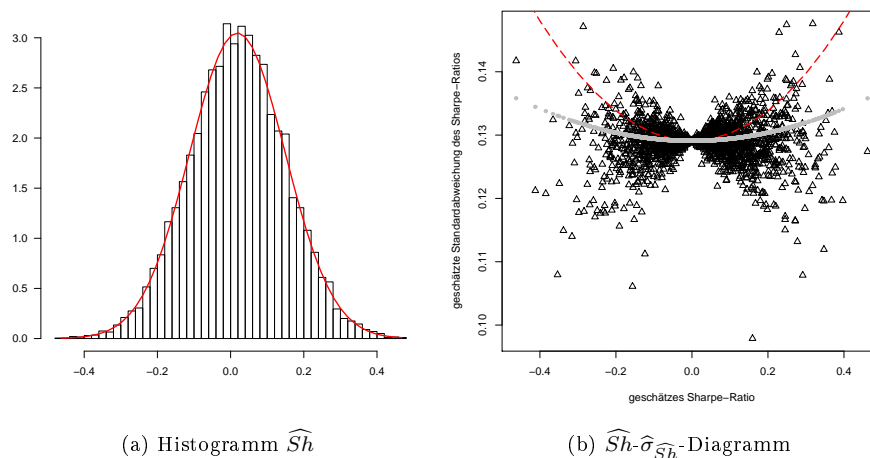
- i.i.d.- $t_{10}$ -verteilte Renditen,
- GARCH(1,1)-Prozesse mit  $\alpha_1 = 0,1$  (MA-Term) und  $\beta_1 = 0,8$  (AR-Term) sowie normalverteilten Störtermen,
- AR(1)-Prozesse mit  $\alpha_1 = 0,25$  und normalverteilten Störtermen.

Die einzelnen Prozesse sind jeweils so kalibriert, dass die unbedingten Momente der Renditen wie zuvor  $\mu_X = 0,002$  und  $\sigma_X = 0,1$  betragen. Die Ergebnisse stehen in Tabelle 5.9.

Die Unterschiede zwischen den Ergebnissen für normal- und  $t_{10}$ -verteilte Daten sind gering. Es gibt kaum Unterschiede zwischen den i.i.d.-Formeln mit und ohne Normalverteilungsannahme. In beiden Fällen liegt die empirische Standardabweichung  $s(t_i)$  erkennbar über den asymptotischen Näherungen, sowohl

<sup>12</sup>In diesem Fall wurde der negative Wert in den weiteren Berechnungen durch das Ergebnis der verteilungsfreien i.i.d.-Formel ersetzt.





Links: Histogramm von  $\widehat{Sh}$  mit angepasster Dichtefunktion der Normalverteilung

Rechts: Punktdiagramm für 2000 Zeitreihen mit  $\widehat{Sh}$  (Abszisse) und zugehörigem  $\widehat{\sigma}_{\widehat{Sh}}$  (Ordinate) für die verteilungsfreie Varianzformel( $\Delta$ ) und die Varianzformel mit Normalverteilungsannahme( $\bullet$ ). Gestrichelte Linie: Asymptotische Varianz gemäß Parametern und Varianzformel

Abbildung 5.3: Grafiken zur Verteilung von  $\widehat{Sh}$  für i.i.d.- $t_{10}$ -verteilte Daten mit  $Sh = 0,02$  und Zeitreihenlänge  $n = 60$ .

der theoretischen ( $\sigma_0(\hat{\theta})$ ) als auch dem Mittel der geschätzten ( $\overline{s}(\hat{\theta}_i)$ ). Die Näherung der empirischen Verteilung des Schätzers an die Normalverteilung ist trotz der im AD-Test abgelehnten Normalverteilungshypothese fast perfekt (siehe Abb. 5.3a). Im  $\widehat{Sh}-\widehat{\sigma}_{\widehat{Sh}}$ -Diagramm (Abb. 5.3b) folgen die Schätzwerte mittels der i.i.d.-NV-Formel weiter der theoretischen Vorgabe. Zusätzlich ist die theoretische asymptotische Standardabweichung in Abhängigkeit vom Sharpe-Ratio gemäß der allgemeinen i.i.d.-Formel als gestrichelte Kurve eingezeichnet. In der Grafik ist jedoch kein Zusammenhang zwischen den mittels dieser Formel geschätzten Standardabweichungen und der gestrichelten Kurve zu erkennen.

Die i.i.d.-Annahme, die bei der Herleitung der Varianzformel gemacht wurde, ist bei GARCH-Zeitreihen nicht mehr erfüllt. Insbesondere ist zu erwarten, dass die empirischen Varianzen der einzelnen Zeitreihen stark streuen. Zudem existiert das vierte Moment der Rendite bei der genannten Parameterkombination nicht.<sup>13</sup>

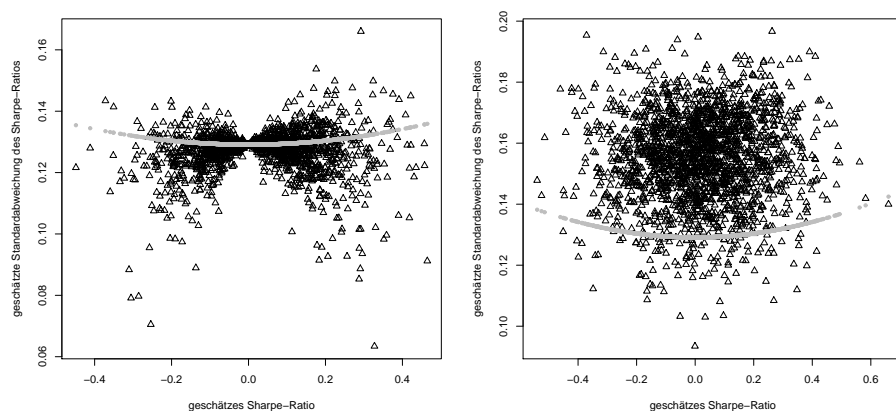
Beim Blick auf die Daten in Tabelle 5.9 fällt zunächst auf, dass der Mittelwert des geschätzten Sharpe-Ratios mit 0,03123 ein gutes Stück über dem theoretischen Wert von 0,02 liegt. Man könnte vermuten, dass einzelne Extremwerte aus der stark leptokurtischen Renditeverteilung den Durchschnitt verzerren.

<sup>13</sup>Das vierte Moment eines GARCH(1,1)-Prozesses mit normalverteilten Innovationen existiert, wenn  $3\alpha^2 + 2\alpha\beta + \beta^2 < 1$  (Bollerslev 1986). Hier:  $3 \cdot 0,8^2 + 2 \cdot 0,8 \cdot 0,1 + 0,1^2 = 2,09 \not< 1$ .

Das ist aber nicht der Fall, wie ein Blick auf Abb. 5.4a zeigt. Die Verzerrung ist systematisch und sie ist auch erklärbar, hier allgemein für positive erwartete Überrenditen  $\mu_X > 0$ : Die bedingten und unbedingten Erwartungswerte der Renditen aller Zeitreihen haben denselben Wert  $\mu_X$ . Die bedingten Varianzen der Zeitreihen hingegen differieren je nach Stärke der Ausschläge in den Vorperioden. Liegt nun eine Zeitreihe ohne große Ausschläge und somit geringer Varianz vor, liegen die Werte der Zeitreihe relativ eng um den positiven Erwartungswert  $\mu_X$ . Die geringe Variation führt zu einem kleinen Schätzwert  $\hat{\sigma}_X$  im Nenner und somit zu einem relativ deutlich positiven Sharpe-Ratio. Kommen dagegen starke Ausschläge vor, führen diese zu einem hohen  $\hat{\sigma}_X$  für die jeweilige Zeitreihe und  $\widehat{Sh}_X$  bleibt betragsmäßig klein. Zudem heben sich deutliche Ausschläge nach oben und unten hinsichtlich  $\hat{\mu}_X$  gegenseitig auf, aber nicht hinsichtlich  $\hat{\sigma}_X$ . Dazu passt, dass der Schätzer  $\widehat{Sh}_X$  ausweislich des  $p$ -Werts des Anderson-Darling-Tests annähernd perfekt normalverteilt ist. Die empirische Standardabweichung des Schätzers ist im Vergleich zum i.i.d.-Fall fast unverändert.

Die in der Tabelle angegebene asymptotische Standardabweichung  $\sigma_0(\hat{\theta})$  für den Non-i.i.d.-Ansatz liegt der empirischen Standardabweichung recht nahe, ist aber nicht mit den anderen Werten in dieser Spalte vergleichbar, da sie simuliert und nicht aus den Parametern berechnet wurde. Die Überdeckungshäufigkeit der damit gebildeten Konfidenzintervalle ist dagegen noch schlechter als im i.i.d.-Fall. Unter den i.i.d.-Ansätzen liefert die Formel mit Normalverteilungsannahme bessere Ergebnisse als jene ohne Verteilungsannahme. Das kann darauf zurückzuführen sein, dass in die verteilungsfreie Formel Schätzungen von Schiefe und Kurtosis der Renditeverteilung eingehen und der Schätzfehler dadurch steigt. Wie an der Skala von Abb. 5.4a erkennbar ist, streuen die Varianzschätzungen hier stärker als bei temporär unabhängigen Zeitreihen (vgl. Abb. 5.2b und 5.3b). Wenn die Non-i.i.d.-Formel für GARCH-Daten keine guten Ergebnisse liefert, muss man bedenken, dass sie nicht für die Behandlung autokorrelierter Varianzen entwickelt wurde, sondern für den Fall autokorrelierter Renditen. Diesen Fall untersuchen wir als nächstes (Tab. 5.9, letzter Abschnitt) anhand von AR(1)-Renditen.

Der Bias von  $\widehat{Sh}$  ist hier wie im i.i.d.-Fall eher geringfügig, jedoch ist die empirische Standardabweichung der Schätzwerte deutlich erhöht. Die Anpassung an die Normalverteilung ist gut, aber nicht perfekt. Die Varianzformeln für unabhängig und identisch verteilte Renditen berücksichtigen die zeitliche Abhängigkeit nicht, ergeben identische bzw. sehr ähnliche Werte wie bei den i.i.d.-Zeitreihen und unterschätzen damit die Varianz des Schätzers deutlich. Die Konfidenzintervalle sind dementsprechend zu schmal. Die Non-i.i.d.-Formel mit  $j = 1$ , die Autokorrelation erster Ordnung einbezieht, ist die für einen AR(1)-Prozess theoretisch exakt passende Formel und liefert bessere Ergebnisse. Ihr theoretischer Wert liegt sehr nahe an der empirischen Standardabweichung von  $\widehat{Sh}$ . Die Schätzungen aus den Daten heraus sind jedoch im Mittel etwas zu niedrig. Im Streudiagramm (Abb. 5.4b) ist erkennbar, dass die Schätzwerte für die Standardabweichung von  $\widehat{Sh}$  erheblich streuen und im Mittel deutlich über den

(a)  $\widehat{Sh}-\widehat{\sigma}_{\widehat{Sh}}$ -Diagramm GARCH(b)  $\widehat{Sh}-\widehat{\sigma}_{\widehat{Sh}}$ -Diagramm AR

Punkt-diagramme für 2000 Zeitreihen mit  $\widehat{Sh}$  (Abszisse) und zugehörigem  $\widehat{\sigma}_{\widehat{Sh}}$  (Ordinate). Renditezeitreihen aus GARCH(1,1)-Prozess (links) bzw. AR(1)-Prozess(rechts). Angegeben sind die verteilungsfreie i.i.d.-Varianzformel (links) bzw. die verallgemeinerte Varianzformel (non-i.i.d.) mit  $j = 1$  (rechts) ( $\Delta$ ) sowie die Varianzformel mit Normalverteilungsannahme( $\bullet$ ).

Abbildung 5.4: Grafiken zur Verteilung von  $\widehat{Sh}$  für GARCH- und AR-Zeitreihen mit  $Sh = 0,02$  und Zeitreihenlänge  $n = 60$ .

Werten der i.i.d.-NV-Formel liegen. Auch ist hier kein Zusammenhang zwischen dem Betrag von  $\widehat{Sh}$  und dem Schätzwert der Standardabweichung zu erkennen. Die Überdeckungshäufigkeit der Konfidenzintervalle von 92,59 % ist vergleichbar mit den Ergebnissen dieses Ansatzes im i.i.d.-Fall. Offenbar kann man mit der Non-i.i.d.-Formel die Autokorrelation der Zeitreihe korrekt verarbeiten, muss aber bei einer Zeitreihenlänge von 60 Renditen weiterhin eine hohe Schätzungenauigkeit in Kauf nehmen. Versucht man die Autokorrelation bis zur fünften Ordnung in die Varianzschätzung miteinzubeziehen ( $j = 5$ ), sind die Ergebnisse hinsichtlich der durchschnittlichen Schätzvarianz mit dem ( $j = 1$ )-Ansatz vergleichbar, aber die Überdeckungshäufigkeit der Konfidenzintervalle ist schlecht.

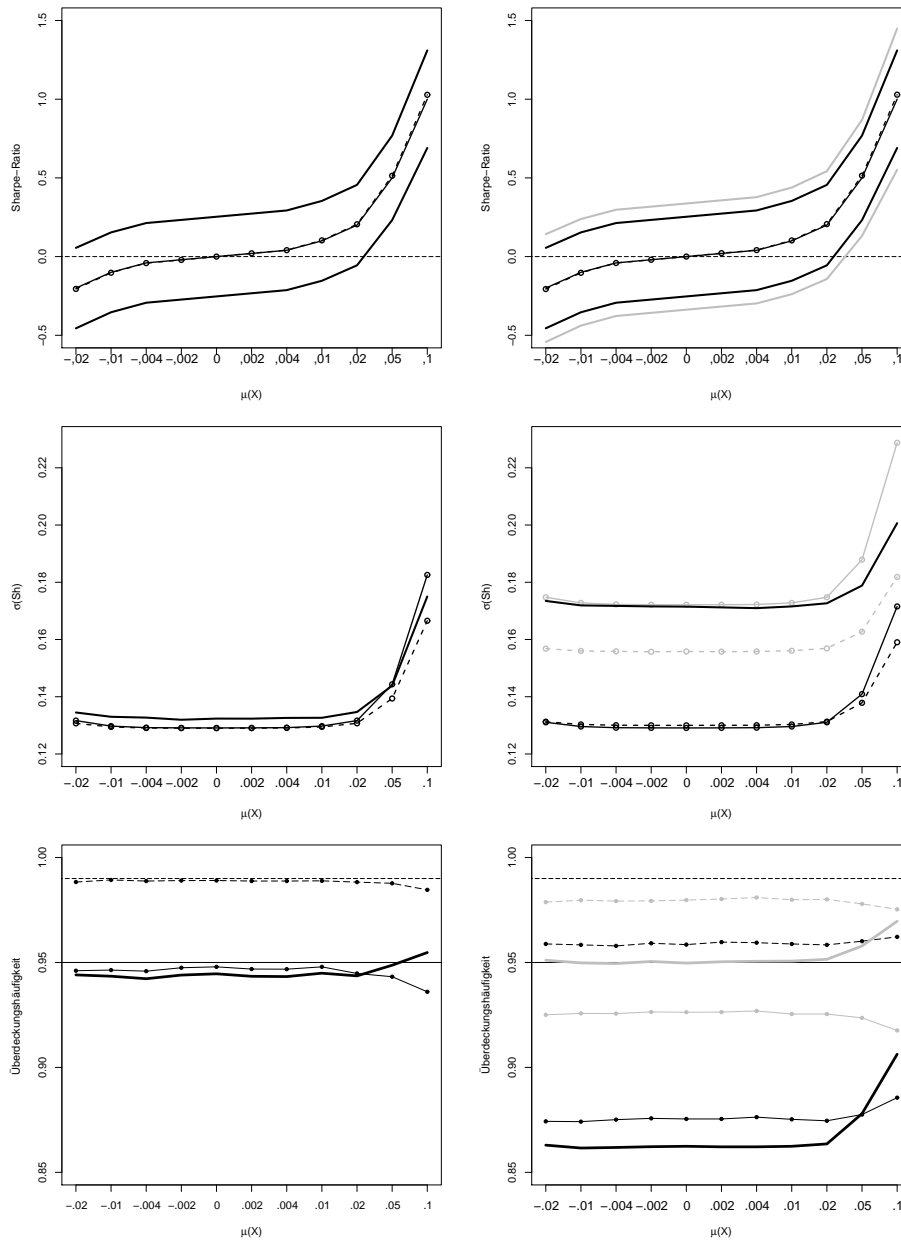
Nun verlassen wir die Referenz-Parameterkombination und lassen den Parameter  $\mu_X$  Werte zwischen -0,02 und 0,1 annehmen; alle anderen Parameter werden beibehalten. Das Wertespektrum ist in ökonomischer Hinsicht sehr weit gewählt, bereits Monats-Überrenditen von 1 % sind kaum realistisch, aber es soll hier in erster Linie um die statistischen Eigenschaften des Maßes gehen. In Abbildung 5.5 sind die wichtigsten der bisher in Tabellen aufgelisteten Simulationsstatistiken visualisiert. Von den sechs Tafeln der Abbildung bezieht sich die linke Hälfte auf i.i.d.- $t_{10}$ -verteilte Renditen, die rechte Hälfte zeigt Ergebnisse für AR(1)-Renditezeitreihen mit denselben Parametern wie im vorigen Abschnitt. In allen sechs Tafeln wird auf der Abszisse die erwartete Rendite  $\mu_X$  variiert,

wobei die Abstände nicht maßstabgetreu sind.

Die oberen beiden Tafeln zeigen als durchgezogene Linie die theoretischen Werte des Sharpe-Ratios bei den jeweiligen Werten für  $\mu_X$ ; als gestrichelte Linie mit Datenpunkten liegt das Mittel der Schätzwerte fast genau darüber und ist in beiden Tafeln vom theoretischen Wert kaum zu unterscheiden. Die Verzerrung des Schätzers ist also relativ gering. Um die Linie herum ist in schwarz der  $1,96\sigma$ -Bereich gemäß der asymptotischen i.i.d.-Varianzformel eingezeichnet, dessen Breite sich über den Wertebereich kaum verändert. Die rechte Tafel enthält zusätzlich den  $1,96\sigma$ -Bereich nach verallgemeinerter Varianzformel mit  $j = 1$  (grau), der über den gesamten Wertebereich etwas breiter ist.

Die beiden mittleren Tafeln zeigen Standardabweichungen und deren Schätzungen. Links ( $t_{10}$ -Daten) zeigt die starke schwarze Linie ohne Datenpunkte die empirischen Standardabweichungen  $s(t_i)$  der Performancemaß-Schätzwerte bei verschiedenen erwarteten Renditen an. Die durchgezogene bzw. die gestrichelte Linie mit Datenpunkten zeigen die aus den Parametern berechnete asymptotische Standardabweichung  $\sigma_0(\hat{\theta})$  bzw. die Mittelung aus den Schätzungen der Standardabweichungen  $\bar{s}(\hat{\theta}_i)$ . Die Werte bleiben für den ökonomisch relevanten Bereich von  $-0,01$  bis  $+0,01$  nahezu konstant, wobei die empirische Standardabweichung stets etwas höher ist als die asymptotische und die aus den Daten geschätzte Näherung. Erst für stark von Null abweichendes  $\mu_X$  steigt die Standardabweichung des Schätzers merklich an. Rechts ( $AR(1)$ -Daten) sind in grauer Farbe zusätzlich die entsprechenden Werte von  $\sigma_0(\hat{\theta})$  und  $\bar{s}(\hat{\theta}_i)$  aus der Non-i.i.d.-Varianzformel mit  $j = 1$  eingezeichnet. Die Schätzwerte gemäß der i.i.d.-Formel sind hier nahezu gleich denen in der linken Tafel. Die empirische Standardabweichung der Schätzer liegt konstant etwa  $0,03$  höher als im i.i.d.-Fall, was von der verallgemeinerten asymptotischen Varianzformel für den ökonomisch relevanten Bereich sehr gut nachvollzogen wird, wenn man die Standardabweichung aus den Parametern berechnet. Sind die Parameter unbekannt und schätzt man die Standardabweichungen aus den Daten, sind die Schätzwerte tendenziell etwas zu niedrig, jedoch deutlich besser als die Ergebnisse der i.i.d.-Formel.

Die untere Reihe schließlich zeigt empirische Anteile, entweder den Anteil der Schätzwerte, die in einem  $1,96\sigma$ -Bereich liegen oder der Anteil der Konfidenzintervalle, die den wahren Wert umschließen. Links zeigt die breite schwarze Linie, dass meist geringfügig weniger als  $95\%$  der Schätzwerte innerhalb der  $1,96\sigma$ -Bereiche liegen, was zu erwarten war, da die asymptotische Wert der Standardabweichung etwas unter der empirischen Standardabweichung liegt (siehe mittlere Reihe). Die Überdeckungshäufigkeiten der Konfidenzintervalle zu  $95\%$  (durchgezogen, mit Datenpunkten) und  $99\%$  (gestrichelt) liegen ebenfalls geringfügig unter dem vorgegebenen Niveau. Rechts liegen alle mittels der i.i.d.-Formel berechneten Statistiken weit unter den jeweiligen Zielwerten. Die grau eingezeichneten Werte für die Non-i.i.d.-Formel sind deutlich besser. Allerdings ist nur der aus den Parametern berechnete  $1,96\sigma$ -Bereich relativ exakt, die aus den Daten geschätzten Konfidenzintervalle haben zu niedrige Überdeckungshäufigkeiten (ca.  $93\%$  und  $98\%$  anstatt  $95\%$  und  $99\%$ ).



*Erste Reihe:* Theoretischer Wert  $Sh$  und Mittelwert der Schätzungen (praktisch deckungsgleich); umliegend die theoretische  $1,96 - \sigma$ -Umgebung mit der i.i.d.-Formel (schwarz) und der allgemeinen Formel mit  $j = 1$  (nur rechts, grau).

*Zweite Reihe:* Empirische Standardabweichung der Schätzer (durchgehende Linie, ohne Datenpunkte); asymptotische Standardabweichung (durchgehend, mit Datenpunkten) und mittlere geschätzte Standardabweichung (gestrichelt) mittels i.i.d.-Formel, rechts auch in grau für die allgemeine Formel mit  $j = 1$ .

*Dritte Reihe:* Anteil der Schätzwerte in  $1,96\sigma$ -Umgebung (breite Linien), Überdeckungshäufigkeit des 95%-KI (durchgehende Linie, mit Datenpunkten) und des 99%-KI (gestrichelte Linie, m.D.) für die i.i.d.-Formel (schwarz); rechts auch in grau für die allgemeine Formel mit  $j = 1$ .

Abbildung 5.5: Simulationsstatistiken des Sharpe-Ratios bei Variation von  $\mu_X$  für i.i.d.-Daten (links) und AR(1)-Daten (rechts)

Verteilung	$\sigma_0(\hat{\theta})$	$\bar{s}(\hat{\theta}_i)$	$p_{\sigma^2=\sigma_0^2}$	$[u_{0,05}]$	$t_i$ in $u$	$KI_{95\%}$ um $\theta$
<b>NV</b>	$\theta = 0,01$ ; $\bar{t} = 0,0107$ ; $s(t_i) = 0,09389$ ; $p_{AD} = 0,0249$ ; $\hat{\psi} = 0,0292$					
i.i.d., NV	0,09351	0,09280	0,2070	[-0,1733;0,1933]	95,01%	94,99%
i.i.d.	0,09351	0,09191	0,2070	[-0,1733;0,1933]	95,01%	94,61%
<b>t<sub>10</sub></b>	$\theta = 0,01$ ; $\bar{t} = 0,0098$ ; $s(t_i) = 0,09432$ ; $p_{AD} = 0,0047$ ; $\hat{\psi} = 0,0261$					
i.i.d., NV	0,09351	0,09289	0,0065	[-0,1733;0,1933]	94,78%	94,93%
i.i.d.	0,09143*	0,09169	$< 10^{-99}$	[-0,1692;0,1892]	94,26%	94,47%
<b>t<sub>5</sub></b>	$\theta = 0,01$ ; $\bar{t} = 0,0104$ ; $s(t_i) = 0,09469$ ; $p_{AD} = 0,0002$ ; $\hat{\psi} = 0,0208$					
i.i.d., NV	0,09351	0,09322	$< 10^{-4}$	[-0,1733;0,1933]	94,63%	95,07%
i.i.d.	0,09162*	0,09143	$< 10^{-99}$	[-0,1696;0,1896]	94,15%	94,32%
<b>GARCH(1)</b>	$\theta = 0,01$ ; $\bar{t} = 0,0109$ ; $s(t_i) = 0,09405$ ; $p_{AD} = 0,0098$ ; $\hat{\psi} = 0,0245$					
i.i.d., NV	0,09351	0,09282	0,0699	[-0,1733;0,1933]	94,80%	95,11%
i.i.d.	0,09143*	0,09185	$< 10^{-99}$	[-0,1692;0,1892]	94,32%	94,63%
* aus den Daten simulierte Werte						
Erläuterung der Kürzel für die Statistiken in Tab. 5.2 auf S. 114.						

Tabelle 5.10: Ergebnisse für die Differenz der Sharpe-Ratios zweier Portfolios  $X$  und  $Y$  mit Zeitreihenlänge  $n = 60$ , Ausgangsparameter

Insgesamt kann man feststellen, dass die weiter oben anhand der Standard-Parameterkombination getroffenen Aussagen im ökonomisch relevanten Bereich qualitativ nicht von der erwarteten Überrendite der Anlage abhängen. Für betragsmäßig sehr hohe theoretische Werte des Sharpe-Ratios scheint das Verhalten der Schätzer zumindest bei  $n = 60$  deutlicher von der Asymptotik abzuweichen; insbesondere nimmt die Abweichung der Verteilung der Schätzwerte von der Normalverteilung zu.

Bei der Simulation der **Differenz des Sharpe-Ratios zweier Portfolios**  $X$  und  $Y$  werden Renditetupel  $(X, Y)$  simuliert. Außer den Parametern für die Zeitreihenlänge, die Zeitreihenstruktur und den Verteilungstyp der multivariaten Verteilung haben wir fünf Parameter für die ersten beiden (gemischten) Momente:  $\mu_X, \mu_Y, \sigma_X, \sigma_Y$  und  $\rho_{X,Y}$ . Wie beim Sharpe-Ratio für ein einzelnes Portfolio wird zunächst eine Ausgangs-Parameterkombination

Variable	$\mu$	$\sigma$	Korrelationsmatrix	
$X$	0,003	0,1	1	0,75
$Y$	0,002	0,1	0,75	1
$n = 60, R = 50\,000$				

genauer untersucht; anschließend werden einzelne Parameter variiert.

Eine grobe Vorstellung von der gemeinsamen Verteilung von  $\widehat{Sh}_X$  und  $\widehat{Sh}_Y$  für die angegebene Kombination ermöglicht (für i.i.d.- $t_{10}$ -verteilte Renditen) das Streudiagramm Abb. 5.6a mit 95%- und 99%-Zentralbereichen einer angepassten bivariaten Normalverteilung<sup>14</sup>.

<sup>14</sup>Hierbei handelt es sich um die Zentralbereiche einer Normalverteilung, deren Verteilungsparameter aus der Punktwolke geschätzt werden.

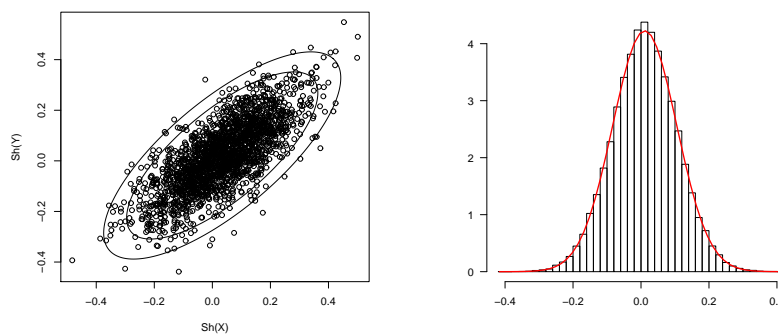
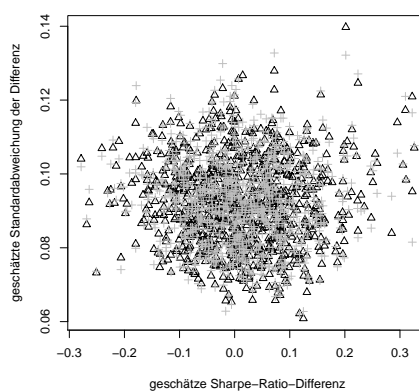
Detaillierte Ergebnisse für *i.i.d.*-normal-,  $-t_{10}$ -,  $-t_5$ - und  $MGARCH(1)$ <sup>15</sup>-verteilte Renditen sind in Tabelle 5.10 aufgelistet. Sie unterscheiden sich für die einzelnen Zeitreihen- und Verteilungstypen nur graduell. Der Schätzer  $\hat{\Delta}_{Sh}$  ist tendenziell geringfügig nach oben verzerrt und gemäß der Verteilungsstatistiken  $p_{AD}$  und  $\hat{\psi}$  fast exakt normalverteilt (siehe auch Abb. 5.6b). Hinsichtlich der Varianzformeln schneidet die Formel für normalverteilte Werte gut ab. Ihre Werte kommen der empirischen Standardabweichung sehr nahe und der  $\chi^2$ -Test auf  $H_0 : \sigma^2 = \sigma_0^2$  lehnt nur bei den  $t$ -verteilten Renditen ab, und das angesichts der hohen Zahl von Simulationsläufen nur knapp. Entsprechend ist auch der Anteil der Schätzwerte innerhalb des  $u_{95\%}$ -Bereichs nahe 95 %. Im Mittel liegen die geschätzten Standardabweichungen ( $\overline{s}(\hat{\theta}_i)$ ) etwas unter der empirischen Standardabweichung der Schätzwerte  $s(t_i)$ , die Überdeckungshäufigkeit der Konfidenzintervalle bleibt jedoch gut. Etwas anders verhält es sich mit den Ergebnissen für die Varianzformel ohne Verteilungsannahme. Die (simulierten) asymptotischen Standardabweichungen sowie die mittlere geschätzte Standardabweichung sind erkennbar niedriger als die empirische Standardabweichung der Schätzwerte. Auch die Überdeckungshäufigkeiten der Konfidenzintervalle liegen etwa einen halben Prozentpunkt unter dem vorgegebenen Niveau.

In Abb. 5.6c wird sichtbar, dass die Schätzungen der Standardabweichungen  $\hat{\sigma}_{\hat{\Delta}_{Sh}}$  für beide Varianzformeln keine erkennbare Abhängigkeit von  $\hat{\Delta}_{Sh}$  aufweisen. Sie streuen aber deutlich stärker als die Schätzungen der Standardabweichung von  $\hat{Sh}$  (vgl. Abb. 5.3b, S. 125).

Die Ergebnisse für verschiedene Parameterkombination sind (im Wesentlichen analog zu Abb. 5.5) in Abb. 5.7 dargestellt. Hierbei wird von der oben verwendeten Parameterkombination ( $\mu_X = 0,003, \mu_Y = 0,002, \sigma_X = \sigma_Y = 0,1, \rho_{X,Y} = 0,75$ ) ausgegangen und jeweils einer der drei Parameter  $\mu_X, \sigma_X$  und  $\rho_{X,Y}$  variiert, entsprechend sind die drei Tafeln der Grafik horizontal in drei Abschnitte aufgeteilt. Die Ausgangsparameterkombination ist jeweils mit einer senkrechten gestrichelten Linie gekennzeichnet. Das verwendete Verteilungsmodell der Simulationsdaten ist ein  $MGARCH$ -Prozess, dasjenige unter den verwendeten Modellen, das realen Finanzmarktdaten wohl am nächsten kommt. Die verwendeten Varianzformeln sind die *i.i.d.*-Formel für beliebig verteilte Daten und jene für *i.i.d.*-normalverteilte Daten. Bei einer Zeitreihenlänge von  $n = 60$  wurden je 50 000 Simulationsläufe durchgeführt.

In der oberen Tafel stellen die mittleren, optisch praktisch nicht unterscheidbaren Linien den theoretischen Wert sowie das Mittel der Schätzwerte von  $\Delta_{Sh}$  dar. Die beiden äußeren Linien sind die Grenzen der theoretischen  $[\Delta_{Sh} \pm 1,96\sigma_{\hat{\Delta}_{Sh},0}]$ -Intervalle. Die Differenz der Sharpe-Ratios verändert sich mit  $\mu_X$  (steigende Überrendite von  $X$  führt zu höherem  $Sh(X)$ ) und  $\sigma_X$  (bei positivem  $\mu_X$  nimmt  $Sh(X)$  mit steigendem  $\sigma_X$  ab), während es bei Variationen von  $\rho_{X,Y}$  konstant bleibt. Die Verzerrung von  $\hat{\Delta}_{Sh}$  beträgt bei allen untersuchten Parameterkombinationen nur wenige Prozent. Die Breite des (theoretischen)

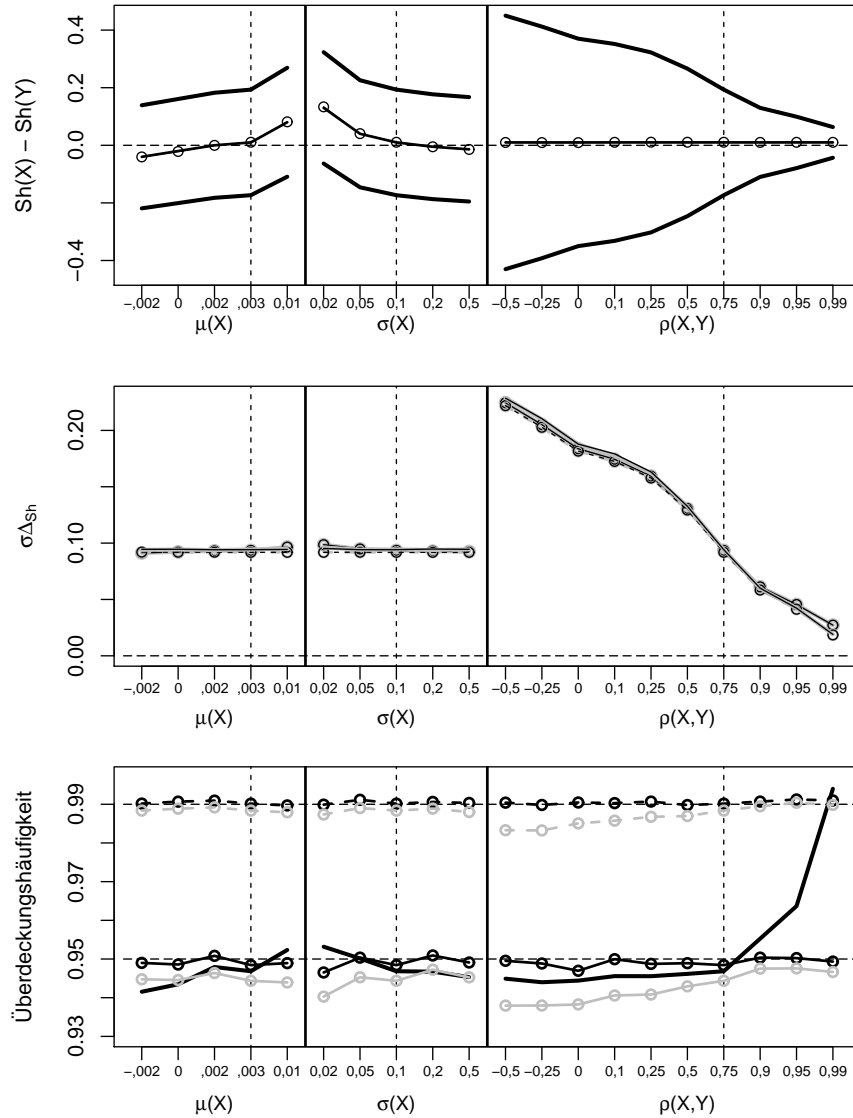
<sup>15</sup>gemäß Modell 2.1, siehe S. 20, mit  $\beta_1 = 0,8, \alpha_1 = 0,1$  und normalverteilten Innovationen

(a)  $\widehat{Sh}_X$ - $\widehat{Sh}_Y$ -Streudiagramm(b) Histogramm  $\widehat{\Delta}_{Sh}$ (c)  $\widehat{\Delta}_{Sh}$ - $\widehat{\sigma}_{\widehat{\Delta}_{Sh}}$ -Diagramm

oben links: Streudiagramm  $\widehat{Sh}_X$ - $\widehat{Sh}_Y$  mit angepassten 95%- und 99%-NV-Zentralbereichen  
 oben rechts: Histogramm von  $\widehat{\Delta}_{Sh}$  mit angepasster Normalverteilung  
 unten:  $\widehat{\Delta}_{Sh}$ (Abszisse) gegen  $\widehat{\sigma}_{\widehat{\Delta}_{Sh}}$  (Ordinate) für 1000 Zeitreihen. Varianzschätzung nach Formel für normalverteilte Renditen ( $\Delta$ ) bzw. ohne Verteilungsannahme (+).

Abbildung 5.6: Grafiken zur Verteilung von  $\widehat{\Delta}_{Sh}$ , i.i.d.- $t_{10}$ -verteilte Renditen, Zeitreihenlänge  $n = 60$ .





Verteilungsstatistiken bei Variation der Simulationsparameter  $\mu_X$ ,  $\sigma_X$  und  $\rho_{X,Y}$ . Ausgangsparameterkombination:  $\mu_X = 0,003$ ,  $\mu_Y = 0,002$ ,  $\sigma_X = \sigma_Y = 0,1$ ,  $\rho_{X,Y} = 0,75$  (jeweils senkrechte gestrichelte Linien).

Daten: MGARCH(1)-Prozesse,  $n = 60$ , je 50 000 Simulationsläufe.

*oben:* Theoretischer Wert  $\Delta_{Sh}$  und Mittelwert der Schätzungen (praktisch deckungsgleich); umliegend die theoretische  $1,96 - \sigma$ -Umgebung mit der i.i.d.-NV-Formel.

*Mitte:* Empirische Standardabweichung der Schätzer (durchgehende Linie, ohne Datenpunkte); asymptotische Standardabweichung (durchgehend, mit Datenpunkten) und mittlere geschätzte Standardabweichung (gestrichelt) mittels i.i.d.-NV-Formel (schwarz), und verteilungsfreier i.i.d.-Formel (grau).

*unten:* Anteil der Schätzwerte in  $1,96\sigma$ -Umgebung nach i.i.d. NV-Formel (Linie ohne Datenpunkte), Überdeckungshäufigkeit des 95%-KI (durchgehende Linie mit Datenpunkten) und des 99%-KI (gestrichelte Linie, m.D.) für die i.i.d.-NV-Formel (schwarz) und die verteilungsfreie i.i.d.-Formel (grau).

Abbildung 5.7: Grafiken zu  $\hat{\Delta}_{Sh}$  bei Parametervariation

Streuungsbereichs nimmt mit steigender Korrelation ab.

In der mittleren Tafel sind vier verschiedene Statistiken der Standardabweichung abgetragen: Die asymptotische Standardabweichung nach der i.i.d.-NV-Formel, die empirische Standardabweichung der Schätzwerte und die gemittelten geschätzten Standardabweichungen nach i.i.d.-NV-Formel und verteilungsfreier i.i.d.-Formel. Alle vier Kurven liegen bis auf wenige Ausnahmen nahezu genau aufeinander. Für Änderungen von  $\mu_X$  und  $\sigma_X$  bleibt die Standardabweichung fast konstant, lediglich bei  $\mu_X = 0,01$  sowie  $\sigma_X = 0,02$  liegt die Mittelung der nach der verteilungsfreien Formel geschätzten Standardabweichungen geringfügig höher. Das liegt daran, dass hier  $Sh_X$  relativ hoch ist und die Standardabweichung von  $\widehat{Sh}_X$  mit  $|Sh_X|$  steigt.<sup>16</sup> Bei sehr hoher Korrelation von  $X$  und  $Y$  ist außerdem die theoretische Standardabweichung erkennbar höher als die empirische und die mittlere geschätzte Standardabweichung.

Die untere Tafel vergleicht die Anpassung von Häufigkeitsstatistiken an vorgegebene Niveaus von 95% bzw. 99%. Die schwarze Linie ohne Datenpunkte stellt den Anteil der in den theoretischen  $1,96\sigma$ -Bereich um  $Sh_X$  fallenden Schätzwerte dar. In den meisten Fällen wird die Marke von 95 % leicht unterschritten; bei sehr hohen Korrelationen fallen deutlich mehr als 95 % der Schätzwerte in das Intervall, was also offenbar (siehe auch voriger Absatz) zu breit ist. Die Linien mit Datenpunkten stellen Überdeckungshäufigkeiten des wahren Wertes durch Konfidenzintervalle dar. Durchgehende Linien beziehen sich auf 95 %-Intervalle, gestrichelte Linien auf 99 %-Intervalle. In schwarz sind Werte zu Intervallen auf Basis der i.i.d.-NV-Formel abgebildet, graue Werte beziehen sich auf Intervalle auf Basis der verteilungsfreien Formel. Offenbar liefert die Formel mit Normalverteilungsannahme verlässlich die besseren Ergebnisse. Auf ihr basierende Konfidenzintervalle halten die vorgegebenen Niveaus ziemlich verlässlich ein, ohne dass systematische Verzerrungen oder Trends bei bestimmten Parametern erkennbar wären. Konfidenzintervalle auf Basis der verteilungsfreien Varianzformel schließen den wahren Wert dagegen etwas zu selten ein. Insbesondere bei schwach oder negativ korrelierten Renditen wird das vorgegebene Niveau merklich unterschritten, die Überdeckungshäufigkeiten liegen hier eher bei knapp 94% bzw. gut 98%.

In der Grafik nicht dargestellt ist die Anpassungsgüte der Schätzer an die Normalverteilung. Diese ist, mit  $\widehat{\psi}$  gemessen, durchweg gut. Die p-Werte der AD-Tests sind in den meisten Fällen im Annahmehereich der gängigen Testniveaus oder zumindest größer als 0,0001. Lediglich bei hohen Werten für  $Sh_X$  sowie bei hoher Korrelation von  $X$  und  $Y$  sind die p-Werte kleiner.

### 5.3.2 Treynor-Ratio

Die Ergebnisse für verschiedene Verteilungsmodelle bei der Parameterkombination

<sup>16</sup>vgl. Abb. 5.5, mittlere Reihe links

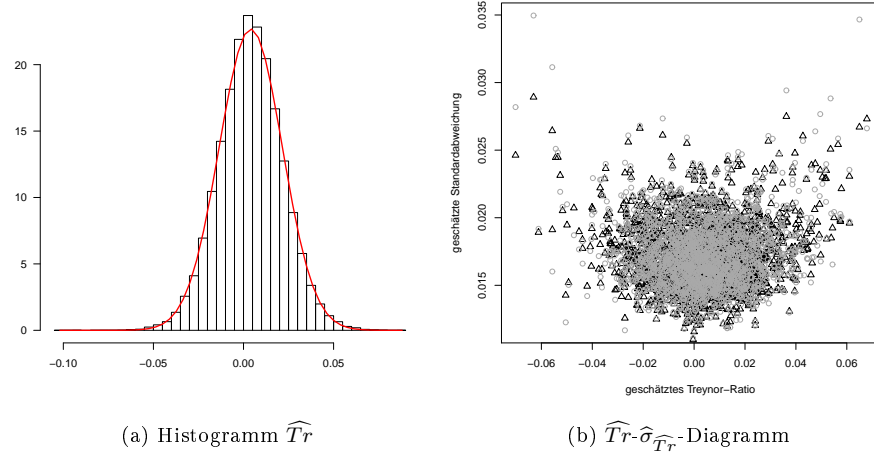
Verteilung	$\sigma_0(\hat{\theta})$	$\bar{s}(\hat{\theta}_i)$	$p_{\sigma^2=\sigma_0^2}$	$[u_{0,05}]$	$t_i$ in $u$	$KI_{95\%}$ um $\theta$
<b>NV</b>	$\theta = 0,004$ ; $\bar{t} = 0,00413$ ; $s(t_i) = 0,0176$ ; $p_{AD} < 10^{-10}$ ; $\hat{\psi} = 0,0293$					
i.i.d., NV	0,01722	0,01731	$< 10^{-10}$	[-0,0298;0,0378]	94,32%	95,47%
i.i.d.	0,01722	0,01717	$< 10^{-10}$	[-0,0298;0,0378]	94,31%	94,24%
<b>t<sub>10</sub></b>	$\theta = 0,004$ ; $\bar{t} = 0,00415$ ; $s(t_i) = 0,0177$ ; $p_{AD} < 10^{-9}$ ; $\hat{\psi} = 0,0318$					
i.i.d., NV	0,01722	0,01735	$< 10^{-15}$	[-0,0298;0,0378]	94,27%	95,40%
i.i.d.	0,01722*	0,01720	$< 10^{-14}$	[-0,0298;0,0378]	94,28%	94,07%
<b>t<sub>5</sub></b>	$\theta = 0,004$ ; $\bar{t} = 0,00412$ ; $s(t_i) = 0,0189$ ; $p_{AD} < 10^{-99}$ ; $\hat{\psi} = 0,0475$					
i.i.d., NV	0,01722	0,07520	$< 10^{-99}$	[-0,0298;0,0378]	94,22%	95,55%
i.i.d.	0,01725*	0,20741	$< 10^{-99}$	[-0,0298;0,0378]	94,25%	93,99%
<b>MGARCH</b>	$\theta = 0,004$ ; $\bar{t} = 0,00408$ ; $s(t_i) = 0,0176$ ; $p_{AD} < 10^{-13}$ ; $\hat{\psi} = 0,0304$					
i.i.d., NV	0,01722	0,01731	$< 10^{-11}$	[-0,0298;0,0378]	94,32%	95,55%
i.i.d.	0,01721*	0,01716	$< 10^{-11}$	[-0,0297;0,0377]	94,31%	94,24%
* aus den Daten simulierte Werte						
Erläuterung der Kürzel für die Statistiken in Tab. 5.2 auf S. 114.						

Tabelle 5.11: Ergebnisse für das Treynor-Ratio bei Zeitreihenlänge  $n = 60$ , Ausgangsparameter

Variable	$\mu$	$\sigma$	Korrelationsmatrix	
$M$	0,002	0,1	1	0,75
$X$	0,003	0,1	0,75	1
$n = 60, R = 50\,000$				

sind in Tabelle 5.11 zu finden. Erneut kann man eine leichte Verzerrung des Schätzers nach oben beobachten. Die Hypothese normalverteilter Schätzer wird jeweils klar abgelehnt. Aber nur bei den  $t_5$ -verteilten Daten signalisiert  $\hat{\psi}$  durch einen erhöhten Wert, dass die Schätzerverteilung deutlicher von der Normalverteilung abweicht. Der Ursprung der Abweichung ist im Histogramm Abb. 5.8a für  $t_{10}$ -verteilte Daten implizit durch die Breite der Abszisse erkennbar. Während der Hauptteil der Verteilung eine gute Anpassung an die Normalverteilung zeigt, liegen einzelne Schätzwerte deutlich abseits. Konkret liegen 50% der Werte zwischen den beiden Quartilen -0,0074 und 0,0155, aber die beiden Extremwerte sind -0,102 und 0,088.

Bei normal-,  $t_{10}$ - und MGARCH(1)-verteilten Daten sind die asymptotische Standardabweichung  $\sigma_0(\hat{\theta})$  und die Mittelung von dessen Schätzwerten  $\bar{s}(\hat{\theta}_i)$  etwas niedriger als die beobachtete Standardabweichung der Schätzwerte  $s(t_i)$ . Schätzungen nach der verteilungsfreien Formel weichen dabei noch etwas stärker nach unten ab als jene nach der NV-Formel. Die Werte für  $t_5$ -verteilte Daten sind durch Ausreißer nach oben stark verzerrt. Da die aus den Parametern berechnete asymptotische Standardabweichung etwas kleiner als die empirische ist, liegen jeweils etwas weniger als 95 % der Schätzwerte im  $1,96\sigma$ -Intervall. Die Überdeckungshäufigkeiten der Konfidenzintervalle sind bei Anwendung der NV-Formel etwa bei 95,5%. Wie schon beim Sharpe-Ratio gesehen sind die Überdeckungshäufigkeiten leicht überhöht, obwohl die geschätzten Standardabweichungen im



Links: Histogramm von  $\widehat{Tr}$  mit angepasster Dichtefunktion der Normalverteilung  
 Rechts: Punktdiagramm für 2000 Zeitreihen mit  $\widehat{Tr}$  (Abszisse) und zugehörigem  $\widehat{\sigma}_{\widehat{Tr}}$  (Ordinate) für die verteilungsfreie Varianzformel( $\circ$ ) und die Varianzformel mit Normalverteilungsannahme( $\Delta$ ).

Abbildung 5.8: Grafiken zur Verteilung von  $\widehat{Tr}$  für i.i.d.- $t_{10}$ -verteilte Daten mit  $Tr = 0,004$  und Zeitreihenlänge  $n = 60$ .

Mittel etwas zu niedrig sind. Auch beim Treynor-Ratio zeigt sich, dass deutlich von Null abweichende Schätzwerte mit hohen Varianzschätzungen einhergehen. Im  $\widehat{Tr}-\widehat{\sigma}_{\widehat{Tr}}$ -Diagramm (Abb. 5.8b) erkennt man eine U-Form der Punktwolke, wenn man die äußeren, dünn besetzten Bereiche einbezieht.<sup>17</sup>

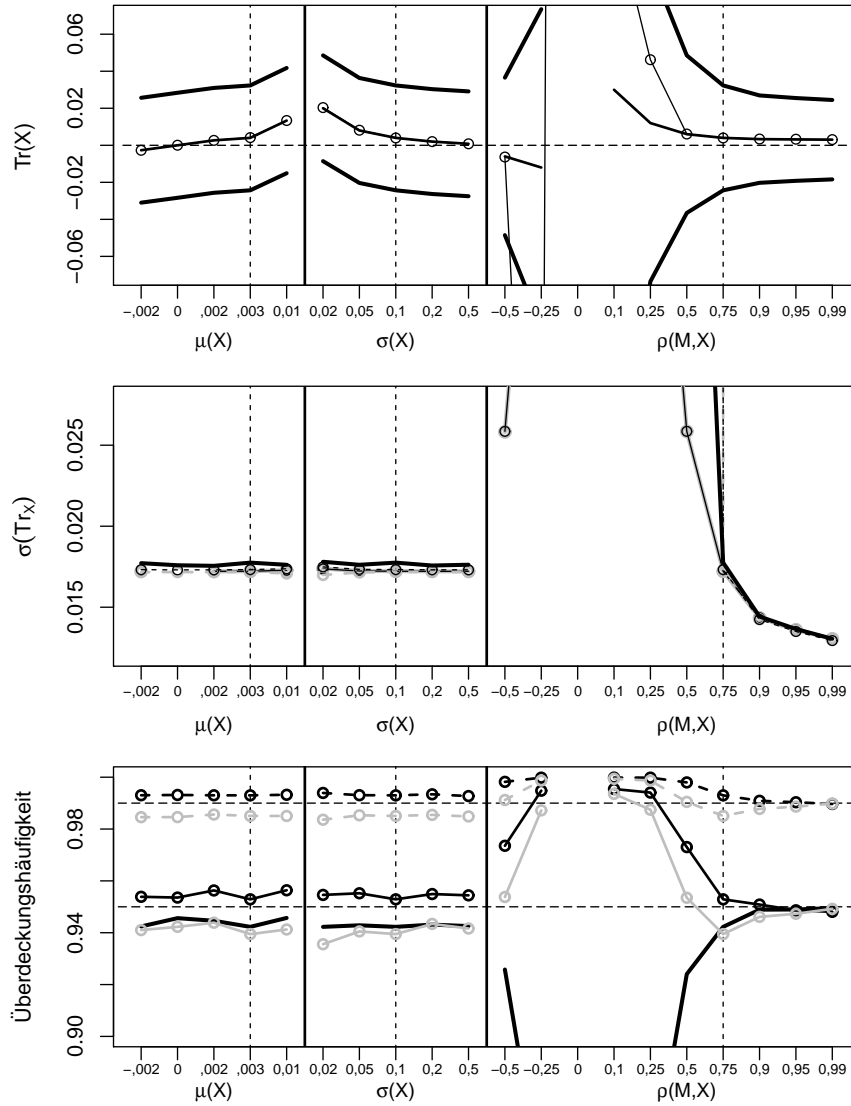
Klare Aussagen liefern die Grafiken zum Verhalten von  $\widehat{Tr}_X$  bei Variation von  $\mu_X$ ,  $\sigma_X$  und  $\rho_{M,X}$  (Abb. 5.9). Hinsichtlich der Variation von  $\mu_X$  und  $\sigma_X$  gibt es wenige Auffälligkeiten.  $Tr_X$  steigt mit  $\mu_X$  und fällt mit  $\sigma_X$ ; nennenswerte Verzerrungen des Schätzers  $\widehat{Tr}_X$  werden nicht beobachtet. Die Standardabweichung des Schätzers zeigt für verschiedene  $\mu_X$  und  $\sigma_X$  praktisch keine Veränderung und es gibt keine größeren Abweichungen zwischen theoretischen, geschätzten und empirischen Standardabweichungen. Kleinere Abweichungen werden auf der unteren Tafel erkennbar. Die Konfidenzintervalle auf Basis der i.i.d.-NV-Formel sind tendenziell etwas zu breit (Überdeckungshäufigkeiten im Durchschnitt ca. 95,4 % bzw. 99,3 %), jene nach verteilungsfreier Formel eher zu schmal (94 % bzw. 98,5 %).

<sup>17</sup> Anders als im Fall des Sharpe-Ratios (vgl. Abb. 5.2b) ziehen sich die Schätzwerte  $\widehat{\sigma}_{\widehat{Tr}}$  bei  $\widehat{Tr}_X \approx 0$  nicht in Richtung eines Punktes zusammen. Zwar taucht  $\mu_X^2$  im Zähler des Vorfaktors der Varianzformel auf, so dass man annehmen könnte, dass die Varianz für  $\widehat{Tr}_X \rightarrow 0$  gegen null geht, aber  $\mu_X^2$  steht auch im Zähler einer der Terme innerhalb der Klammer.

Deutlichere Unterschiede gibt es bei der Variation von  $\rho_{M,X}$ . Je näher  $\rho_{M,X}$  an der Obergrenze 1 liegt, desto besser sind die Anpassungsstatistiken. Für  $\rho_{M,X} = 0,95$  und  $\rho_{M,X} = 0,99$  führt weder der Anderson-Darling-Test noch der  $\chi^2$ -Test zur Ablehnung. Die Überdeckungshäufigkeiten liegen fast exakt bei 95 % bzw. 99 %. In Richtung  $\rho_{M,X} = 0$  nimmt die Qualität der Schätzungen stark ab. Bereits bei der Ausgangsparameterkombination sind die oben beschriebenen leichten Abweichungen erkennbar und bei  $\rho_{M,X} < 0,5$  sind die Ergebnisse kaum noch brauchbar. Die Verzerrung des Schätzers nimmt stark zu und die Varianzschätzung wird vollständig unzuverlässig. Für  $\rho_{M,X} = 0,5$  beispielsweise beträgt die aus den Parametern berechnete asymptotische Standardabweichung 0,026, die empirische Standardabweichung ist mehr als doppelt so hoch und die gemittelte geschätzte Standardabweichung liegt bei etwa 44. Geht der Parameter gegen null, ist der Schätzer nur noch in weniger als 85 % der Fälle im (stark verbreiterten)  $1,96\sigma$ -Intervall und die Schätzungen der Standardabweichung sind so hoch, dass die Konfidenzintervalle den wahren Wert fast immer einschließen. Der Grund hierfür ist, dass  $\rho_{M,X}$  einerseits ein Faktor des Nenners ist ( $\beta_{M,X} = \sigma_X/\sigma_M \cdot \rho_{M,X}$ ) und andererseits nur verhältnismäßig ungenau geschätzt werden kann.<sup>18</sup> Ist also die tatsächliche Korrelation von  $X$  und  $M$  nicht allzu hoch, kann es gut sein, dass der empirische Korrelationskoeffizient  $r_{M,X}$  einer bestimmten Stichprobe nahe null liegt. Dann ergibt sich ein betragsmäßig großer Schätzwert. Für  $\rho_{M,X} = 0$  schließlich ist  $Tr_X$  gar nicht definiert, da  $\rho_{M,X}$  ein Faktor des Nenners von  $Tr_X$  ist. Inferenzstatistisch brauchbar ist das Treynor-Ratio also nur dann, wenn das untersuchte Portfolio und das Marktportfolio stark korreliert sind. Bei  $n = 60$  sollte der Korrelationskoeffizient zwischen Portfolio und Markt auf jeden Fall höher als 0,5 sein. In der Praxis muss bei der Entscheidung, ob der Einsatz des Treynor-Ratios für eine bestimmte Portfolio-Markt-Konstellation geeignet ist, berücksichtigt werden, dass man seinerseits meist nur Schätzwerte für  $\rho_{M,X}$  zur Verfügung hat. Schon eine verhältnismäßig geringe Korrelation zwischen Portfolio und Markt kann Werte für  $r_{M,X}$  über 0,5 hervorbringen. Außerdem führen Renditeverteilungen mit starken Flanken (z.B.  $t_5$ -Verteilung, verschiedene GARCH-Prozesse) häufiger zu Ausreißern, die die in der Stichprobe beobachtete Korrelation in eine Richtung ziehen. Liegen solche Renditeverteilungen vor, sollte man sich noch stärker auf die Analyse hoch korrelierter Renditezeitreihen beschränken.

Zugunsten besserer Eigenschaften des Schätzers  $\hat{\beta}_{M,X}$  könnte man die Stichprobengröße dadurch erhöhen, dass man die Wertentwicklung in kürzeren Frequenzen misst, etwa indem man von Monats- auf Tagesrenditen wechselt. Für einen Beobachtungszeitraum von fünf Jahren würde man dann nicht auf Basis von 60 Monatsrenditen rechnen, sondern (bei ungefähr 250 Börsentagen/Jahr) auf Basis von etwa 1250 Tagesrenditen. Natürlich würde das daraus resultierende Ergebnis qualitativ verändert sein (Umrechnung der Ergebnisse erfordert Annahme unabhängig und identisch verteilter Tagesrenditen; andere Rendite-

<sup>18</sup>Bei  $n = 60$  und  $\alpha = 5\%$  ist z.B. der Annahmehereich eines Tests auf  $\rho = 0$  bei normalverteilten Daten das Intervall  $[-0,258; 0,258]$ , vgl. Sachs und Hedderich (2006, S. 544 ff.).



Verteilungsstatistiken bei Variation der Simulationsparameter  $\mu_X$ ,  $\sigma_X$  und  $\rho_{X,M}$ . Ausgangsparameterkombination:  $\mu_X = 0,003, \mu_M = 0,002, \sigma_X = \sigma_M = 0,1, \rho_{X,M} = 0,75$  (jeweils senkrechte gestrichelte Linien).

Daten: MGARCH(1)-Prozesse,  $n = 60$ , je 50 000 Simulationsläufe.

*oben:* Theoretischer Wert  $Tr_X$  und Mittelwert der Schätzungen (mit Datenpunkten) ; umliegend die theoretische  $1,96 - \sigma$ -Umgebung mit der i.i.d.-NV-Formel.

*Mitte:* Empirische Standardabweichung der Schätzer (durchgehende Linie, ohne Datenpunkte); asymptotische Standardabweichung (durchgehend, mit Datenpunkten) und mittlere geschätzte Standardabweichung (gestrichelt) mittels i.i.d.-NV-Formel (schwarz), und verteilungsfreier i.i.d.-Formel (grau).

*unten:* Anteil der Schätzwerte in  $1,96\sigma$ -Umgebung nach i.i.d. NV-Formel (Linie ohne Datenpunkte), Überdeckungshäufigkeit des 95%-KI (durchgehende Linie mit Datenpunkten) und des 99%-KI (gestrichelte Linie, m.D.) für die i.i.d.-NV-Formel (schwarz) und die verteilungsfreie i.i.d.-Formel (grau).

Abbildung 5.9: Grafiken zu  $\widehat{Tr}_X$  bei Parametervariation

verteilung; tägliche Korrelation ist nicht monatliche Korrelation; ggf. Probleme mit Wochentags- und Uhrzeiteffekten). Simulieren wir trotz dieser Einschränkungen einmal ein Szenario für Tagesdaten. Hierfür wird die Zeitreihenlänge um den Faktor  $d = \frac{250}{12} = 20\frac{5}{6}$  auf 1 250 erhöht, im Gegenzug werden die Parameter für Periodenrendite durch  $d$  und jene für die Periodenstandardabweichung durch  $\sqrt{d}$  geteilt. Hier eine Übersicht der Simulationsparameter:

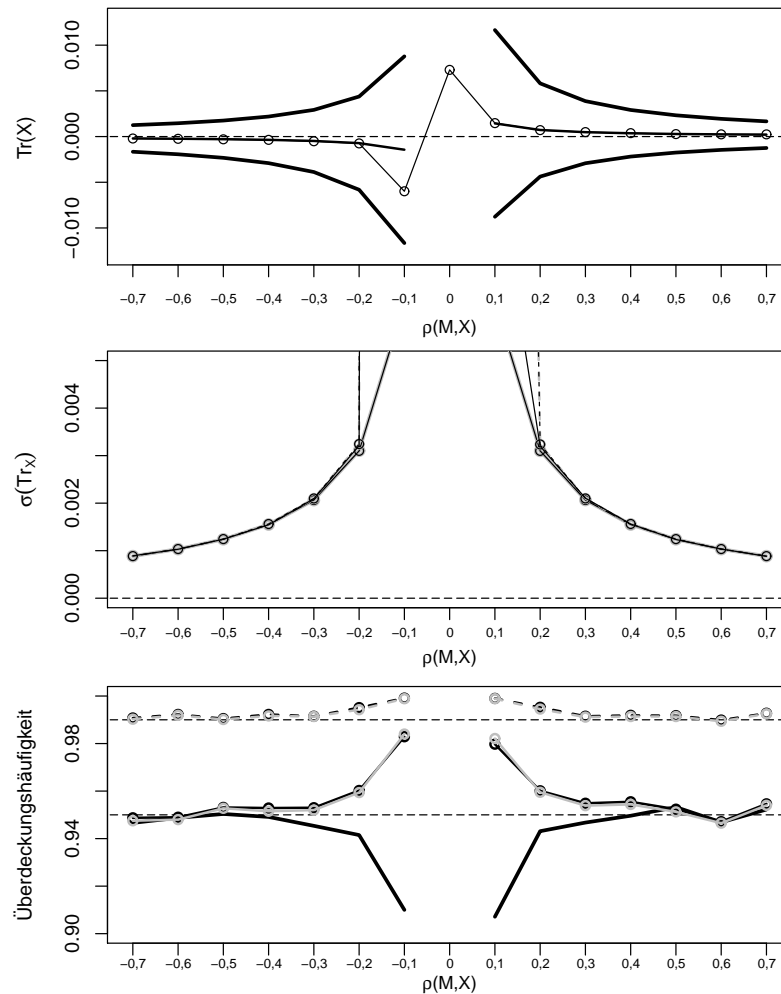
Variable	$\mu$	$\sigma$	Korrelationsmatrix	
$M$	0,000096	0,02191	1	$\rho$
$X$	0,000144	0,02191	$\rho$	1
$n = 1\,250, R = 8\,000, \text{MGARCH}(1)$				

Die Ergebnisse in Abb. 5.10 zeigen, dass durch die Erhöhung der Renditefrequenz brauchbare Schätzungen des Treynor-Ratios nun schon ab etwa  $|\rho_{M,X}| \geq 0,2$  erzielt werden können. Trotzdem bleibt es dabei, dass das Treynor-Ratio für seine zahlenmäßige Aussagekraft eine deutliche Korrelation zwischen Portfolio und Markt benötigt. Aus inferenzstatistischer Sicht erscheint das Treynor-Ratio also für die Bewertung von Portfolios geeignet, die schon anhand ihrer Anlagestrategie an einem bestimmten Marktportfolio ausgerichtet sind, nicht aber für solche – etwa bestimmte Hedgefonds –, die versuchen, ohne größere Exposure gegenüber dem Risiko des Gesamtmarktes Überrenditen zu generieren. Auffällig ist, dass der Schätzer weiterhin sehr stark streut. Die  $1,96\sigma$ -Intervalle liegen weiterhin fast symmetrisch um den Nullpunkt herum, obwohl die von den Parameterangaben implizierte jährliche Überrendite von ca. 3,6 % ökonomisch erheblich ist. Die Schätzgenauigkeit des Treynor-Ratios lässt sich also durch eine höhere Datenfrequenz und damit genauere Schätzung von  $\sigma_X$ ,  $\sigma_M$  und  $\rho_{M,X}$  nicht wesentlich erhöhen. Das Problem bleibt die hohe Streuung der Rendite bei einem relativ geringen Erwartungswert.

Die **Differenz  $\Delta_{Tr_{X,Y}}$  der Treynor-Ratios zweier Portfolios  $X$  und  $Y$ , bezogen auf dasselbe Marktportfolio  $M$** , untersuchen wir zunächst wieder anhand einer Ausgangsparameterkombination

Variable	$\mu$	$\sigma$	Korrelationsmatrix		
$M$	0,002	0,1	1	0,75	0,75
$X$	0,003	0,1	0,75	1	0,75
$Y$	0,002	0,1	0,75	0,75	1
$n = 60, R = 20\,000$					

und den vier Zeitreihenmodellen i.i.d.-normal, i.i.d.- $t_5$ , i.i.d.- $t_{10}$  und MGARCH(1) (mit  $\beta_1 = 0,8$  und  $\alpha_1 = 0,1$ ). In Tabelle 5.12 erkennt man Ergebnisse, die denen in Tab. 5.11 zu  $\widehat{Tr}_X$  sehr ähneln. Durch die positive Korrelation von  $X$  und  $Y$  (und somit auch  $\widehat{Tr}_X$  und  $\widehat{Tr}_Y$ ) ist die Standardabweichung von  $\widehat{\Delta}_{Tr_{X,Y}}$  etwas kleiner als jene von  $\widehat{Tr}_X$ , das oben mit denselben Parametern simuliert wurde. Außerdem ist die Überdeckungshäufigkeit des wahren Werts durch die mittels der verteilungsfreien Varianzformel berechneten Konfidenzintervalle nicht mehr



Verteilungsstatistiken bei Variation von  $\rho_{X,M}$ . Feste Parameter:  $\mu_X = 0,000144, \mu_M = 0,000096, \sigma_X = \sigma_M = 0,02191$

Daten: MGARCH(1)-Prozesse,  $n = 1250$ , je 8 000 Simulationsläufe.

*oben:* Theoretischer Wert  $Tr_X$  (zentrale Linien ohne Datenpunkte) und Mittelwert der Schätzungen (mit Datenpunkten); umliegend die theoretische  $1,96 - \sigma$ -Umgebung mit der i.i.d.-NV-Formel.

*Mitte:* Empirische Standardabweichung der Schätzer (durchgehende Linie, ohne Datenpunkte); asymptotische Standardabweichung (durchgehend, mit Datenpunkten) und mittlere geschätzte Standardabweichung (gestrichelt) mittels i.i.d.-NV-Formel (schwarz), und verteilungsfreier i.i.d.-Formel (grau).

*unten:* Anteil der Schätzwerte in  $1,96\sigma$ -Umgebung nach i.i.d. NV-Formel (Linie ohne Datenpunkte), Überdeckungshäufigkeit des 95%-KI (durchgehende Linie mit Datenpunkten) und des 99%-KI (gestrichelte Linie, m.D.) für die i.i.d.-NV-Formel (schwarz) und die verteilungsfreie i.i.d.-Formel (grau).

Abbildung 5.10: Grafiken zu  $\widehat{Tr}_X$  bei Variation des Korrelationskoeffizienten und  $n = 1250$



Verteilung	$\sigma_0(\hat{\theta})$	$\bar{s}(\hat{\theta}_i)$	$p_{\sigma^2=\sigma_0^2}$	$[u_{0,05}]$	$t_i$ in $u$	$KI_{95\%}$ um $\theta$
<b>NV</b>	$\theta = 0,00133$ ; $\bar{t} = 0,00136$ ; $s(t_i) = 0,01259$ ; $p_{AD} = 0,0016$ ; $\hat{\psi} = 0,0344$					
i.i.d., NV	0,01218	0,01254	$< 10^{-10}$	[-0,0225;0,0252]	94,06%	95,05%
i.i.d.	0,01218	0,01241	$< 10^{-10}$	[-0,0225;0,0252]	94,06%	94,76%
<b>t<sub>10</sub></b>	$\theta = 0,00133$ ; $\bar{t} = 0,00128$ ; $s(t_i) = 0,01267$ ; $p_{AD} < 10^{-4}$ ; $\hat{\psi} = 0,0359$					
i.i.d., NV	0,01218	0,01262	$< 10^{-14}$	[-0,0225;0,0252]	94,06%	95,46%
i.i.d.	0,01218*	0,01252	$< 10^{-14}$	[-0,0225;0,0252]	94,08%	95,23%
<b>t<sub>5</sub></b>	$\theta = 0,00133$ ; $\bar{t} = 0,00155$ ; $s(t_i) = 0,01339$ ; $p_{AD} < 10^{-99}$ ; $\hat{\psi} = 0,0481$					
i.i.d., NV	0,01218	0,01338	$< 10^{-99}$	[-0,0225;0,0252]	93,39%	95,46%
i.i.d.	0,01219*	0,01339	$< 10^{-99}$	[-0,0226;0,0252]	93,42%	95,52%
<b>MGARCH</b>	$\theta = 0,00133$ ; $\bar{t} = 0,00134$ ; $s(t_i) = 0,01262$ ; $p_{AD} < 10^{-4}$ ; $\hat{\psi} = 0,0353$					
i.i.d., NV	0,01218	0,01256	$< 10^{-12}$	[-0,0225;0,0252]	93,89%	95,23%
i.i.d.	0,01219*	0,01243	$< 10^{-12}$	[-0,0226;0,0252]	93,95%	95,01%
* aus den Daten simulierte Werte						
Erläuterung der Kürzel für die Statistiken in Tab. 5.2 auf S. 114.						

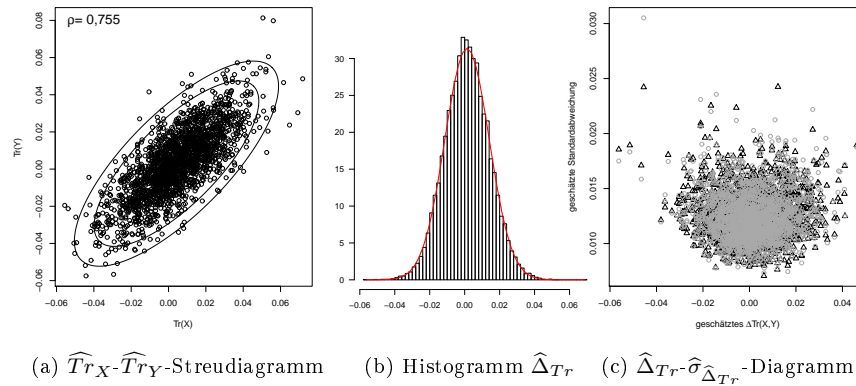
Tabelle 5.12: Ergebnisse für die Differenz der Treynor-Ratios zweier Portfolios bei Zeitreihenlänge  $n = 60$ , Ausgangsparameter

wie oben deutlich unter den vorgegebenen 95 %. Die Abweichungen betragen jetzt nur noch bis zu einem halben Prozentpunkt. In den Grafiken in Abb. 5.11 sieht man, dass  $\hat{\Delta}_{Tr_{X,Y}}$  fast exakt normalverteilt ist (Abb. 5.11b) und dass auch die gemeinsame Verteilung von  $\hat{Tr}_X$  und  $\hat{Tr}_Y$  der Hypothese einer gemeinsamen Normalverteilung nicht widerspricht (Abb. 5.11a). Die gemessene Korrelation zwischen  $\hat{Tr}_X$  und  $\hat{Tr}_Y$  von 0,755 entspricht nahezu der Korrelation von  $X$  und  $Y$  ( $\rho_{X,Y} = 0,75$ ). In Abb. 5.11c erkennt man bestenfalls eine schwache positive Abhängigkeit zwischen  $|\hat{\Delta}_{Tr_{X,Y}}|$  und  $\hat{\sigma}_{\hat{\Delta}_{Tr}}$ .

Da es bei der Simulation der Differenz zweier Treynor-Ratios zusätzlich zu Verteilungsmodell und Stichprobenlänge neun variierbare Parameter gibt, soll hier aus Platzgründen darauf verzichtet werden, die Simulationsergebnisse bei Parametervariationen systematisch darzustellen. Statt dessen eine Übersicht über einige wichtige Punkte:

- Wenn mindestens eines der Portfolios  $X$  und  $Y$  eine geringe Korrelation zum Marktportfolio  $M$  (kleiner als 0,5) aufweist, ist die Qualität der Schätzergebnisse – kaum überraschend – ähnlich jener, die wir weiter oben beim Treynor-Ratio eines Portfolios mit geringer Korrelation zum Marktportfolio gesehen hatten. Extreme Ausreißer sind häufig und die Anpassung der Verteilung des Schätzers an die Normalverteilung ist schlecht. Das gilt auch dann, wenn die Korrelation zwischen  $X$  und  $Y$  hoch und jene der beiden zum Markt niedrig ist (z.B.  $\rho_{XY} = 0,99$ ,  $\rho_{MX} = \rho_{MY} = 0,5$ ).
- Wenn die Korrelationen zwischen den drei Portfolios bei 0,9 und höher liegen, sind alle Prüfgrößen (AD-Test,  $\chi^2$ -Varianztest, Überdeckungshäufigkeiten) unauffällig, unabhängig von den anderen Verteilungsparametern.<sup>19</sup>

<sup>19</sup>Voraussetzungen: Sonstige Parameter in ökonomisch sinnvollem Bereich; vierte Momente der Renditeverteilung sind endlich.



links: Streudiagramm  $\widehat{Tr}_X - \widehat{Tr}_Y$  mit angepassten 95%- und 99%-NV-Zentralbereichen für 2 000 Zeitreihen

Mitte: Histogramm von  $\widehat{\Delta}_{Tr}$  mit angepasster Normalverteilung

rechts:  $\widehat{\Delta}_{Tr}$  (Abszisse) gegen  $\widehat{\sigma}_{\widehat{\Delta}_{Tr}}$  (Ordinate) für 2 000 Zeitreihen. Varianzschätzung nach Formel für normalverteilte Renditen ( $\Delta$ ) bzw. ohne Verteilungsannahme ( $\circ$ ).

Abbildung 5.11: Grafiken zur Verteilung von  $\widehat{\Delta}_{Tr}$ , i.i.d.- $t_{10}$ -verteilte Renditen, Zeitreihenlänge  $n = 60$ .

- Die asymptotischen  $1,96\sigma$ -Intervalle beinhalteten in fast allen simulierten Parameterkombinationen die Null. Ausnahmen gab es nur, wenn der Renditeunterschied zwischen den beiden Portfolios sehr hoch war (z.B. 1 Prozentpunkt pro Monat) bei gleichzeitiger starker Korrelation ( $\rho_{XY} \geq 0,95$ ).
- Die Mittelwerte  $\mu_M, \mu_X, \mu_Y$  und die Standardabweichungen  $\sigma_M, \sigma_X, \sigma_Y$  haben nur geringen Einfluss auf die statistischen Eigenschaften des Schätzers  $\widehat{\Delta}_{Tr}$ .

### 5.3.3 Jensens Alpha

Die Varianz des Schätzers von Jensens Alpha für ein Portfolio  $X$  auf Basis eines Marktportfolios  $M$  ist unter Annahme normalverteilter Renditen nichts anderes als die Varianz des Steigungsparameters einer linearen Einfachregression mit normalverteilten Residuen (siehe Abschnitt 4.1.2.3, S. 80). Neu hinzu kommt dessen oben vorgestellte Verallgemeinerung für beliebige Verteilungen. In Tab. 5.13 stehen die Ergebnisse für die Ausgangsparameterkombination ( $\mu_X = 0,003$ ,  $\mu_M = 0,002$ ,  $\sigma_X = \sigma_M = 0,1$ ,  $\rho_{M,X} = 0,75$ ,  $n = 60$ ,  $R = 50\,000$ , vgl. S. 134).

Die Anpassung der empirischen Verteilung der Schätzwerte an eine Normalverteilung ist ausweislich der AD-Tests und des Histogramms in Abb. 5.12a sehr

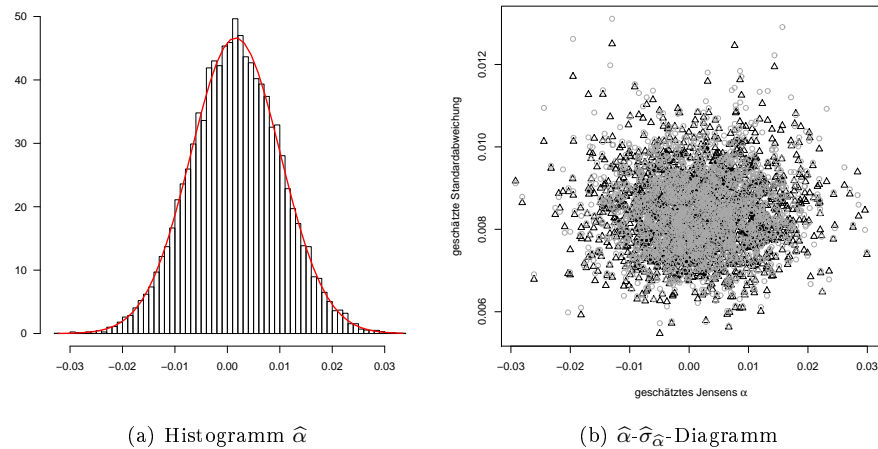
Verteilung	$\sigma_0(\hat{\theta})$	$\bar{s}(\hat{\theta}_i)$	$p_{\sigma^2=\sigma_0^2}$	$[u_{0,05}]$	$t_i$ in $u$	$KI_{95\%}$ um $\theta$
<b>NV</b>	$\theta = 0,0015$ ; $\bar{t} = 0,001533$ ; $s(t_i) = 0,00864$ ; $p_{AD} = 0,2483$ ; $\hat{\psi} = 0,0232$					
i.i.d., NV	0,00854	0,00847	0,0004	[-0,0152;0,0182]	94,66%	94,51%
i.i.d.	0,00854	0,00839	0,0004	[-0,0152;0,0182]	94,66%	94,22%
<b>t<sub>10</sub></b>	$\theta = 0,0015$ ; $\bar{t} = 0,001599$ ; $s(t_i) = 0,00860$ ; $p_{AD} = 0,3566$ ; $\hat{\psi} = 0,0288$					
i.i.d., NV	0,00854	0,00845	0,0213	[-0,0152;0,0182]	94,82%	94,53%
i.i.d.	0,00854*	0,00836	0,0304	[-0,0152;0,0182]	94,84%	94,30%
<b>t<sub>5</sub></b>	$\theta = 0,0015$ ; $\bar{t} = 0,001536$ ; $s(t_i) = 0,00856$ ; $p_{AD} = 0,3609$ ; $\hat{\psi} = 0,0269$					
i.i.d., NV	0,00854	0,00840	0,5143	[-0,0152;0,0182]	95,01%	94,62%
i.i.d.	0,00856*	0,00827	0,9176	[-0,0153;0,0183]	95,03%	94,25%
<b>MGARCH</b>	$\theta = 0,0015$ ; $\bar{t} = 0,001548$ ; $s(t_i) = 0,00862$ ; $p_{AD} = 0,0877$ ; $\hat{\psi} = 0,0272$					
i.i.d., NV	0,00854	0,00846	0,0042	[-0,0152;0,0182]	94,71%	94,67%
i.i.d.	0,00854*	0,00839	0,0060	[-0,0152;0,0182]	94,72%	94,44%
* aus den Daten simulierte Werte						
Erläuterung der Kürzel für die Statistiken in Tab. 5.2 auf S. 114.						

Tabelle 5.13: Ergebnisse für Jensens Alpha bei Zeitreihenlänge  $n = 60$ , Ausgangsparameter

gut, auch bei den Verteilungsmodellen mit stärkeren Flanken.<sup>20</sup> Auch asymptotische und empirische Standardabweichung weisen nur geringe Unterschiede auf. Beide Varianzformeln liefern gute Ergebnisse, wobei die Formel mit NV-Annahme in Hinblick auf die Überdeckungshäufigkeiten der 95%-Konfidenzintervalle etwas besser abschneidet. In Abb. 5.12b erkennt man keine Abhängigkeit zwischen Schätzer und geschätzter Standardabweichung. Auch sind keine nennenswerten Unterschiede in den statistischen Eigenschaften zwischen Varianzschätzungen auf Basis der NV-Formel und der verteilungsfreien Formel erkennbar. Insbesondere gibt es keine Ausreißer in den geschätzten Standardabweichungen. Im Vergleich zum Treynor-Ratio (Abb. 5.8b) und zum Sharpe-Ratio (Abb. 5.3b) ist die Schwankungsbreite der Schätzwerte für die Standardabweichung des Performancemaßschätzers verhältnismäßig hoch.

Die Ergebnisse bei Parametervariation sind Abb. 5.13 zu entnehmen, wiederum für je 50 000 MGARCH(1)-Zeitreihen der Länge  $n = 60$ . In der oberen Tafel ist keine Verzerrung des Schätzers auszumachen, und wenn, wäre sie kaum relevant im Vergleich zu seiner großen Streuung, dargestellt durch die theoretischen  $1,96\sigma$ -Bereiche. In der mittleren Tafel ist bei allen drei Parametern jeweils praktisch nur eine einzige Linie zu sehen. Die empirische Standardabweichung der Schätzer, die asymptotische Standardabweichung  $\sigma_0(\hat{\theta})$  und das Mittel der geschätzten Standardabweichungen  $\bar{s}(\hat{\theta}_i)$  mit und ohne Normalverteilungsannahme unterscheiden sich nicht nennenswert. Die in der unteren Tafel dargestellten Überdeckungshäufigkeiten des wahren Wertes durch verschiede-

<sup>20</sup>Da die Mittel der Schätzwerte aller vier Simulationsreihen über dem theoretischen Wert liegen, würde man eine Verzerrung des Schätzers in Erwägung ziehen, die theoretisch nicht vorkommen dürfte. Der KQ-Schätzer für den Achsenabschnitt einer linearen Einfachregression ist nämlich unverzerrt, siehe z.B. Auer (2007, S. 83 f.). Es handelt sich hier offenbar um Zufall, denn die Abweichungen sind nicht signifikant und in Kontroll-Simulationsreihen traten Mittelwerte unter 0,0015 auf.



Links: Histogramm von  $\hat{\alpha}$  mit angepasster Dichtefunktion der Normalverteilung

Rechts: Punktdiagramm für 2000 Zeitreihen mit  $\hat{\alpha}$  (Abszisse) und zugehörigem  $\hat{\sigma}_{\hat{\alpha}}$  (Ordinate) für die verteilungsfreie Varianzformel( $\circ$ ) und die Varianzformel mit Normalverteilungsannahme( $\Delta$ ).

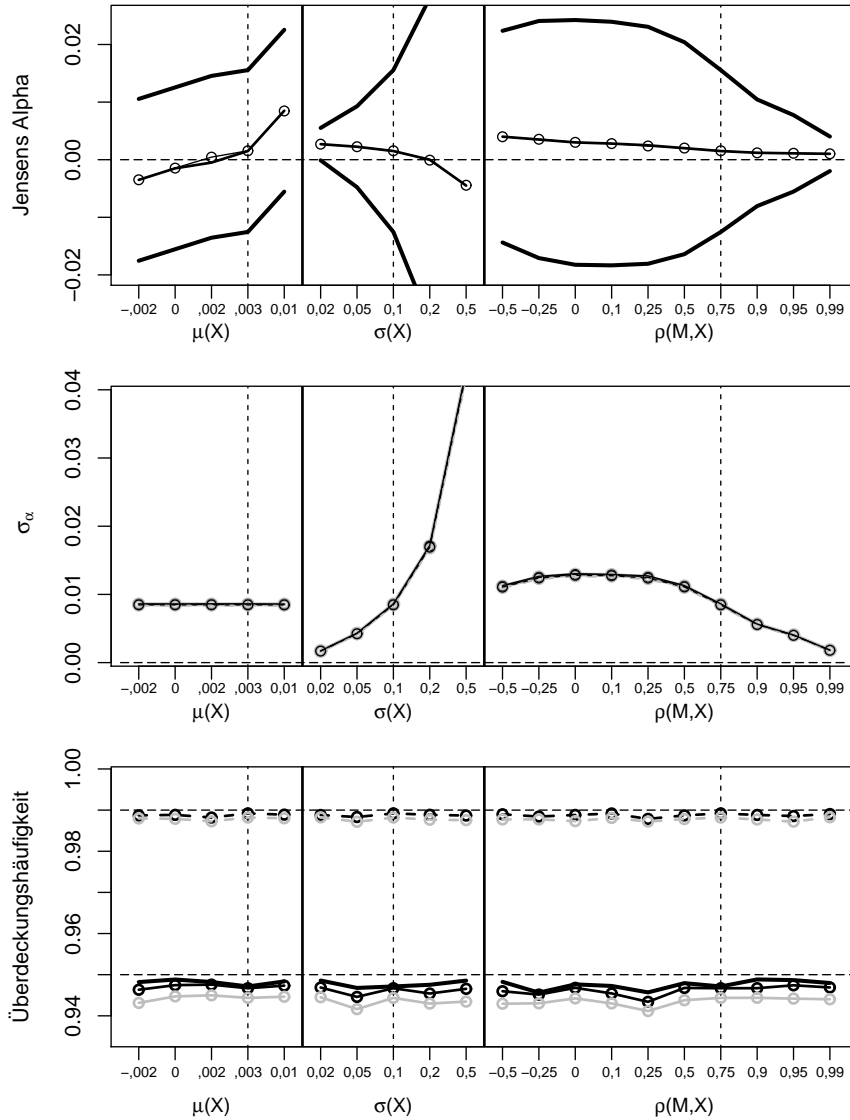
Abbildung 5.12: Grafiken zur Verteilung von Jensen's  $\hat{\alpha}$  für i.i.d.- $t_{10}$ -verteilte Daten mit  $\alpha = 0,0015$  und Zeitreihenlänge  $n = 60$ .

ne Konfidenzintervalle hängen nicht erkennbar von der Parameterkonstellation ab. Stets wird das vorgegebene Niveau leicht unterschritten (bis max. ein Prozentpunkt). Konfidenzintervalle auf Basis der verteilungsfreien Varianzformel weichen stets etwas stärker nach unten ab.

Insgesamt erweist sich die Varianzformel mit Normalverteilungsannahme in den Simulationen als zuverlässiger als die verteilungsfreie Formel. Es kann natürlich sein, dass sich diese Aussage bei längeren Zeitreihen und stark schiefen Renditeverteilungen umkehrt. Anders als das Treynor-Ratio kann man Jensen's Alpha ohne Rücksicht auf die Verteilungsparameter inferenzstatistisch valide schätzen (aber zumindest bei  $n = 60$  keineswegs genau).<sup>21</sup>

Zur Differenz  $\Delta_{\alpha_X, Y}$  der Alphas zweier Portfolios  $X$  und  $Y$ , bezogen auf dasselbe Marktportfolio  $M$  stehen in Tabelle 5.14 die Ergebnisse für die Ausgangsparameterkombination bei verschiedenen Verteilungsmodellen. Ähnlich wie schon bei Jensen's Alpha für ein einzelnes Portfolio ist die empirische Verteilung der Schätzwerte per Anderson-Darling-Test nicht von einer Normalverteilung unterscheidbar. Theoretische und empirische Standardabweichungen passen nach den Ergebnissen der  $\chi^2$ -Tests sehr gut zusammen. Lediglich

<sup>21</sup>Bei Simulationen mit sehr extremen Parameterkonstellationen erwies sich die verteilungsfreie Formel als weniger robust als die NV-Formel. In solchen Fällen waren die Konfidenzintervalle auf Basis der NV-Formel tendenziell zu liberal, diejenigen auf Basis der verteilungsfreien Formel zu konservativ.



Verteilungsstatistiken bei Variation der Simulationsparameter  $\mu_X$ ,  $\sigma_X$  und  $\rho_{X,M}$ . Ausgangskombination:  $\mu_X = 0,003$ ,  $\mu_M = 0,002$ ,  $\sigma_X = \sigma_M = 0,1$ ,  $\rho_{X,M} = 0,75$  (jeweils senkrechte gestrichelte Linien).

Daten: MGARCH(1)-Prozesse,  $n = 60$ , je 50 000 Simulationsläufe.

*oben:* Theoretischer Wert  $\alpha_X$  und Mittelwert der Schätzungen (mit Datenpunkten); umliegend die theoretische  $1,96 - \sigma$ -Umgebung mit der i.i.d.-NV-Formel.

*Mitte:* Empirische Standardabweichung der Schätzer (durchgehende Linie, ohne Datenpunkte); asymptotische Standardabweichung (durchgehend, mit Datenpunkten) und mittlere geschätzte Standardabweichung (gestrichelt) mittels i.i.d.-NV-Formel (schwarz), und verteilungsfreier i.i.d.-Formel (grau).

*unten:* Anteil der Schätzwerte in  $1,96\sigma$ -Umgebung nach i.i.d. NV-Formel (Linie ohne Datenpunkte), Überdeckungshäufigkeit des 95%-KI (durchgehende Linie mit Datenpunkten) und des 99%-KI (gestrichelte Linie, m.D.) für die i.i.d.-NV-Formel (schwarz) und die verteilungsfreie i.i.d.-Formel (grau).

Abbildung 5.13: Grafiken zu  $\hat{\alpha}_X$  bei Parametervariation

Verteilung	$\sigma_0(\hat{\theta})$	$\bar{s}(\hat{\theta}_i)$	$p_{\sigma^2=\sigma_0^2}$	$[u_{0,05}]$	$t_i$ in $u$	$KI_{95\%}$ um $\theta$
<b>NV</b>	$\theta = 0,001$ ; $\bar{t} = 0,000995$ ; $s(t_i) = 0,00922$ ; $p_{AD} = 0,869$ ; $\hat{\psi} = 0,0316$					
i.i.d., NV	0,00913	0,00907	0,053	[-0,0169;0,0189]	94,66%	94,49%
i.i.d.	0,00913	0,00899	0,053	[-0,0169;0,0189]	94,66%	94,26%
<b>t<sub>10</sub></b>	$\theta = 0,001$ ; $\bar{t} = 0,000935$ ; $s(t_i) = 0,00915$ ; $p_{AD} = 0,841$ ; $\hat{\psi} = 0,0365$					
i.i.d., NV	0,00913	0,00902	0,663	[-0,0169;0,0189]	94,94%	94,81%
i.i.d.	0,00914*	0,00892	0,849	[-0,0169;0,0189]	94,97%	94,49%
<b>t<sub>5</sub></b>	$\theta = 0,001$ ; $\bar{t} = 0,001134$ ; $s(t_i) = 0,00913$ ; $p_{AD} = 0,247$ ; $\hat{\psi} = 0,0333$					
i.i.d., NV	0,00913	0,00898	0,910	[-0,0169;0,0189]	95,07%	94,80%
i.i.d.	0,00915*	0,00885	0,641	[-0,0169;0,0189]	95,11%	94,51%
<b>MGARCH</b>	$\theta = 0,001$ ; $\bar{t} = 0,001013$ ; $s(t_i) = 0,00922$ ; $p_{AD} = 0,113$ ; $\hat{\psi} = 0,0353$					
i.i.d., NV	0,00913	0,00905	0,050	[-0,0169;0,0189]	94,79%	94,61%
i.i.d.	0,00913	0,00897	0,040	[-0,0169;0,0189]	94,77%	94,40%
* aus den Daten simulierte Werte						
Erläuterung der Kürzel für die Statistiken in Tab. 5.2 auf S. 114.						

Tabelle 5.14: Ergebnisse für die Differenz der Jensens Alphas zweier Portfolios bei Zeitreihenlänge  $n = 60$ , Ausgangsparameter

die Überdeckungshäufigkeiten der Konfidenzintervalle liegen stets etwas unter 95 %, wobei die Abweichung bei der verteilungsfreien Formel etwas größer ist.

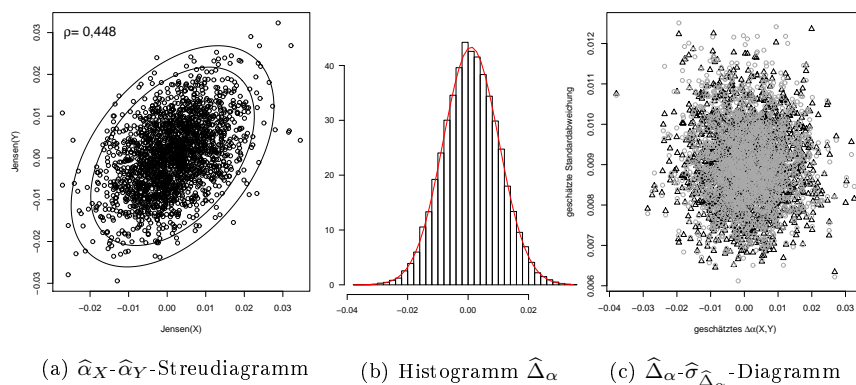
In Abb. 5.14 erkennt man in der linken Grafik, dass  $\hat{\alpha}_X$  und  $\hat{\alpha}_Y$  eine regelmäßige Punktwolke bilden, ohne dass von einer bivariaten Normalverteilung abweichende Strukturen zu finden wären. Die beiden Schätzer sind positiv korreliert, jedoch – anders als bei  $\Delta_{Tr}$  – deutlich schwächer als  $X$  und  $Y$  (hier:  $\rho_{\hat{\alpha}_X, \hat{\alpha}_Y} \approx 0,45$  bei  $\rho_{X,Y} = 0,75$ ). Die mittlere Tafel verdeutlicht noch einmal, dass der Schätzer  $\hat{\Delta}_\alpha$  bei i.i.d.- $t_{10}$ -verteilten Werten schon bei  $n = 60$  geradezu perfekt normalverteilt ist. Die rechte Tafel, die eine kreisrunde, gleichmäßige Punktwolke zeigt, lässt auch keinen Zusammenhang zwischen den Schätzwerten und den zugehörigen Schätzungen der Standardabweichung erkennen.

Zum Abschluss wieder einige Hinweise zum Verhalten des Schätzers bei Veränderung der Ausgangsparameter  $\mu_X, \mu_Y, \mu_M, \sigma_X, \sigma_Y, \sigma_M, \rho_{M,X}, \rho_{M,Y}$  und  $\rho_{X,Y}$  (i.i.d.- $t_{10}$ -Zeitreihen,  $n = 60, R = 20\,000$ ):

- Normalverteilung der Schätzer: Die Normalverteilungshypothese konnte mittels des Anderson-Darling-Test – mit einer Ausnahme<sup>22</sup> – bei keiner Parameterkombination zum 0,01 %-Niveau abgelehnt werden.<sup>23</sup>

<sup>22</sup>Die Ausnahme trat auf, wenn eines der beiden Wertpapiere sehr viel stärker mit dem Markt korreliert ist als das Wertpapier und der Markt mit dem alternativen Wertpapier.  $\hat{\psi}$  signalisiert hier jedoch, dass die Abweichung zwischen empirischer Verteilungsfunktion und angepasster Normalverteilung immer noch relativ unbedeutend ist.

<sup>23</sup>Da hier Simulationsdaten zu etwa 150 verschiedenen Parameterkombinationen unabhängig voneinander auf Signifikanz getestet wurden (Data Mining), sollte ein deutlich strengeres nominelles Testniveau zur Entscheidung über die Nullhypothese verwendet werden. Würde man Tests zum Niveau 5% oder 1% verwenden, wäre eine erhebliche Zahl falsch positiver Resultate (Fehler 1. Art) zu erwarten.



links: Streudiagramm  $\hat{\alpha}_X$ - $\hat{\alpha}_Y$  mit angepassten 95%- und 99%-NV-Zentralbereichen für 2000 Zeitreihen

Mitte: Histogramm von  $\hat{\Delta}_\alpha$  mit angepasster Normalverteilung

rechts:  $\hat{\Delta}_\alpha$  (Abszisse) gegen  $\hat{\sigma}_{\hat{\Delta}_\alpha}$  (Ordinate) für 2000 Zeitreihen. Varianzschätzung nach Formel für normalverteilte Renditen ( $\Delta$ ) bzw. ohne Verteilungsannahme ( $\circ$ ).

Abbildung 5.14: Grafiken zur Verteilung von  $\hat{\Delta}_\alpha$ , i.i.d.- $t_{10}$ -verteilte Renditen, Zeitreihenlänge  $n = 60$ .

- Die Überdeckungshäufigkeiten der 95%- und 99%- Konfidenzintervalle lagen stets ausreichend nahe an den theoretischen Niveaus (meist leicht darunter), außer wenn die Standardabweichung eines der beiden Wertpapier weit über der des anderen Wertpapiers und des Marktportfolios liegt. In diesem Fall steigen die Überdeckungshäufigkeiten in Richtung eins an.
- Während das Mittel der aus den Daten geschätzten Standardabweichungen stets nahe der empirischen Standardabweichung liegt, gibt es oft große Unterschiede zwischen diesen beiden Werten und der aus den Parametern berechneten asymptotischen Standardabweichung. Die aus den Parametern berechnete asymptotische Varianz ist zu klein, wenn das Marktportfolio mit einem oder beiden Portfolios deutlich stärker korreliert ist als die beiden Portfolios untereinander. Im Gegenzug ist die asymptotische Varianz zu groß, sobald die Korrelation zwischen den beiden Portfolios  $X$  und  $Y$  deutlich größer ist als die Korrelationen zwischen den Portfolios und dem Marktportfolio. Wenn man also Varianzschätzungen betreiben möchte, sollte man ein Marktportfolio wählen, das mit den beiden Portfolios etwa ähnlich stark korreliert ist wie die Portfolios untereinander.

Jensens Alpha ist vom inferenzstatistischen Standpunkt her wesentlich problematischer als das Treynor-Ratio. Bei allen untersuchten Parameterkombinationen erwiesen sich die Schätzungen der Standardabweichungen und Konfidenz-

Verteilung	$\sigma_0(\hat{\theta})$	$\bar{s}(\hat{\theta}_i)$	$p_{\sigma^2=\sigma_0^2}$	$[u_{0,05}]$	$t_i$ in $u$	$KI_{95\%}$ um $\theta$
<b>NV</b>	$\theta = 0,003$ ; $\bar{t} = 0,00307$ ; $s(t_i) = 0,01300$ ; $p_{AD} < 10^{-4}$ ; $\hat{\psi} = 0,0284$					
i.i.d., NV	0,01291	0,01285	0,0363	[-0,0223;0,0283]	94,70%	95,15%
i.i.d.	0,01291	0,01345	0,0596	[-0,0223;0,0283]	94,70%	98,01%
<b>t<sub>10</sub></b>	$\theta = 0,003$ ; $\bar{t} = 0,00307$ ; $s(t_i) = 0,01303$ ; $p_{AD} < 10^{-4}$ ; $\hat{\psi} = 0,0285$					
i.i.d., NV	0,01291	0,01285	0,0026	[-0,0223;0,0283]	94,66%	95,09%
i.i.d.	0,01292*	0,01368	0,0047	[-0,0223;0,0283]	94,68%	98,60%
<b>t<sub>5</sub></b>	$\theta = 0,003$ ; $\bar{t} = 0,00306$ ; $s(t_i) = 0,01306$ ; $p_{AD} < 10^{-6}$ ; $\hat{\psi} = 0,0278$					
i.i.d., NV	0,01291	0,01286	0,0005	[-0,0223;0,0283]	94,64%	95,23%
i.i.d.	0,01295*	0,01421	0,0103	[-0,0224;0,0284]	94,72%	99,12%
<b>MGARCH</b>	$\theta = 0,003$ ; $\bar{t} = 0,00303$ ; $s(t_i) = 0,01301$ ; $p_{AD} < 10^{-7}$ ; $\hat{\psi} = 0,0284$					
i.i.d., NV	0,01291	0,01285	0,0204	[-0,0223;0,0283]	94,68%	95,20%
i.i.d.	0,01291*	0,01350	0,0207	[-0,0223;0,0283]	94,68%	98,23%
* aus den Daten simulierte Werte						
Erläuterung der Kürzel für die Statistiken in Tab. 5.2 auf S. 114.						

Tabelle 5.15: Ergebnisse für die Originalversion von  $MM$  bei Zeitreihenlänge  $n = 60$ , Ausgangsparameter

intervalle als einigermaßen bis sehr zuverlässig – abgesehen von dem Hinweis am Ende des letzten Absatzes.

### 5.3.4 Modigliani-Modigliani

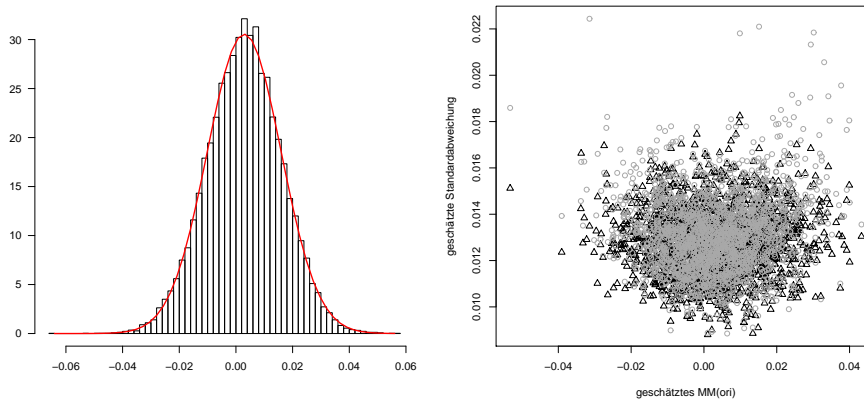
Die Simulationsergebnisse der beiden Versionen des Performancemaßes nach Modigliani/Modigliani<sup>24</sup> ( $MM_{ori}$  und  $MM_{alt}$ ) werden in diesem Abschnitt nacheinander vorgestellt. Die Differenz des Performancemaßes für zwei Portfolios  $\Delta_{MM}$  ist für beide Versionen identisch und folgt im Anschluss.

Die Ergebnisse für  $MM_{ori}$  mit den Ausgangsparametern (siehe S. 134) sind in Tabelle 5.15 zu finden. Für alle Verteilungsmodelle lassen sich folgende Beobachtungen treffen: Man kann eine positive Verzerrung des Schätzers feststellen, die aber vernachlässigbar erscheint. Ebenfalls erkennbar, aber sehr gering ist eine leichte Unterschätzung der tatsächlichen Varianz durch die asymptotische Varianz sowie durch die geschätzte Varianz bei Verwendung der Formel mit Normalverteilungsannahme. Die Varianz nach der Formel ohne NV-Annahme ist hingegen tendenziell zu groß, was darin zum Ausdruck kommt, dass die 95%-Konfidenzintervalle den wahren Wert deutlich zu häufig einschließen (in etwa 98 - 99 % der Fälle.). Der Schätzer  $\widehat{MM}_{X,ori}$  ist nahezu normalverteilt (siehe auch Abb. 5.15a). In Abb. 5.15b kann man bei ansonsten unauffälligem Befund erkennen, dass die Standardabweichungen nach der Formel ohne NV-Annahme gelegentlich – vor allem bei betragsmäßig großem  $MM_{ori}$  – leicht nach oben ausreißen.

In Abb. 5.16 (Ergebnisse für Parametervariationen) erkennt man in der oberen Tafel, dass für keine der geprüften Parametervariationen eine nennenswerte

<sup>24</sup>siehe Abschnitt 3.2.3.1



(a) Histogramm  $\widehat{MM}_{ori}$ (b)  $\widehat{MM}_{ori}$ - $\widehat{\sigma}_{\widehat{MM}_{ori}}$ -Diagramm

Links: Histogramm von  $\widehat{MM}_{ori}$  mit angepasster Dichtefunktion der Normalverteilung

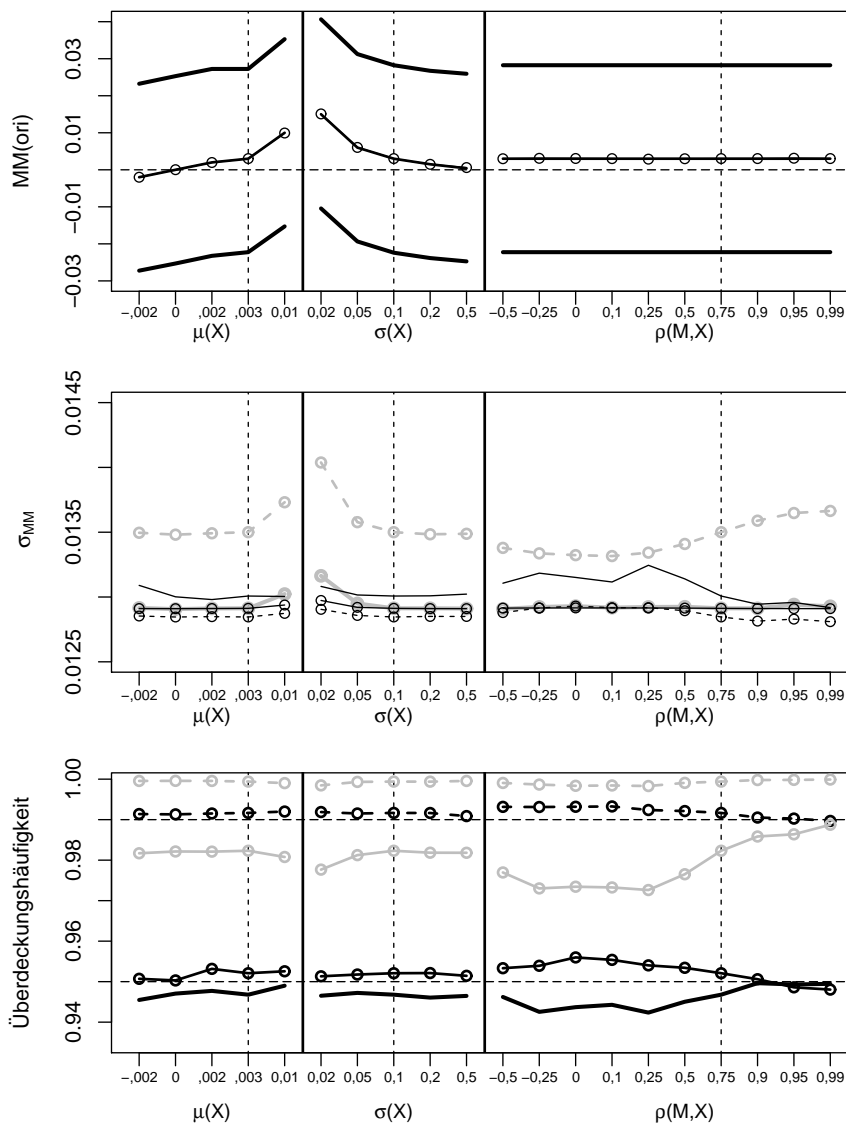
Rechts: Punktdiagramm für 2000 Zeitreihen mit  $\widehat{MM}_{ori}$  (Abszisse) und zugehörigem  $\widehat{\sigma}_{\widehat{MM}_{ori}}$  (Ordinate) für die verteilungsfreie Varianzformel( $\circ$ ) und die Varianzformel mit Normalverteilungsannahme( $\triangle$ ).

Abbildung 5.15: Grafiken zur Verteilung von  $\widehat{MM}_{ori}$  für i.i.d.- $t_{10}$ -verteilte Daten mit Zeitreihenlänge  $n = 60$ .

Verzerrung des Schätzers auffällt. Die 95%-Bereiche sind erneut sehr breit und kommen an keiner Stelle in die Nähe der Nulllinie. Die Standardabweichung wird im ökonomisch sinnvollen Bereich kaum durch die hier variierten Parameter beeinflusst<sup>25</sup> (mittlere Tafel, man beachte die feine Skala). Die Schätzungen der Standardabweichung über die NV-Formel liegen im Mittel sehr nahe an der asymptotischen Standardabweichung und minimal unter der empirischen Standardabweichung. Die Schätzungen über die Formel ohne NV-Annahme sind im Mittel etwas erhöht (siehe die Ausreißer in Abb. 5.15b). Entsprechend sind die Überdeckungshäufigkeiten der mit jenen Schätzungen der Standardabweichung gebildeten Konfidenzintervalle deutlich über den vorgegebenen Niveaus von 95 % bzw. 99 %. Die Konfidenzintervalle auf Basis der Formel mit NV-Annahme sind hier besser.

Schauen wir uns nun die alternative Version des Modigliani/Modigliani-Maßes  $MM_{alt} = MM_{ori} - \mu_M$  an. Die Ergebnisse in Tab. 5.16 wurden aus denselben simulierten Renditedaten berechnet wie jene in Tab. 5.15 und sind somit unmittelbar vergleichbar. Wie bei  $\widehat{MM}_{ori}$  ist eine geringfügige Verzerrung des Schätzers nach oben erkennbar. Auch die Abweichung der Verteilung des Schätzers von der Normalverteilung ist allenfalls minimal, vgl. Abb. 5.17a.

<sup>25</sup>Der absolute Wert der Standardabweichung hängt jedoch stark vom hier nicht variierten Parameter  $\sigma_M$  ab.



Verteilungsstatistiken bei Variation der Simulationsparameter  $\mu_X$ ,  $\sigma_X$  und  $\rho_{X,M}$ . Ausgangsparameterkombination:  $\mu_X = 0,003$ ,  $\mu_M = 0,002$ ,  $\sigma_X = \sigma_M = 0,1$ ,  $\rho_{X,M} = 0,75$  (jeweils senkrechte gestrichelte Linien).

Daten: MGARCH(1)-Prozesse,  $n = 60$ , je 50 000 Simulationsläufe.

*oben*: Theoretischer Wert  $MM_{ori}$  und Mittelwert der Schätzungen (mit Datenpunkten); umliegend die theoretische  $1,96 - \sigma$ -Umgebung mit der i.i.d.-NV-Formel.

*Mitte*: Empirische Standardabweichung der Schätzer (durchgehende Linie, ohne Datenpunkte); asymptotische Standardabweichung (durchgehend, mit Datenpunkten) und mittlere geschätzte Standardabweichung (gestrichelt) mittels i.i.d.-NV-Formel (schwarz), und verteilungsfreier i.i.d.-Formel (grau).

*unten*: Anteil der Schätzwerte in  $1,96\sigma$ -Umgebung nach i.i.d. NV-Formel (Linie ohne Datenpunkte), Überdeckungshäufigkeit des 95%-KI (durchgehende Linie mit Datenpunkten) und des 99%-KI (gestrichelte Linie, m.D.) für die i.i.d.-NV-Formel (schwarz) und die verteilungsfreie i.i.d.-Formel (grau).

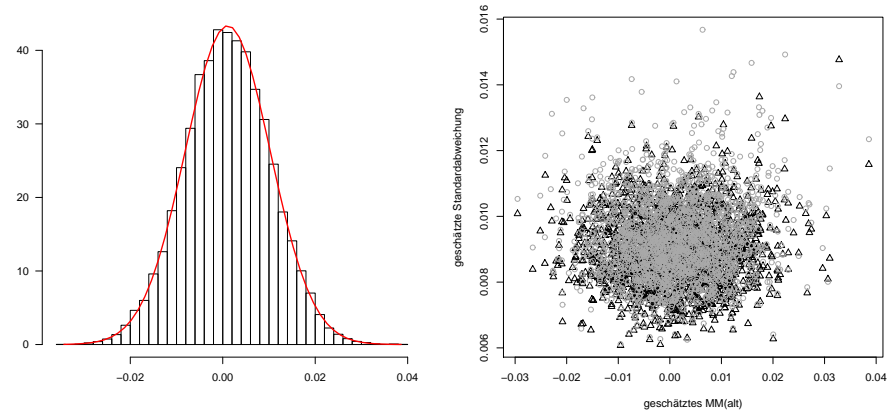
Abbildung 5.16: Grafiken zu  $\widehat{MM}_{ori}(X)$  bei Parametervariation

Verteilung	$\sigma_0(\hat{\theta})$	$\bar{s}(\hat{\theta}_i)$	$p_{\sigma^2=\sigma_0^2}$	$[u_{0,05}]$	$t_i$ in $u$	$KI_{95\%}$ um $\theta$
<b>NV</b>	$\theta = 0,001$ ; $\bar{t} = 0,00104$ ; $s(t_i) = 0,00925$ ; $p_{AD} = 0,0258$ ; $\hat{\psi} = 0,0268$					
i.i.d., NV	0,00913	0,00907	0,0001	[-0,0169;0,0189]	94,70%	94,54%
i.i.d.	0,00913	0,00936	0,0001	[-0,0169;0,0189]	94,70%	95,37%
<b>t<sub>10</sub></b>	$\theta = 0,001$ ; $\bar{t} = 0,00112$ ; $s(t_i) = 0,00921$ ; $p_{AD} = 0,0590$ ; $\hat{\psi} = 0,0273$					
i.i.d., NV	0,00913	0,00906	0,0112	[-0,0169;0,0189]	94,74%	94,74%
i.i.d.	0,00915*	0,00950	0,0653	[-0,0169;0,0189]	94,79%	95,96%
<b>t<sub>5</sub></b>	$\theta = 0,001$ ; $\bar{t} = 0,00105$ ; $s(t_i) = 0,00921$ ; $p_{AD} = 0,0068$ ; $\hat{\psi} = 0,0249$					
i.i.d., NV	0,00913	0,00902	0,0106	[-0,0169;0,0189]	94,84%	94,71%
i.i.d.	0,00920*	0,00976	0,7598	[-0,0170;0,0190]	94,97%	96,37%
<b>MGARCH</b>	$\theta = 0,001$ ; $\bar{t} = 0,00107$ ; $s(t_i) = 0,00920$ ; $p_{AD} = 0,0233$ ; $\hat{\psi} = 0,0287$					
i.i.d., NV	0,00913	0,00907	0,0126	[-0,0169;0,0189]	94,82%	94,90%
i.i.d.	0,00914*	0,00940	0,0285	[-0,0169;0,0189]	94,84%	95,82%
* aus den Daten simulierte Werte						
Erläuterung der Kürzel für die Statistiken in Tab. 5.2 auf S. 114.						

Tabelle 5.16: Ergebnisse für die alternative Version von  $MM$  bei Zeitreihenlänge  $n = 60$ , Ausgangsparameter

Die Standardabweichung von  $\widehat{MM}_{alt}$  ist mit ca. 0,009 etwas kleiner als jene von  $\widehat{MM}_{ori}$  (ca. 0,013). Grund dafür ist die in den Simulationsparametern vorgegebene positive Korrelation zwischen  $X$  und  $M$  von  $\rho_{M,X} = 0,75$ . Hohe Werte von  $\hat{\mu}_X$  gehen tendenziell mit hohen Werten von  $\hat{\mu}_M$  einher und da  $\hat{\mu}_M$  von  $s_M/s_X \cdot \hat{\mu}_X$  abgezogen wird, gleichen sich die Auswirkungen teilweise aus. Die empirische Standardabweichung  $s(t_i)$  liegt bei allen vier Verteilungsmodellen knapp über der asymptotischen Standardabweichung  $\sigma_0(\widehat{MM}_{alt})$ . Bei der Schätzung der Standardabweichung aus den Daten liefert die Varianzformel mit Normalverteilungsannahme zumeist bessere Ergebnisse. Mit der Formel ohne Normalverteilungsannahme werden die Standardabweichungen leicht überschätzt, was zu erhöhten Überdeckungshäufigkeiten der entsprechenden Konfidenzintervalle führt. Diese Abweichung ist weniger stark ausgeprägt als bei  $\widehat{MM}_{ori}$ .

Bei Variation der Simulationsparameter kann man in Abb. 5.18 erkennen, dass sich die Werte für  $MM_{alt}$  und  $\widehat{MM}_{alt}$  im Vergleich zu  $MM_{ori}$  (Abb. 5.16) lediglich um  $\mu_M = 0,002$  nach unten verschieben. Die Varianz von  $\widehat{MM}_{alt}$  ist hingegen von  $\rho_{M,X}$  abhängig. Ist  $\rho_{M,X}$  größer als etwa 0,5, streut der Schätzer der alternativen Version von  $MM$  weniger als jener von  $\widehat{MM}_{ori}$ . Für schwach bzw. negativ korrelierte Portfolio- und Marktrenditen hat  $\widehat{MM}_{alt}$  die stärkere Streuung. Eine nennenswerte Verzerrung von  $\widehat{MM}_{alt}$  und  $\hat{\sigma}_{\widehat{MM}_{alt}}$  ist nur für letzteren Wert bei hoher Korrelation von  $M$  und  $X$  und nur bei Verwendung der Varianzformel ohne Normalverteilungsannahme erkennbar. Das schlägt sich darin nieder, dass die Überdeckungshäufigkeiten der entsprechenden Konfidenzintervalle das theoretische Niveau vor allem bei höherem  $\rho_{M,X}$  deutlich überschreiten. Der Abstand zu den Werten 95 % bzw. 99% ist aber durchweg kleiner als bei  $\widehat{MM}_{ori}$ .

(a) Histogramm  $\widehat{MM}_{alt}$ (b)  $\widehat{MM}_{alt}-\widehat{\sigma}_{\widehat{MM}_{alt}}$ -Diagramm

Links: Histogramm von  $\widehat{MM}_{alt}$  mit angepasster Dichtefunktion der Normalverteilung

Rechts: Punktdiagramm für 2000 Zeitreihen mit  $\widehat{MM}_{alt}$  (Abszisse) und zugehörigem  $\widehat{\sigma}_{\widehat{MM}_{alt}}$  (Ordinate) für die verteilungsfreie Varianzformel( $\circ$ ) und die Varianzformel mit Normalverteilungsannahme( $\triangle$ ).

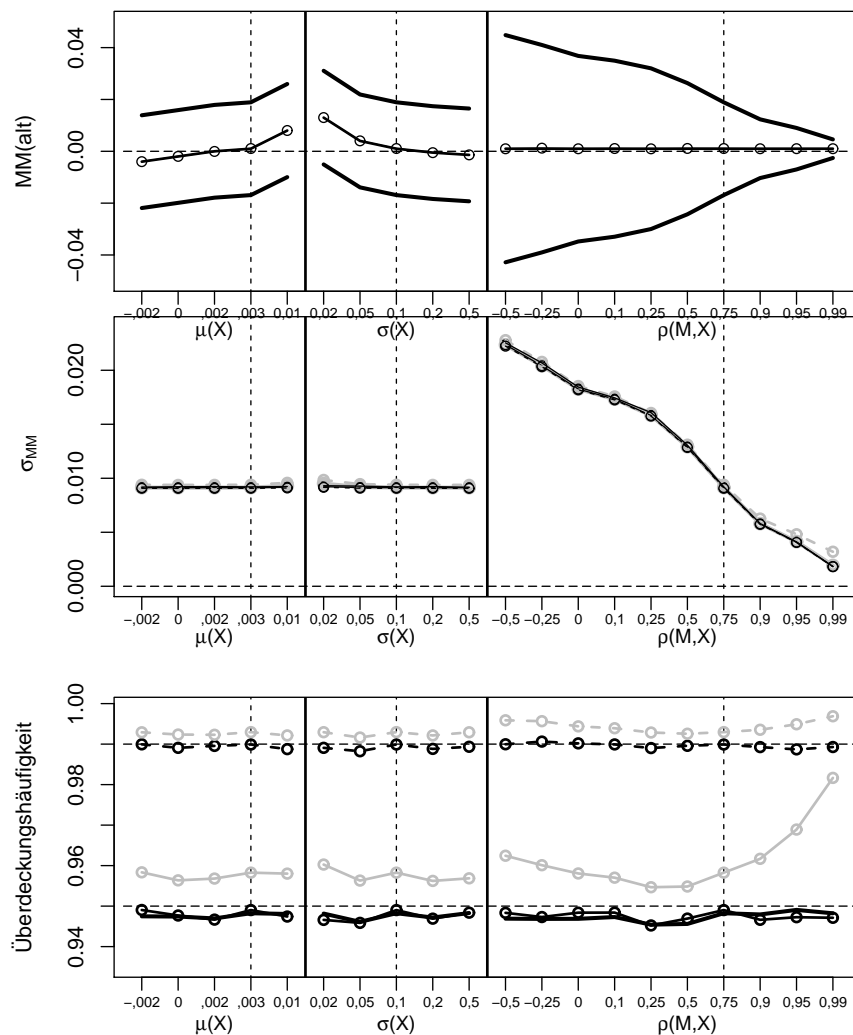
Abbildung 5.17: Grafiken zur Verteilung von  $\widehat{MM}_{alt}$  für i.i.d.- $t_{10}$ -verteilte Daten mit Zeitreihenlänge  $n = 60$ .

Schließlich noch Blick auf die gemeinsame Verteilung der beiden Versionen  $\widehat{MM}_{ori}$  und  $\widehat{MM}_{alt}$ : Abb. 5.19a zeigt ein Streudiagramm der beiden Schätzer, berechnet aus denselben Datenreihen. Es zeigt sich, dass der Zusammenhang zwischen  $\widehat{MM}_{ori}$  und  $\widehat{MM}_{alt}$  recht schwach ausgeprägt ist. Die Korrelation der beiden Varianten miteinander ist mit einem Korrelationskoeffizienten von 0,386 deutlich schwächer als die Korrelation der zugrunde liegenden Renditen,  $\rho_{M,X} = 0,75$ .

Bildet man für zwei Portfolios  $X$  und  $Y$  die **Differenz  $\Delta_{MM_{X,Y}}$  der Werte des Performancemaßes nach Modigliani/Modigliani**, so gibt es, wie bereits erwähnt, keinen Unterschied zwischen den beiden Versionen. In den Abbildungen 5.19b und 5.19c zeigt sich eine deutlich stärkere Korrelation der Performancemaßschätzer für  $MM_{ori}$ , die in etwa der Korrelation von  $X$  und  $M$  entspricht. Die Korrelation von  $\widehat{MM}_{alt,X}$  und  $\widehat{MM}_{alt,Y}$  ist deutlich schwächer.

In den Statistiken zu  $\widehat{\Delta}_{MM_{X,Y}}$  (Tab. 5.17) erkennt man eine deutliche Ähnlichkeit mit den Werten in Tabelle 5.16 ( $\widehat{MM}_{alt,X}$ ). Theoretischer Wert und asymptotische Standardabweichung sind gleich<sup>26</sup> und auch bei den restlichen

<sup>26</sup>Ersteres ergibt sich zufällig aus den gewählten Parametern, da  $\mu_Y = \mu_X$  und somit  $MM_{ori,X} = 0$  bzw.  $MM_{alt,Y} = 0$ .



Verteilungsstatistiken bei Variation der Simulationsparameter  $\mu_X$ ,  $\sigma_X$  und  $\rho_{X,M}$ . Ausgangsparameterkombination:  $\mu_X = 0,003$ ,  $\mu_M = 0,002$ ,  $\sigma_X = \sigma_M = 0,1$ ,  $\rho_{X,M} = 0,75$  (jeweils senkrechte gestrichelte Linien).

Daten: MGARCH(1)-Prozesse,  $n = 60$ , je 50 000 Simulationsläufe.

*oben:* Theoretischer Wert  $MM_{alt}$  und Mittelwert der Schätzungen (mit Datenpunkten); umliegend die theoretische  $1,96 - \sigma$ -Umgebung mit der i.i.d.-NV-Formel.

*Mitte:* Empirische Standardabweichung der Schätzer (durchgehende Linie, ohne Datenpunkte); asymptotische Standardabweichung (durchgehend, mit Datenpunkten) und mittlere geschätzte Standardabweichung (gestrichelt) mittels i.i.d.-NV-Formel (schwarz), und verteilungsfreier i.i.d.-Formel (grau).

*unten:* Anteil der Schätzwerte in  $1,96\sigma$ -Umgebung nach i.i.d. NV-Formel (Linie ohne Datenpunkte), Überdeckungshäufigkeit des 95%-KI (durchgehende Linie mit Datenpunkten) und des 99%-KI (gestrichelte Linie, m.D.) für die i.i.d.-NV-Formel (schwarz) und die verteilungsfreie i.i.d.-Formel (grau).

Abbildung 5.18: Grafiken zu  $\widehat{MM}_{alt}(X)$  bei Parametervariation

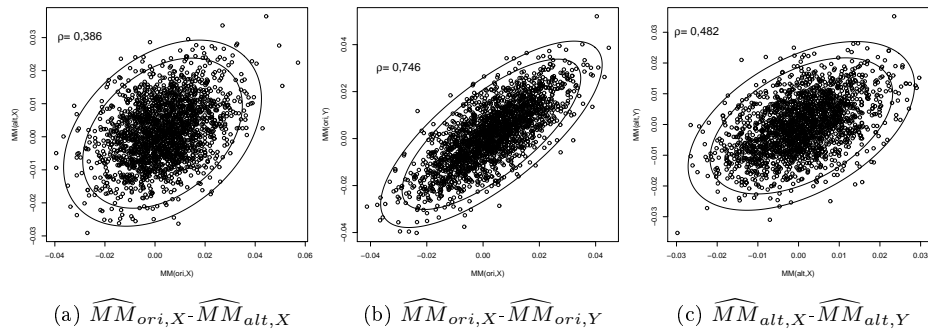
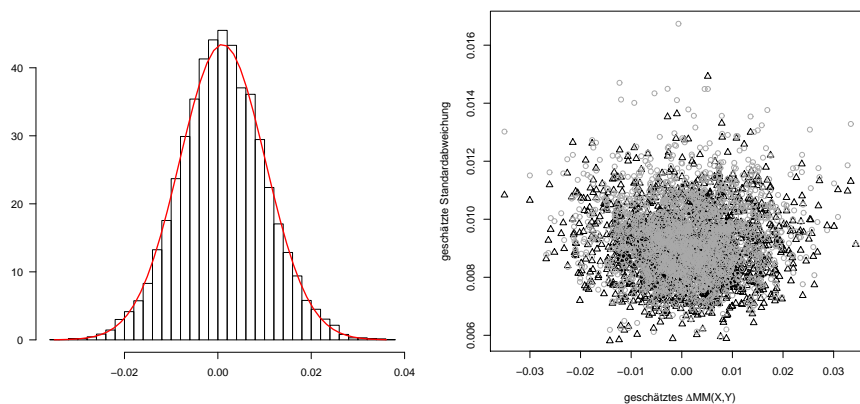


Abbildung 5.19: Streudiagramme zu Modigliani-Modigliani mit 95%- und 99%-NV-Zentralbereichen; i.i.d.- $t_{10}$ -verteilte Renditen, Zeitreihenlänge  $n = 60$ , je 2000 Zeitreihen.

Verteilung	$\sigma_0(\hat{\theta})$	$\bar{s}(\hat{\theta}_i)$	$p_{\sigma^2=\sigma_0^2}$	$[u_{0,05}]$	$t_i$ in $u$	$KI_{95\%}$ um $\theta$
<b>NV</b>	$\theta = 0,001$ ; $\bar{t} = 0,00101$ ; $s(t_i) = 0,00924$ ; $p_{AD} = 0,1424$ ; $\hat{\psi} = 0,0330$					
i.i.d., NV	0,00913	0,00910	0,0207	[-0,0169;0,0189]	94,58%	94,72%
i.i.d.	0,00913	0,00937	0,0201	[-0,0169;0,0189]	94,58%	95,39%
<b>t<sub>10</sub></b>	$\theta = 0,001$ ; $\bar{t} = 0,00095$ ; $s(t_i) = 0,00923$ ; $p_{AD} = 0,0266$ ; $\hat{\psi} = 0,0354$					
i.i.d., NV	0,00913	0,00909	0,0349	[-0,0169;0,0189]	94,69%	95,10%
i.i.d.	0,00912*	0,00950	0,0227	[-0,0169;0,0189]	94,64%	95,99%
<b>t<sub>5</sub></b>	$\theta = 0,001$ ; $\bar{t} = 0,00114$ ; $s(t_i) = 0,00924$ ; $p_{AD} = 0,0023$ ; $\hat{\psi} = 0,0311$					
i.i.d., NV	0,00913	0,00906	0,0133	[-0,0169;0,0189]	94,74%	94,99%
i.i.d.	0,00913*	0,00982	0,0133	[-0,0169;0,0189]	94,74%	96,41%
<b>MGARCH</b>	$\theta = 0,001$ ; $\bar{t} = 0,00100$ ; $s(t_i) = 0,00927$ ; $p_{AD} = 0,0048$ ; $\hat{\psi} = 0,0351$					
i.i.d., NV	0,00913	0,00911	0,0032	[-0,0169;0,0189]	94,61%	94,90%
i.i.d.	0,00914*	0,00942	0,0056	[-0,0169;0,0189]	94,62%	95,66%

\* aus den Daten simulierte Werte  
Erläuterung der Kürzel für die Statistiken in Tab. 5.2 auf S. 114.

Tabelle 5.17: Ergebnisse für die Differenz der  $MM$  zweier Portfolios bei Zeitreihenlänge  $n = 60$ , Ausgangsparameter

(a) Histogramm  $\hat{\Delta}_{MM_{X,Y}}$ (b)  $\hat{\Delta}_{MM_{X,Y}} - \hat{\sigma}_{\hat{\Delta}_{MM(X,Y)}}$ -Diagramm

Links: Histogramm von  $\hat{\Delta}_{MM_{X,Y}}$  mit angepasster Dichtefunktion der Normalverteilung  
 Rechts: Punktdiagramm für 2000 Zeitreihen mit  $\hat{\Delta}_{MM_{X,Y}}$  (Abszisse) und zugehörigem  $\hat{\sigma}_{\hat{\Delta}_{MM(X,Y)}}$  (Ordinate) für die verteilungsfreie Varianzformel( $\circ$ ) und die Varianzformel mit Normalverteilungsannahme( $\Delta$ ).

Abbildung 5.20: Grafiken zur Verteilung von  $\hat{\Delta}_{MM_{X,Y}}$  für i.i.d.- $t_{10}$ -verteilte Daten mit Zeitreihenlänge  $n = 60$ .

Ergebnissen findet man keine gravierende Abweichung. Die aus den Parametern berechnete asymptotische Standardabweichung  $\sigma_0(\theta)$  liegt nahe der beobachteten Standardabweichung der Schätzwerte  $s(t_i)$ . Die Anpassung an die Normalverteilung ist bei allen Verteilungsmodellen befriedigend bis gut (vgl. Abb. 5.20a). Wiederum scheint der Varianzschätzer ohne Normalverteilungsannahme eine leichte Verzerrung nach oben aufzuweisen. Hierzu sieht man in Abb. 5.20b, dass Schätzungen der Standardabweichung nach der Formel ohne Normalverteilungsannahme häufiger oberhalb des Kerns der Punktwolke liegen. Eine Abhängigkeit vom Schätzwert  $\hat{\Delta}_{MM_{X,Y}}$  ist nicht erkennbar.

Werden die Verteilungsparameter verändert, ändern sich die obigen Ergebnisse qualitativ nicht. Die Anpassung an die Normalverteilung wird lediglich dann schlechter, wenn die Korrelation zwischen den Portfolios  $X$  und  $Y$  sehr hoch und deren jeweilige Korrelation mit dem Marktportfolio niedrig ist. Abgesehen von diesem Fall sind die mittels NV-Formel aus den Parametern berechnete Standardabweichung und die mit dieser Formel vorgenommenen Schätzungen der Konfidenzintervalle sehr zuverlässig. Konfidenzintervalle auf Basis der Formel ohne Normalverteilungsannahme sind dagegen um so konservativer (d.h. die Überdeckungshäufigkeiten gehen über das vorgegebene Niveau hinaus), je stärker  $X$  und  $Y$  korreliert sind.

### 5.3.5 Kappa

Das Performancemaß Kappa unterscheidet sich in zweierlei Hinsicht von den bisherigen Maßen: Erstens bestanden alle bisherigen Maße und deren Varianzformeln ausschließlich aus gewöhnlichen und zentralen Momenten der (ggf. multivariaten) Renditeverteilung, wohingegen Kappa auch partielle Momente enthält. Schon in Abschnitt 5.2 wurde sichtbar, dass die Anpassung der Verteilung des Schätzers an die Normalverteilung schlechter ist als bei den anderen Performancemaßen. Zweitens hatten die Performancemaße der vorigen Abschnitte keine Parameter: bei gegebenen Daten steht die Berechnung des Maßes fest. Bei Kappa muss der Anwender die Parameter  $\alpha$  und  $z$  festlegen. Mittels  $\alpha \in \mathbb{N}$  wird die Gewichtung der Verluste untereinander bestimmt: je höher  $\alpha$ , desto stärker die Gewichtung extremer Verluste. Jede Ausprägung von  $\alpha$  stellt quasi ein eigenes Performancemaß dar. In diesem Abschnitt werden die Ausprägungen  $\alpha \in \{1, 2, 3\}$  mittels Simulation getestet. Der Parameter  $z$  lässt die Wahl der Vergleichsrendite offen, die bei den anderen Performancemaßen auf den risikolosen Zinssatz festgelegt ist. Wir können aber darauf verzichten, Variationen des Parameters  $z$  gesondert zu simulieren, da bei Kappa im Zähler und Nenner lediglich Abstände zu  $z$  gebildet werden.<sup>27</sup>

Die Simulationsergebnisse in Tabelle 5.18 sind zunächst nach dem Parameter  $\alpha$  und anschließend nach Verteilungsmodell sortiert.<sup>28</sup> Es gibt hier keine gesonderte Varianzformel für normalverteilte Daten.

Wie sich schon in den Ergebnissen für  $n = 500$  (Tab. 5.4, S. 116; Tab. 5.7, S. 121) andeutete, haben die Schätzer für die Kappa-Maße in deutlich geringerem Maße jene wünschenswerten Eigenschaften (Unverzerrtheit, approximative Normalverteilung), die bei den Schätzern der anderen Performancemaße in mehr oder weniger hinreichendem Grade gegeben waren.

Für  $\alpha = 1$  ist schon bei normalverteilten Daten die Abweichung zwischen theoretischem Wert<sup>29</sup>  $\theta = 0,0781$  und dem Mittelwert der Schätzungen  $\bar{t} = 0,1425$  erheblich. Für Verteilungsmodelle mit höherer Kurtosis wächst diese Differenz weiter an.<sup>30</sup> Die Verteilung von  $\hat{\kappa}(\alpha = 1)$  weicht erheblich von der Normalverteilung ab. Die Maße  $\hat{\psi}$  schlagen stets deutlich nach oben aus und die systematisch abweichende Verteilungsform ist im Histogrammen links oben in Abb. 5.21 auf den ersten Blick erkennbar. Auch die theoretische Standardabweichung  $\sigma_0$  liegt deutlich unter der empirischen Standardabweichung  $s(\hat{\kappa}_i)$  in Höhe von 0,3936. Entsprechend enthält die 1,96 $\sigma$ -Umgebung des theoretischen Werts nur 91,36 % der Schätzwerte. Die geschätzten Standardabweichungen liegen mit einem Mittelwert von 0,3982 nahe an der empirischen Standardabweichung und die Überdeckungshäufigkeit der Konfidenzintervalle ist mit 93,77 % deutlich näher

<sup>27</sup>Man kann für  $b \in \mathbb{R}$  leicht zeigen, dass  $\kappa_X(\alpha, z) = \kappa_{X+b}(\alpha, z + b)$ .

<sup>28</sup>Weitere Parameter:  $\mu_X = 0,003$ ,  $\sigma_X = 0,1$ ,  $z = 0$ ,  $n = 60$ ,  $R = 50\,000$

<sup>29</sup>Die theoretischen Werte der Performancemaße wurden – außer bei den GARCH-Daten – mittels numerischer Integration der Normal- bzw.  $t$ -Verteilung ermittelt.

<sup>30</sup>Der hohe theoretische Wert  $\theta = 0,1256$  bei GARCH-Daten ist Ergebnis einer Simulation (Schätzung anhand einer Zeitreihe mit  $n = 3\,000\,000$ ) und nicht mit den theoretischen Werten der i.i.d.-Verteilungsmodelle vergleichbar. Dieser (unrealistische) Wert ist Ausdruck dessen, dass der Schätzer bei den gegebenen GARCH-Parametern nicht konvergiert.



Verteilung	$\sigma_0(\hat{\theta})$	$\bar{s}(\hat{\theta}_i)$	$p_{\sigma^2=\sigma_0^2}$	$[u_{0,05}]$	$t_i$ in $u$	$KI_{95\%}$ um $\theta$
$\alpha = 1$						
<b>NV</b>	$\theta = 0,0781$ ; $\bar{t} = 0,1425$ ; $s(t_i) = 0,3936$ ; $p_{AD} < 10^{-99}$ ; $\hat{\psi} = 0,1327$					
i.i.d.	0,3240	0,3982	$< 10^{-99}$	[-0,5569;0,7131]	91,36%	93,77%
<b>t<sub>10</sub></b>	$\theta = 0,0719$ ; $\bar{t} = 0,1516$ ; $s(t_i) = 0,4073$ ; $p_{AD} < 10^{-99}$ ; $\hat{\psi} = 0,1392$					
i.i.d.	0,2962	0,4131	$< 10^{-99}$	[-0,5088;0,6525]	88,01%	93,90%
<b>t<sub>5</sub></b>	$\theta = 0,0653$ ; $\bar{t} = 0,1567$ ; $s(t_i) = 0,4323$ ; $p_{AD} < 10^{-99}$ ; $\hat{\psi} = 0,1463$					
i.i.d.	0,2663	0,4346	$< 10^{-99}$	[-0,4567;0,5872]	83,09%	93,76%
<b>t<sub>3</sub></b>	$\theta = 0,0559$ ; $\bar{t} = 0,1943$ ; $s(t_i) = 0,5088$ ; $p_{AD} < 10^{-99}$ ; $\hat{\psi} = 0,1662$					
i.i.d.	0,2207	0,5021	$< 10^{-99}$	[-0,3767;0,4885]	70,14%	93,44%
<b>GARCH</b>	$\theta = 0,1256^*$ ; $\bar{t} = 0,2265$ ; $s(t_i) = 0,4867$ ; $p_{AD} < 10^{-99}$ ; $\hat{\psi} = 0,1569$					
i.i.d.	0,5172*	0,4840	$< 10^{-99}$	[-0,8881;1,1392]	95,31%	93,32%
$\alpha = 2$						
<b>NV</b>	$\theta = 0,0435$ ; $\bar{t} = 0,0669$ ; $s(t_i) = 0,2067$ ; $p_{AD} < 10^{-99}$ ; $\hat{\psi} = 0,0886$					
i.i.d.	0,1827	0,2050	$< 10^{-99}$	[-0,3146;0,4016]	92,40%	94,94%
<b>t<sub>10</sub></b>	$\theta = 0,0387$ ; $\bar{t} = 0,0711$ ; $s(t_i) = 0,2097$ ; $p_{AD} < 10^{-99}$ ; $\hat{\psi} = 0,0995$					
i.i.d.	0,1620	0,2077	$< 10^{-99}$	[-0,2787;0,3562]	88,60%	94,89%
<b>t<sub>5</sub></b>	$\theta = 0,0334$ ; $\bar{t} = 0,0729$ ; $s(t_i) = 0,2165$ ; $p_{AD} < 10^{-99}$ ; $\hat{\psi} = 0,1131$					
i.i.d.	0,1379	0,2117	$< 10^{-99}$	[-0,2369;0,3038]	82,57%	94,65%
<b>GARCH</b>	$\theta = 0,0483^*$ ; $\bar{t} = 0,1053$ ; $s(t_i) = 0,2319$ ; $p_{AD} < 10^{-99}$ ; $\hat{\psi} = 0,1253$					
i.i.d.	0,2753*	0,2226	$< 10^{-99}$	[-0,4912;0,5878]	96,51%	94,87%
$\alpha = 3$						
<b>NV</b>	$\theta = 0,0330$ ; $\bar{t} = 0,0480$ ; $s(t_i) = 0,1560$ ; $p_{AD} < 10^{-99}$ ; $\hat{\psi} = 0,7493$					
i.i.d.	0,1393	0,1536	$< 10^{-99}$	[-0,2401;0,3059]	92,39%	95,13%
<b>t<sub>10</sub></b>	$\theta = 0,0284$ ; $\bar{t} = 0,0509$ ; $s(t_i) = 0,1558$ ; $p_{AD} < 10^{-99}$ ; $\hat{\psi} = 0,0870$					
i.i.d.	0,1192	0,1528	$< 10^{-99}$	[-0,2053;0,2621]	88,09%	95,14%
<b>GARCH</b>	$\theta = 0,0176^*$ ; $\bar{t} = 0,0743$ ; $s(t_i) = 0,1662$ ; $p_{AD} < 10^{-99}$ ; $\hat{\psi} = 0,1208$					
i.i.d.	0,3879*	0,1563	$< 10^{-99}$	[-0,7427;0,7778]	99,72%	95,72%
* aus den Daten simulierte Werte						
Erläuterung der Kürzel für die Statistiken in Tab. 5.2 auf S. 114.						

Tabelle 5.18: Ergebnisse für Kappa( $\alpha, z = 0$ ) bei Zeitreihenlänge  $n = 60$ , Ausgangsparameter

am Zielwert. Bei i.i.d.- $t$ -verteilten Daten sind die gerade beschriebenen Beobachtungen mit sinkender Zahl an Freiheitsgraden noch akzentuierter sichtbar. Da für die Konvergenz von  $\sigma_{\hat{\kappa}_1}$  nur die ersten beiden Momente der Renditeverteilung endlich sein müssen, wurden hier zusätzlich  $t_3$ -Daten simuliert und ausgewertet.

Für  $\kappa_2$  und  $\kappa_3$  gilt qualitativ dasselbe wie für  $\kappa_1$ , nur dass die Abweichung der Schätzerverteilung von der Normalverteilung geringer ausfällt. Auch die Unterschiede zwischen den theoretischen Werten  $\theta$  bzw.  $\sigma_0$  und den entsprechenden empirischen Größen werden kleiner. Wir werden allerdings weiter unten sehen, dass die Überdeckungshäufigkeiten der Konfidenzintervalle, die in der Tabelle recht nahe an 95 % liegen, stark vom Simulationsparameter  $\mu_X$  abhängen.

In den Streudiagrammen der Performancemaß-Schätzwerte mit den zugehörigen Schätzungen der Standardabweichung (Abb. 5.21 rechts) fällt die starke Korrelation zwischen den beiden Größen auf, die in dieser Stärke bei den anderen Maßen nirgends feststellbar ist. Die Werte der empirischen Korrelationskoeffizienten  $r_{\hat{\kappa}_\alpha, \hat{\sigma}(\hat{\kappa}_\alpha)}$  betragen 0,9904 ( $\alpha = 1$ ), 0,9766 ( $\alpha = 2$ ) und 0,9234 ( $\alpha = 3$ ). Gerade bei  $\kappa_1$  (= Omega - 1) sind Schätzwert und Schätzvarianz also nahezu perfekt linear abhängig.

Eine starke Korrelation kann man auch zwischen den drei Schätzern  $\hat{\kappa}_1$ ,  $\hat{\kappa}_2$  und  $\hat{\kappa}_3$  beobachten. Die empirische Korrelationsmatrix zwischen den drei Werten war (bei den i.i.d.- $t_{10}$ -verteilten Simulationswerten)

$$\begin{pmatrix} 1 & 0,9937 & 0,9874 \\ 0,9937 & 1 & 0,9986 \\ 0,9874 & 0,9986 & 1 \end{pmatrix}.$$

Die Ergebnisse der drei Performancemaße  $\kappa_1$ ,  $\kappa_2$  und  $\kappa_3$  für ein bestimmtes Portfolio unterscheiden sich also bis auf den Absolutwert des Maßes kaum. Dieser Sachverhalt wird auch in den Streudiagrammen in Abb. 5.21 sichtbar, die aus denselben Renditedaten berechnet wurden: An den äußeren Enden der Punktwolken kann man Muster einzelner Datenpunkte erkennen, die sich vor allem bei  $\alpha = 2$  und  $\alpha = 3$  wiederholen und offenbar aus denselben Simulationszeitreihen stammen.

Abbildung 5.22 zeigt die bekannten Grafiken bei Parametervariation. Es wird dabei nur der Parameter  $\mu_X$  für  $\alpha = 1, 2, 3$  verändert,  $\sigma_X = 0, 1$  und  $z = 0$  bleiben gleich. In der oberen Tafel sieht man, dass der Mittelwert der Schätzungen von Kappa bei sehr niedrigen Werten für  $\mu_X$  nahe am theoretischen Wert liegt. Mit steigendem  $\mu_X$  zeigt sich eine zunehmende und sehr deutliche Aufwärtsverzerrung des Performancemaßschätzers. Auch die Abweichung der Schätzerverteilung von der Normalverteilung, gemessen mit  $\hat{\psi}$  (nicht dargestellt), nimmt mit  $\mu_X$  zu. Das ist dadurch erklärbar, dass bei steigender „Überrendite“  $\mu_X - z$  (bei gegebenen  $\sigma_X$  und  $n$ ) tendenziell immer weniger der  $n$  Renditen in den Bereich unterhalb von  $z$  fallen und damit in die Berechnung des partiellen Moments eingehen. Außerdem sind  $(\bar{X} - z)$  und  $\hat{E}(z - X)_+^\alpha$  negativ korreliert<sup>31</sup>, so

<sup>31</sup>Hat eine Einzelrendite  $X_i$  statt dem Wert  $x_i$  den höheren Wert  $x_i^+$ , steigt  $\bar{X}$  an und  $\hat{E}(z - X)_+^\alpha$  sinkt (wenn  $x_i < z$ ) oder bleibt gleich (wenn  $x_i \geq z$ ).

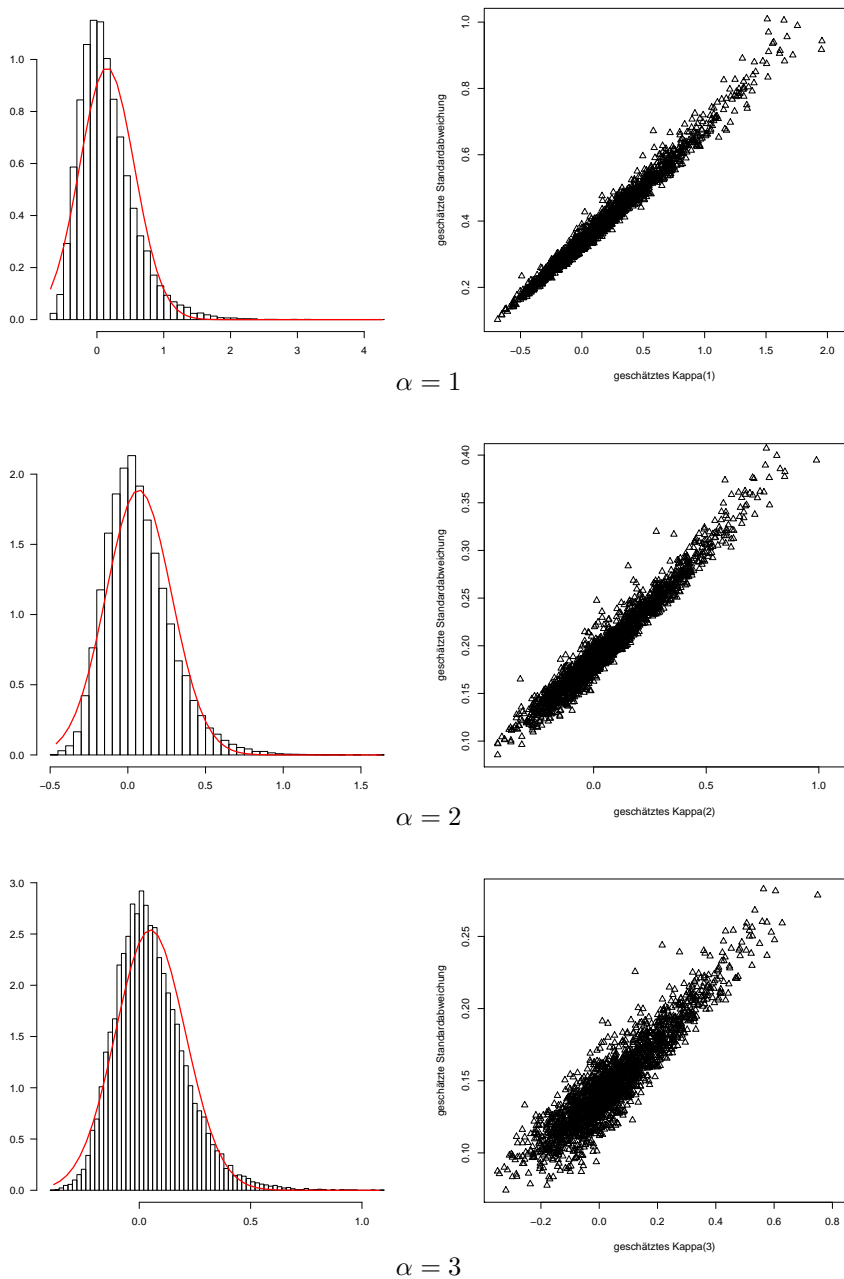


Abbildung 5.21: Grafiken für  $\text{Kappa}_X(\alpha)$  bei verschiedenen Werten für  $\alpha$  mit i.i.d- $t_{10}$ -verteilten Daten und Zeitreihenlänge  $n = 60$ : Histogramme (20 000 Werte, links); Punktdiagramme  $\hat{\kappa}_X(\alpha, z=0)$  gegen  $\hat{\sigma}_{\kappa_X(\alpha, z=0)}$  (2 000 Werte, rechts).

dass bei einer Zeitreihe mit besonders hoher Rendite nicht nur der Zähler von Kappa hoch, sondern auch der Nenner niedrig ist.

Die asymptotische Standardabweichung von  $\widehat{\kappa}(\alpha, 0)$  variiert im untersuchten Bereich  $\mu_X \in [-0,02; 0,02]$  bei  $\sigma_X = 0,1$  kaum. Die empirische Standardabweichung der Schätzwerte steigt dagegen ebenso wie der Mittelwert der Schätzungen von  $\sigma_{\widehat{\kappa}}$  mit  $\mu_X$  deutlich an. Ersteres ist auf im vorigen Absatz beschriebenen Effekt zurückzuführen, letzteres auch auf die bereits bei Abb. 5.21 thematisierte starke lineare Abhängigkeit von  $\widehat{\kappa}$  und  $\widehat{\sigma}_{\widehat{\kappa}}$ .

In der unteren Tafel von Abb. 5.22 ist sofort erkennbar, dass die asymptotische  $1,96\sigma$ -Umgebung vor allem bei hohem  $\mu_X$  weitaus weniger als 95 % der Schätzwerte umfasst (schwarze durchgezogene Linie). Bei niedrigem  $\mu_X$  sind dagegen die Überdeckungshäufigkeiten der Konfidenzintervalle zu niedrig. Alle diese Auffälligkeiten sind bei  $\alpha = 1$  deutlich stärker ausgeprägt als bei höheren Alphas.

Insgesamt lassen sich die Eigenschaften des Performancemaßes Kappa (und damit auch Omega und Sortino-Ratio) deutlich schlechter mit den hier genutzten asymptotischen Methoden beschreiben. Gerade bei  $\kappa_1$  (Omega) erscheint eine Berechnung auf Basis von 60 Monatsrenditen problematisch. Eine Möglichkeit ist der Umstieg auf Tagesrenditen. Ähnlich wie für das Treynor-Ratio wurden deswegen Simulationen zur Nachbildung entsprechender Tagesrenditen durchgeführt ( $n = 1250$ ,  $\mu_{X,Tag} = 60/1250 \cdot \mu_X$ ,  $\sigma_{X,Tag} = \sqrt{60/1250} \cdot \sigma_X$ ). In der unteren Tafel von Abb. 5.22 sind die Ergebnisse dieser Simulationen in grau eingezeichnet. Die Überdeckungshäufigkeit der Konfidenzintervalle liegt jetzt stets nahe des theoretischen Niveaus. Die ex-ante- $1,96\sigma$ -Intervalle bleiben aber relativ unzuverlässig. Stets fallen deutlich weniger als 95 % der Schätzwerte in dieses Intervall, auch wenn die Abweichung schwächer als bei  $n = 60$  ist. Für das Performancemaß Omega, ggf. auch das Sortino-Ratio ist also der Einsatz von Tagesdaten sinnvoll, wenn man Varianzen und Konfidenzintervalle nutzen will.

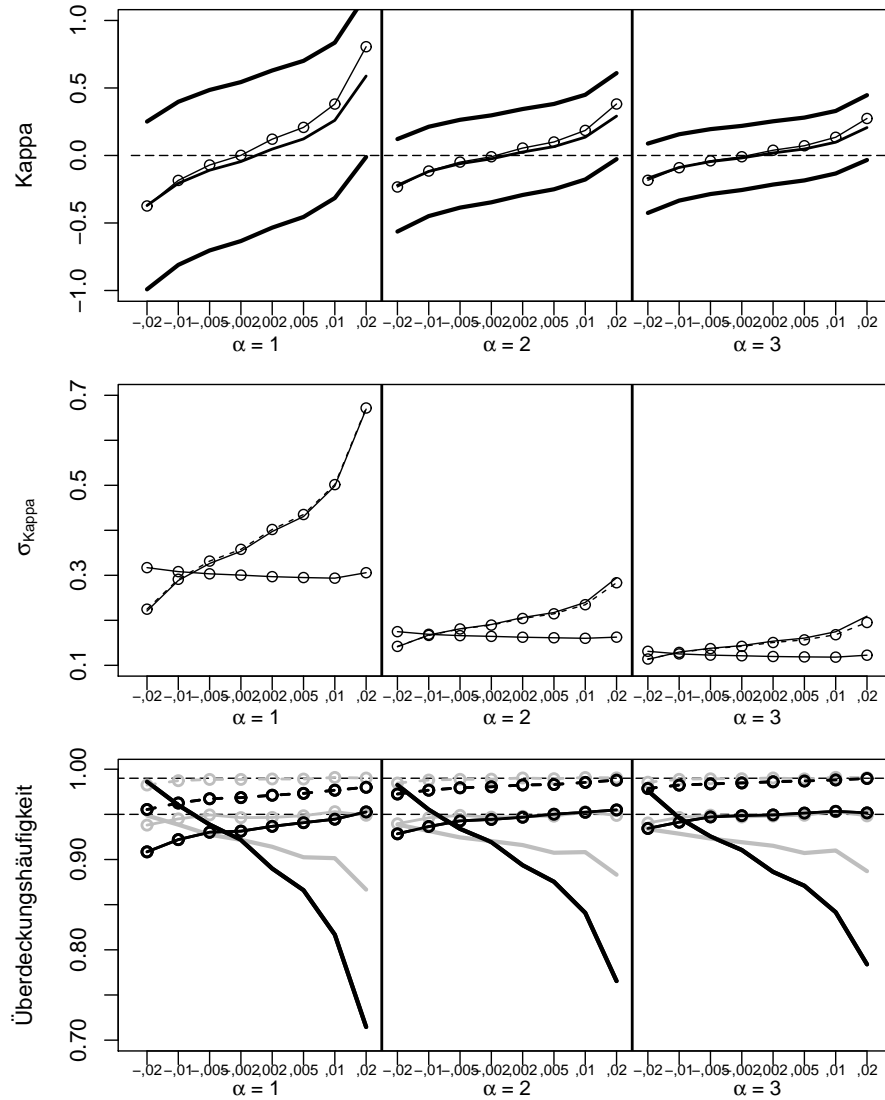
Betrachtet man die Kappa-Schätzwerte für zwei korrelierte Portfolios (siehe Abb. 5.23 links), findet man die Korrelation der Renditen ( $\rho_{X,Y} = 0,75$ ) fast identisch in den Schätzwerten der Performancemaße wieder ( $r_{\widehat{\kappa}_X, \widehat{\kappa}_Y} \approx 0,74$ ). Der Zusammenhang  $r_{\widehat{\kappa}_X, \widehat{\kappa}_Y} \approx \rho_{X,Y}$  konnte auch für andere Werte von  $\rho_{X,Y}$  beobachtet werden.<sup>32</sup> Reißt der Schätzwert eines Portfolios nach oben aus, gilt das meistens auch für den Schätzwert des anderen Portfolios.

Der Schätzer der Differenz der Kappas zweier Portfolios  $\Delta_{\kappa_{X,Y}}$  ist für  $n = 60$  nur in sehr grober Näherung normalverteilt. Tabelle 5.19 zeigt, dass für die Ausgangsparameterkombination<sup>33</sup> die Normalverteilungshypothese für jedes Verteilungsmodell und für jede Ausprägung von  $\alpha$  klar abgelehnt wird.<sup>34</sup> In den Histo-

<sup>32</sup>Das gilt nicht für stark negativ korrelierte Portfolios, da der Zusammenhang zwischen  $\widehat{\kappa}_X$  und  $\widehat{\kappa}_Y$  für  $\rho_{X,Y} \rightarrow -1$  annähernd hyperbolisch ist.

<sup>33</sup> $\mu_X = 0,003$ ,  $\mu_Y = 0,002$ ,  $\sigma_X = \sigma_Y = 0,1$ ,  $\rho_{X,Y} = 0,75$ ,  $n = 60$

<sup>34</sup>Die Berechnung der asymptotischen Standardabweichungen über numerische Integration von multivariaten Verteilungen erwies sich als nicht exakt genug, um sinnvoll Tests durchzuführen, die auf Abweichungen in höheren Nachkommastellen reagieren. Deswegen wurde



Verteilungsstatistiken bei Variation des Simulationsparameters  $\mu_X$  bei Kappa mit  $\alpha = 1, 2, 3$ . Daten: i.i.d.- $t_{10}$ -Zeitreihen,  $n = 60$ , je 50 000 Simulationsläufe.

*oben:* Theoretischer Wert  $\kappa_\alpha$  und Mittelwert der Schätzungen (mit Datenpunkten); umliegend die theoretische  $1,96\sigma$ -Umgebung.

*Mitte:* Empirische Standardabweichung der Schätzer (durchgehende Linie, ohne Datenpunkte); asymptotische Standardabweichung (durchgehend, mit Datenpunkten) und mittlere geschätzte Standardabweichung (gestrichelt).

*unten:* Schwarz: Anteil der Schätzwerte in  $1,96\sigma$ -Umgebung (Linie ohne Datenpunkte), Überdeckungshäufigkeit des 95%-KI (durchgehende Linie mit Datenpunkten) und des 99%-KI (gestrichelte Linie mit Datenpunkten). Grau: ebenso, für simulierte Tagesrenditen ( $n = 1250$ , entsprechend angepasste  $\mu_X$  und  $\sigma_X$ )

Abbildung 5.22: Grafiken zu  $\hat{\kappa}_X(\alpha)$  bei Variation von  $\mu_X$

Verteilung	$\theta$	$\bar{t}$	$p_{AD}$	$\hat{\psi}$	$\bar{s}(\hat{\theta}_i)$	$s(t_i)$	$KI_{95\%}$ um $\theta$
$\alpha = 1$							
NV	0,0267	0,0304	$< 10^{-96}$	0,0700	0,2944	0,2770	99,01 %
<b>t<sub>10</sub></b>	0,0245	0,0286	$< 10^{-99}$	0,0859	0,3077	0,2958	99,07 %
<b>t<sub>5</sub></b>	0,0222	0,0308	$< 10^{-99}$	0,0883	0,3274	0,3165	99,09 %
<b>t<sub>3</sub></b>	0,0190	0,0411	$< 10^{-99}$	0,1099	0,3813	0,3709	99,07 %
$\alpha = 2$							
NV	0,0147	0,0165	$< 10^{-24}$	0,0431	0,1488	0,1447	97,19 %
<b>t<sub>10</sub></b>	0,0131	0,0150	$< 10^{-52}$	0,0555	0,1516	0,1506	97,17 %
<b>t<sub>5</sub></b>	0,0113	0,0155	$< 10^{-81}$	0,0660	0,1571	0,1577	97,40 %
$\alpha = 3$							
NV	0,0111	0,0126	$< 10^{-14}$	0,0428	0,1114	0,1094	96,50 %
<b>t<sub>10</sub></b>	0,0096	0,0113	$< 10^{-41}$	0,0535	0,1116	0,1121	96,58 %
Erläuterung der Kürzel für die Statistiken in Tab. 5.2 auf S. 114.							

Tabelle 5.19: Ergebnisse für die Differenz von  $\text{Kappa}(\alpha, z = 0)$  zweier Portfolios bei Zeitreihenlänge  $n = 60$ , Ausgangsparameter

grammen in Abb. 5.23 Mitte sieht man, dass die Abweichung vor allem durch erhöhte Kurtosis verursacht wird. Die Schätzwert-Varianz-Diagramme rechts zeigen, dass mit höherer Differenz zwischen  $\hat{\kappa}_X$  und  $\hat{\kappa}_Y$  auch die Varianzschätzungen tendenziell höher sind. Deswegen ist es nicht verwunderlich, dass die Überdeckungshäufigkeiten der Konfidenzintervalle teilweise stark überhöht sind. Wiederum sind die Abweichungen von der Asymptotik bei  $\alpha = 1$  am stärksten ausgeprägt.

Die Auswertung weiterer Parameterkombinationen führt zumeist zu ähnlichen Ergebnissen. Die Normalverteilungshypothese wurde (bei 1270 verschiedenen Parameterkombinationen) immer zum 0,1 %-Niveau abgelehnt. Die Überdeckungshäufigkeit der 95%-Konfidenzintervalle lag für  $\alpha = 3$  stets im Bereich 93,7 - 96,7 %. Für  $\alpha = 1$  dagegen reichte die Bandbreite von knapp 92 % bis über 99 %.

Auch beim Vergleich zweier Portfolios hat Kappa hinsichtlich der Inferenzstatistik also ungünstige Eigenschaften.

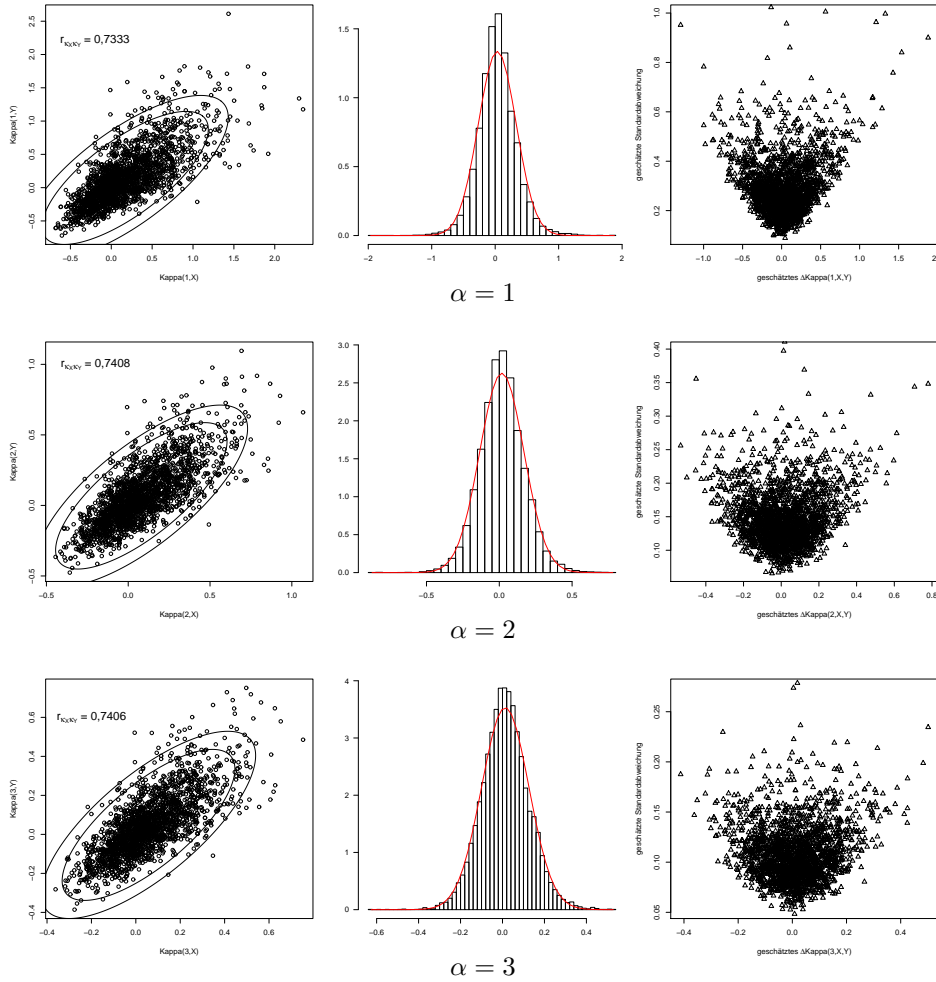
### 5.3.6 Expected Return on Value-at-Risk

Für die Varianz des Schätzers des Expected Return on Value-at-Risk wurde in Kapitel 4 folgende Formel entwickelt (vgl. S. 93):

$$n\sigma_{\widehat{ERVaR}} = \frac{\mu_X^2}{x_\alpha^2} \left( \frac{\sigma_X^2}{\mu_X^2} - \frac{E(|X - x_\alpha|) + (2\alpha - 1)(\mu_X - x_\alpha)}{\mu_X x_\alpha f(x_\alpha)} + \frac{\alpha(1 - \alpha)}{(x_\alpha f(x_\alpha))^2} \right)$$

Bei der Schätzung dieser Varianz sind die Werte  $x_\alpha$  und  $f(x_\alpha)$  geeignet zu schätzen. Ersteres wurde bereits im Abschnitt 3.3 diskutiert und für  $\hat{f}(x_\alpha)$  wurden in Abschnitt 4.1.2.6 zwei Verfahren vorgeschlagen. Erstens ein parametrisches

hier auf diese Berechnungen verzichtet und die daraus resultierenden Angaben ( $\sigma_0$ ,  $\chi^2$ -Test, 1,96 $\sigma$ -Bereiche) entfallen.



Links: Streudiagramme  $\widehat{\kappa}_X(\alpha, 0)$  (Abszisse) gegen  $\widehat{\kappa}_Y(\alpha, 0)$  (Ordinate), 2 000 Werte

Mitte: Histogramme, 20 000 Werte

Rechts: Punktdiagramme  $\widehat{\Delta}_{\kappa_X}(\alpha, z=0)$  gegen  $\widehat{\sigma}_{\widehat{\Delta}_{\kappa_X}(\alpha, z=0)}$  (2 000 Werte, rechts).

Abbildung 5.23: Grafiken für  $\Delta_{\kappa_X}(\alpha)$  bei verschiedenen Werten für  $\alpha$  mit i.i.d- $t_{10}$ -verteilten Daten und Zeitreihenlänge  $n = 60$

Verteilung	$\sigma_0(\hat{\theta})$	$\bar{s}(\hat{\theta}_i)$	$p_{\sigma^2=\sigma_0^2}$	$[u_{0,05}]$	$t_i$ in $u$	$KI_{95\%}$ um $\theta$
<b>NV</b>	$\theta = 0,0123$ ; $\bar{t} = 0,0201$ ; $s(t_i) = 0,0905$ ; $p_{AD} < 10^{-99}$ ; $\hat{\psi} = 0,0732$					
param.	0,0805	0,0910	$< 10^{-99}$	[-0,1454;0,1700]	92,44%	95,99%
nichtpar.	0,0805	0,0923	$< 10^{-99}$	[-0,1454;0,1700]	92,44%	95,68%
<b>t<sub>10</sub></b>	$\theta = 0,0125$ ; $\bar{t} = 0,0238$ ; $s(t_i) = 0,0937$ ; $p_{AD} < 10^{-99}$ ; $\hat{\psi} = 0,0902$					
param.	0,0818	0,0946	$< 10^{-99}$	[-0,1479;0,1729]	92,13%	96,00%
nichtpar.	0,0818	0,0965	$< 10^{-99}$	[-0,1479;0,1729]	92,13%	95,52%
<b>t<sub>5</sub></b>	$\theta = 0,0130$ ; $\bar{t} = 0,0260$ ; $s(t_i) = 0,0982$ ; $p_{AD} < 10^{-99}$ ; $\hat{\psi} = 0,1000$					
param.	0,0853	0,1007	$< 10^{-99}$	[-0,1542;0,1802]	91,84%	95,71%
nichtpar.	0,0853	0,1028	$< 10^{-99}$	[-0,1542;0,1802]	91,84%	95,09%
<b>GARCH</b>	$\theta = 0,0190^*$ ; $\bar{t} = 0,0369$ ; $s(t_i) = 0,1069$ ; $p_{AD} < 10^{-99}$ ; $\hat{\psi} = 0,1252$					
param.	0,1129*	0,1140	$< 10^{-99}$	[-0,2023;0,2404]	95,53%	95,49%
nichtpar.	0,1129*	0,1210	$< 10^{-99}$	[-0,2023;0,2404]	95,53%	94,56%
* aus den Daten simulierte Werte						
Erläuterung der Kürzel für die Statistiken in Tab. 5.2 auf S. 114.						

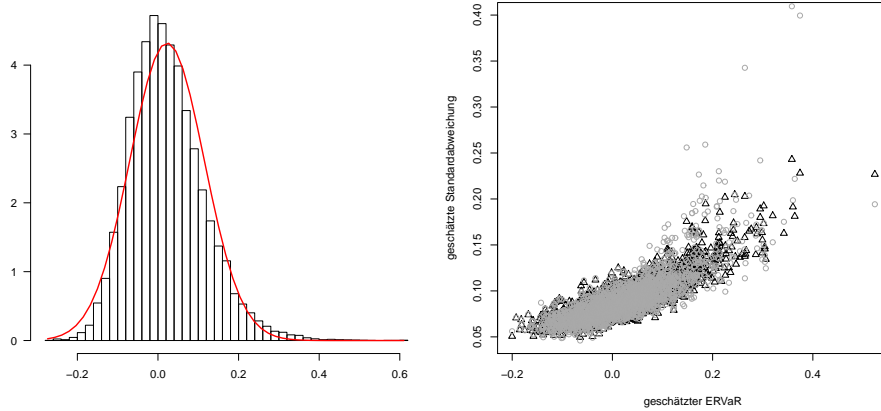
Tabelle 5.20: Ergebnisse für ERVaR( $\alpha = 0,05$ ) bei Zeitreihenlänge  $n = 60$ , Ausgansparameter

Verfahren, bei dem eine  $t_\nu$ -Verteilung angepasst und anschließend der Wert von deren Dichtefunktion am Punkt  $\hat{x}_\alpha$  abgelesen wird. Dieses Verfahren geht davon aus, dass die  $t$ -Verteilung ein einigermaßen geeignetes Verteilungsmodell für die (unbedingte) Renditeverteilung ist. Im zweiten, „nichtparametrischen“ Vorschlag wird versucht, den Wert der Dichtefunktion am Punkt  $\hat{x}_\alpha$  über den Zuwachs der empirischen Verteilungsfunktion in der Umgebung von  $\hat{x}_\alpha$  zu ermitteln. Entsprechend werden in den folgenden Abbildungen und Tabellen eine „parametrische“ und eine „nichtparametrische“ Version unterschieden, die sich aber ausschließlich auf die Schätzungen von Varianzen aus den Daten beziehen. Die aus den Parametern (und den jeweiligen Verteilungsmodellen) berechnete asymptotische Standardabweichung ist eindeutig.

In Tabelle 5.20 sehen wir die Ergebnisse für die Parameter  $\alpha = 0,05$ ,  $\mu_X = 0,002$ ,  $\sigma_X = 0,1$  und  $n = 60$ . Die Verzerrung des Schätzers fällt sofort auf. Das Mittel der Schätzwerte liegt bei allen vier Verteilungsmodellen deutlich über dem theoretischen Wert. Das liegt daran, das Mittelwert und empirisches Quantil positiv korreliert sind. Im Schätzer  $\widehat{ERVaR}_X = \bar{X}/(-\hat{x}_\alpha)$  gehen infolgedessen hohe Werte von  $\bar{X}$  im Zähler tendenziell mit niedrigen Werten ( $-\hat{x}_\alpha$ ) im Nenner einher und umgekehrt, was den Mittelwert der ERVaR-Schätzungen nach oben verzerrt. Auch deshalb weicht die Verteilung von  $\widehat{ERVaR}$  (bei  $n = 60$ ) deutlich von der Normalverteilung ab, siehe auch Abb. 5.24a mit erkennbarer Rechtsschiefe. In Abb. 5.24b ist zudem eine deutliche positive Korrelation des Schätzwerts und der geschätzten Standardabweichung erkennbar, wobei die nichtparametrische Varianzschätzung gelegentlich nach oben ausreißt.<sup>35</sup> Die em-

<sup>35</sup>Führt man die Simulation mit deutlich längeren Zeitreihen durch, sind Schätzwert und Standardabweichung bei parametrischer Varianzschätzung fast perfekt korreliert. Bei Verwendung des Schätzers mit nichtparametrischer Dichteschätzung sind die geschätzten Varianzen weiterhin erkennbar verwaschen.



(a) Histogramm  $\widehat{ERVaR}_X$ (b)  $\widehat{ERVaR}_X$ - $\widehat{\sigma}_{\widehat{ERVaR}_X}$ -Diagramm

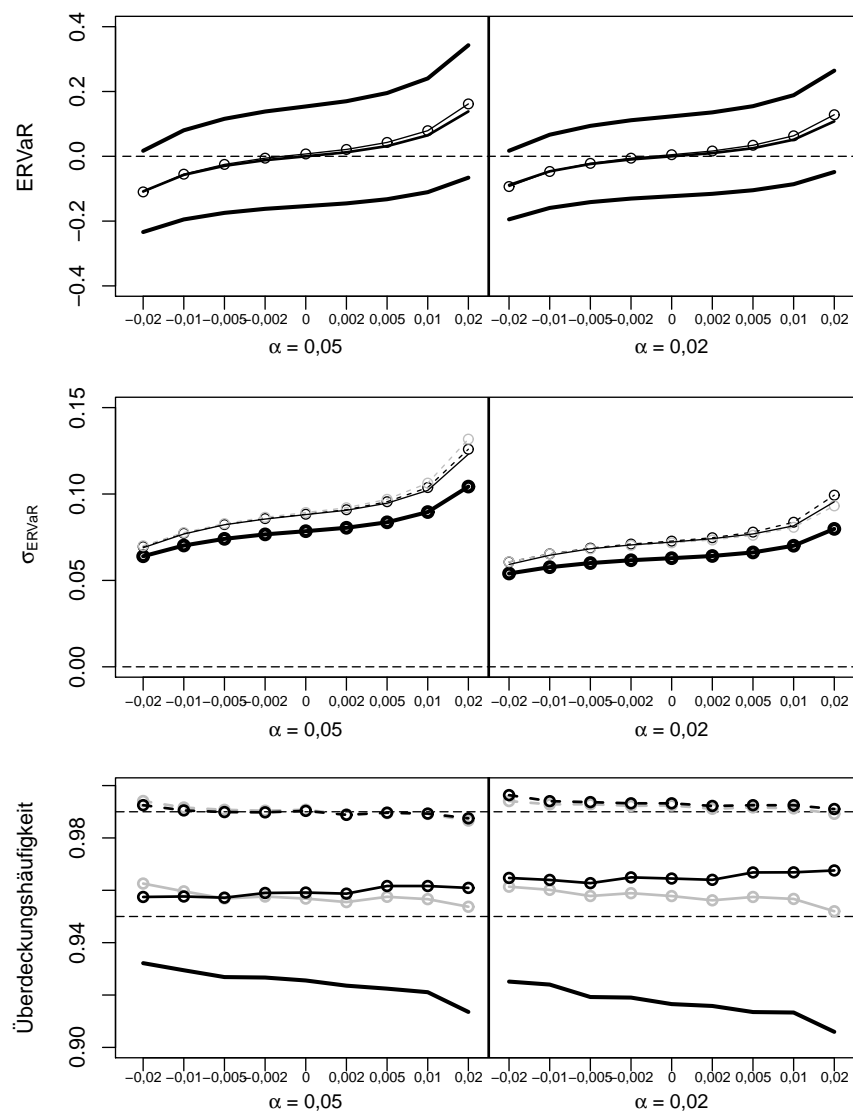
Links: Histogramm von  $\widehat{ERVaR}_X$  mit angepasster Dichtefunktion der Normalverteilung

Rechts: Punktdiagramm für 2000 Zeitreihen mit  $\widehat{ERVaR}_X$  (Abszisse) und zugehörigem  $\widehat{\sigma}_{\widehat{ERVaR}_X}$  (Ordinate) für die nichtparametrische ( $\circ$ ) und die parametrische Varianzformel ( $\triangle$ ).

Abbildung 5.24: Grafiken zur Verteilung von  $\widehat{ERVaR}_X$  für i.i.d.- $t_{10}$ -verteilte Daten mit Zeitreihenlänge  $n = 60$ .

pirische Standardabweichung liegt deutlich über der asymptotischen Standardabweichung, so dass die 1,96 $\sigma$ -Bereiche deutlich weniger als 95% der Schätzwerte enthalten. Die Schätzungen der Standardabweichung liegt dagegen im Mittel recht nahe an der empirischen Standardabweichung. Auch die Konfidenzintervalle halten das vorgegebene Niveau einigermaßen genau ein.

In Abb. 5.25 schließlich werden die Ergebnisse für verschiedene Werte von  $\mu_X$  bei i.i.d.-normalverteilten Renditen mit  $\alpha = 5\%$  bzw.  $\alpha = 2\%$  dargestellt. Demnach steigt die Aufwärtsverzerrung des Performancemaßschätzers in absoluten Zahlen mit höherer erwarteter Rendite. Die Standardabweichung des Schätzers steigt mit  $\mu_X$  leicht an. Die aus den Daten geschätzten Standardabweichungen liegen im Mittel sehr nahe an der empirischen Standardabweichung der Schätzwerte, wobei sich die beiden Verfahren kaum unterscheiden. Die asymptotische Standardabweichung ist stets erkennbar niedriger als die empirische Standardabweichung. Entsprechend enthalten die 1,96- $\sigma$ -Bereiche stets deutlich weniger als 95% der Schätzwerte. Die 95%-Konfidenzintervalle überdecken den wahren Wert etwas häufiger als ihrem theoretischen Niveau entsprechend, wobei die Abweichung bei der nichtparametrischen Formel etwas geringer ist. Die Überdeckungshäufigkeiten der 99%-Konfidenzintervalle liegen recht nahe an 99%. Insgesamt hängen die statistischen Eigenschaften von  $\widehat{ERVaR}$  nur schwach von Verteilungsparametern ab.



Verteilungsstatistiken bei Variation des Simulationsparameters  $\mu_X$  bei ERVaR mit  $\alpha = 0,05$  und  $\alpha = 0,02$ .

Daten: i.i.d.-normalverteilte Zeitreihen,  $n = 60$ , je 50 000 Simulationsläufe.

*oben:* Theoretischer Wert  $ERVaR_\alpha$  und Mittelwert der Schätzungen (mit Datenpunkten); umliegend die theoretische  $1,96\sigma$ -Umgebung.

*Mitte:* Empirische Standardabweichung der Schätzer (durchgehende, dünne Linie); asymptotische Standardabweichung (durchgehend fett, mit Datenpunkten) und mittlere geschätzte Standardabweichung (gestrichelt) nach parameterischer Formel (schwarz) und nichtparameterischer Formel (grau).

*unten:* Anteil der Schätzwerte in  $1,96\sigma$ -Umgebung (Linie ohne Datenpunkte), Überdeckungshäufigkeit des 95%-KI (durchgehende Linien mit Datenpunkten) und des 99%-KI (gestrichelte Linien mit Datenpunkten) bei Nutzung der parameterischen Formel (schwarz) bzw. der nichtparameterischen Formel (grau).

Abbildung 5.25: Grafiken zu  $\widehat{ERVaR}_X(\alpha)$  bei Variation von  $\mu_X$

## 5.4 Konfidenzintervalle

Im Rahmen des vorhergehenden Abschnitts wurden bereits Konfidenzintervalle des Typs

$$KI_{1-\alpha} = \hat{\theta} \pm u_{1-\frac{\alpha}{2}} \cdot \hat{\sigma}_{\hat{\theta}}$$

geschätzt und es wurden relative Häufigkeiten der Überdeckung des wahren Wertes  $\theta$  durch die Konfidenzintervalle ermittelt. Das geschah in erster Linie zur Diagnose der Eigenschaften der Schätzer  $\hat{\theta}$  und  $\hat{\sigma}_{\hat{\theta}}$ . In diesem Abschnitt werden die in Abschnitt 4.2.2 vorgestellten Berechnungsansätze für Konfidenzintervalle auf Basis des nichtparametrischen Bootstraps zum Vergleich herangezogen. Für jedes Performancemaß werden Konfidenzintervalle mit allen genannten Methoden berechnet und hinsichtlich ihrer Überdeckungshäufigkeit verglichen.

In der Regel werden sieben Varianten von Konfidenzintervallen berechnet (Kürzel in Klammern):

- $\hat{\sigma}_{\hat{\theta}}$ -Konfidenzintervalle: Die im Vorabschnitt verwendeten Konfidenzintervalle mit Normalverteilungsquantil und aus den Daten geschätzter Standardabweichung des Performancemaßschätzers:
  - $\hat{\sigma}_{\hat{\theta}}$ -Konfidenzintervall mit Varianzformel für unabhängig und identisch verteilte Renditen ( $\hat{\sigma}_{\hat{\theta}}$ -i.i.d.-Konfidenzintervall, SD-DF),
  - $\hat{\sigma}_{\hat{\theta}}$ -Konfidenzintervall mit vereinfachter Varianzformel für unabhängig und identisch normalverteilte Renditen ( $\hat{\sigma}_{\hat{\theta}}$ -i.i.d.-NV-Konfidenzintervall, SD-NV)<sup>36</sup>,
- Bootstrap-Konfidenzintervalle (vgl. Abschnitt 4.2.2):
  - Bootstrap- $t$ -Konfidenzintervall ohne Bias-Korrektur (TKI),
  - Bootstrap- $t$ -Konfidenzintervall mit Bias-Korrektur (TBias),
  - Studentisiertes Bootstrap-Konfidenzintervall (Student),
  - Bootstrap-Perzentil-Konfidenzintervall (Perzent),
  - $BC_{\alpha}$ -Konfidenzintervall (BCA).

Dabei wurde wiederum darauf geachtet, für alle Berechnungsansätze soweit wie möglich dieselben Daten und Zwischenergebnisse zu verwenden. Für ein Maß und eine Parameterkombination wurden alle Konfidenzintervalle aus denselben Simulationsdaten geschätzt. Bei den Bootstrap-Konfidenzintervallen war dieselbe Bootstrap-Simulationsreihe Datengrundlage für alle Berechnungsvarianten. Hierdurch sollen die zufälligen Effekte bei der Daten- bzw. der Bootstrap-Simulation bei allen Berechnungsvarianten in dieselbe Richtung wirken, so dass die Ergebnisse besser vergleichbar sind.

Aufgrund der wesentlich höheren Rechenzeitanforderungen bei Bootstrapsimulationen wurde je Performancemaß meist nur eine Parameterkombination mit

<sup>36</sup>Für Kappa wurde keine vereinfachte Formel hergeleitet, bei ERVaR beziehen sich die beiden Varianten der Konfidenzintervalle auf unterschiedliche Schätzmethoden im Nenner der Varianzformel.

im Vergleich zum Vorabschnitt reduzierter Anzahl von Simulationsläufen untersucht.

Die wichtigste Statistik zur Bewertung der Qualität der Konfidenzintervalle ist die relative Häufigkeit der Überdeckung des wahren Parameterwerts durch das Konfidenzintervall. Diese relative Überdeckungshäufigkeit  $\pi_{1-\alpha}$  sollte möglichst nahe am vorgegebenen Niveau des Konfidenzintervalls  $1 - \alpha$  liegen. Die Anpassungsgüte kann man mit einem approximativen Gauß-Anteilswerttest testen. Die Hypothesen lauten

$$H_0 : \pi_{1-\alpha} = \pi_0 (= 1 - \alpha) \quad \text{gegen} \quad H_1 : \pi_{1-\alpha} \neq \pi_0.$$

Sei  $\hat{\pi}_{1-\alpha}$  der Anteil der Konfidenzintervalle in einer einfachen Stichprobe der Länge  $R$ , die den wahren Wert umschließen. Dann gilt approximativ<sup>37</sup>

$$\frac{\hat{\pi}_{1-\alpha} - \pi_0}{\sqrt{\frac{\pi_0(1-\pi_0)}{R}}} \sim N(0, 1).$$

Tabelle 5.21 enthält eine Übersicht über Annahmebereiche des Gauß-Anteilwerttests für die Wahrscheinlichkeiten  $\alpha_{KI} = 5\%$  und  $\alpha_{KI} = 1\%$  (entsprechend 95%- und 99%-Konfidenzintervallen) bei verschiedenen Stichprobengrößen  $R$  und Fehlerwahrscheinlichkeiten 1. Art  $\alpha_{Test}$ .

$\alpha_{Test}$	$\alpha_{KI} = 5\%$			
	0,1	0,05	0,01	0,001
$R = 5\,000$	0,9449 - 0,9551	0,9440 - 0,9560	0,9421 - 0,9579	0,9399 - 0,9601
$R = 10\,000$	0,9464 - 0,9536	0,9457 - 0,9543	0,9444 - 0,9556	0,9428 - 0,9572
$R = 20\,000$	0,9475 - 0,9525	0,9470 - 0,9530	0,9460 - 0,9540	0,9449 - 0,9551
$R = 50\,000$	0,9484 - 0,9516	0,9481 - 0,9519	0,9475 - 0,9525	0,9468 - 0,9532
$R = 100\,000$	0,9489 - 0,9511	0,9486 - 0,9514	0,9482 - 0,9518	0,9477 - 0,9523
$\alpha_{Test}$	$\alpha_{KI} = 1\%$			
	0,1	0,05	0,01	0,001
$R = 5\,000$	0,9877 - 0,9923	0,9872 - 0,9928	0,9864 - 0,9936	0,9854 - 0,9946
$R = 10\,000$	0,9884 - 0,9916	0,9880 - 0,9920	0,9874 - 0,9926	0,9867 - 0,9933
$R = 20\,000$	0,9888 - 0,9912	0,9886 - 0,9914	0,9882 - 0,9918	0,9877 - 0,9923
$R = 50\,000$	0,9893 - 0,9907	0,9891 - 0,9909	0,9889 - 0,9911	0,9885 - 0,9915
$R = 100\,000$	0,9895 - 0,9905	0,9894 - 0,9906	0,9892 - 0,9908	0,9890 - 0,9910

Tabelle 5.21: Annahmebereiche von  $H_0$  bei zweiseitigem Gauß-Anteilswerttest

Eine weitere wünschenswerte Eigenschaft von Konfidenzintervallen ist die Überdeckungssymmetrie, also dass mittels des Konfidenzintervalls  $[\hat{\theta}_-(\alpha); \hat{\theta}_+(\alpha)]$  der wahre Wert mit gleicher Wahrscheinlichkeit über- und unterschätzt wird.

<sup>37</sup>Zur Kontrolle der ausreichenden Approximation dieses Gauß-Tests an einen exakten Binomialtest findet man in der Literatur die Faustformel  $R\pi_0(1-\pi_0) > 9$ , vgl. z.B. Sachs und Hedderich (2006, S. 349). Diese ist für 99%-Konfidenzintervalle bereits ab einer Stichprobenlänge  $R$  von  $9/(0,01 \cdot 0,99) \approx 910$  erfüllt.

Diese Eigenschaft wird mit dem Anteil der überhalb des wahren Wertes liegenden Konfidenzintervalle an allen den wahren Wert nicht enthaltenden Konfidenzintervallen

$$q_{\theta < KI} = \frac{\#(i|\theta < \hat{\theta}_{i-}(\alpha))}{\#(i|\theta \notin KI_i(\alpha))} = \frac{\#(i|\theta < \hat{\theta}_{i-}(\alpha))}{\#(i|\theta < \hat{\theta}_{i-}(\alpha)) + \#(i|\theta > \hat{\theta}_{i+}(\alpha))} \quad (5.2)$$

überprüft. Hierbei zählt der Ausdruck  $\#(i|B)$  die Anzahl der Realisationen  $i$ , welche die Bedingung  $B$  erfüllen. Deutliche Abweichungen von 50 % weisen auf eine Asymmetrie bzw. Verzerrung des Konfidenzintervalls hin.

Wenn nicht anders erwähnt, werden jene Parameterkombinationen zur Simulation benutzt, die im Vorabschnitt als Ausgangsparameterkombination benutzt werden:

- Für eindimensionale Zeitreihen( $X$ ): unabhängig und identisch  $t_{10}$ -verteilte Zeitreihen mit Länge  $n = 60$ , Überrendite  $\mu_X = 0,002$  und Standardabweichung der Rendite  $\sigma_X = 0,1$ .
- Für alle mehrdimensionalen Zeitreihen  $((X, Y); (M, X); (M, X, Y))$ : unabhängig und identisch multivariat  $t_{10}$ -verteilte Zeitreihen der Länge  $n = 60$  mit den Parametern aus folgender Tabelle (bzw. entsprechenden Teilmengen davon):

Variable	$\mu$	$\sigma$	Korrelationsmatrix		
$M$	0,002	0,1	1	0,75	0,75
$X$	0,003	0,1	0,75	1	0,75
$Y$	0,002	0,1	0,75	0,75	1

Die Anzahl der Simulationsläufe  $R$  wurde nach Rechenzeitgesichtspunkten unterschiedlich gewählt und wird jeweils genannt. Die Anzahl der Bootstrap-Ziehungen je Simulationslauf  $B$  beträgt in der Regel 2 000, wird aber ebenfalls stets mit angegeben.

Beginnen wir wieder mit den Ergebnissen für das **Sharpe-Ratio** eines Portfolios  $X$  (Tabelle 5.22) mit Zeitreihen der Länge  $n = 60$ . Aufgrund der schnellen Berechenbarkeit der Schätzwerte für das Sharpe-Ratio und dessen Varianz konnte hier mit  $R = 100\,000$  Zeitreihen gerechnet werden. Keines der hier getesteten Konfidenzintervalle weicht sehr weit vom vorgegebenen Niveau ab, aber durch die hohe Zahl simulierter Zeitreihen fallen im Anteilswerttest schon relativ kleine Abweichungen der Überdeckungshäufigkeiten vom Zielniveau auf. Die erste Zahl links oben im Datenbereich der Tabelle sagt z.B. aus, dass das geschätzte  $\hat{\sigma}_{\hat{\theta}}$ -i.i.d.-NV-95 %-Konfidenzintervall bei 95,23 % der 100 000 Zeitreihen den wahren Wert einschloss. Das ist etwas mehr als die vorgegebenen 95 % und liegt (vgl. Tab. 5.21) gerade noch im Annahmehereich von  $H_0$  eines Anteilswerttests mit  $\alpha_{Test} = 0,1\%$ , aber außerhalb des Annahmehereiches eines Tests mit  $\alpha_{Test} = 1\%$ . Die folgende eingeklammerte Zahl (51 %) ist  $q_{\theta < KI}$ , das heißt also, dass in 51 % der Fälle, in denen das Konfidenzintervall den wahren

$n = 60, B = 2000$				
	$Sh_X, R = 100\,000$		$\Delta Sh_{XY}, R = 10\,000$	
KI-Niveau	95 %	99 %	95 %	99 %
SD-NV	95,23 %** (51%)	99,17 %*** (50%)	94,72 % (52%)	99,00 % (51%)
SD-DF	94,63 %*** (52%)	98,90 %** (51%)	94,65 % (51%)	98,89 % (50%)
TKI	94,86 % (51%)	98,87 %*** (51%)	95,17 % (52%)	98,95 % (53%)
TBias	95,19 %** (51%)	99,00 % (49%)	95,54 %* (50%)	99,05 % (54%)
Student	95,33 %*** (51%)	99,14 %*** (50%)	95,25 % (50%)	99,04 % (52%)
Perzent	94,18 %*** (52%)	98,60 %*** (52%)	94,19 %*** (52%)	98,54 %*** (54%)
BCA	95,37 %*** (51%)	99,14 %*** (50%)	94,60 % (52%)	98,65 %*** (54%)

Zweiseitige Anteilwerttests auf Einhaltung des KI-Niveaus, Ablehnung von  $H_0$ :  
\* zum Niveau  $\alpha = 5\%$ ; \*\* zum Niveau  $\alpha = 1\%$ ; \*\*\* zum Niveau  $\alpha = 0,1\%$   
In Klammern:  $q_{\theta < KI}$  (vgl. Formel 5.2, S. 169)

Tabelle 5.22: Überdeckungshäufigkeiten verschiedener Konfidenzintervalle für  $Sh_X$  und  $\Delta Sh_{XY}$

Wert nicht einschloss, das Konfidenzintervall überhalb des wahren Wertes lag – und entsprechend in  $(1 - q_{\theta < KI}) = 49\%$  unterhalb des wahren Wertes. Abweichungen nach unten und oben kommen also in etwa gleich häufig vor, was eine wünschenswerte Eigenschaft darstellt (Überdeckungssymmetrie). Der erste Wert in der Zeile darunter gibt an, dass 94,63 % der  $\hat{\sigma}_{\hat{\theta}}$ -i.i.d.-Konfidenzintervalle (Formel ohne Normalverteilungsannahme) den wahren Wert einschlossen. Dieser Wert ist zum 0,1%-Niveau nicht mit einer theoretischen Überdeckungshäufigkeit von 95 % vereinbar. Dieses Ergebnis, dass Konfidenzintervalle auf Basis der Varianzformel mit NV-Annahme etwas konservativ sind und jene auf Basis der Varianzformel ohne Verteilungsannahme etwas zu liberal, kennen wir schon aus der vergleichbaren Tabelle im Vorabschnitt (Tab. 5.9, S. 124, erste beiden Zahlen der letzten Spalte).

Unter den Bootstrap-Konfidenzintervallen liegen bei einem Konfidenzniveau von 95 % die Überdeckungshäufigkeiten der Bootstrap- $t$ -Intervalle mit und ohne Biaskorrektur relativ nahe am vorgegebenen Niveau. Das studentisierte Bootstrap-Konfidenzintervall und das BCA-Konfidenzintervall sind etwas zu konservativ, während das Perzentil-Konfidenzintervall deutlich zu liberal ist. Alle Konfidenzintervalle sind nahezu überdeckungssymmetrisch. Alle diese Beobachtungen findet man in der nächsten Spalte (99%-Konfidenzintervalle) fast identisch wieder. Die Untersuchung für die Differenz des Sharpe-Ratios zweier Portfolios  $X$  und  $Y$  wurde aus Rechenzeitgründen nur mit 10 000 Zeitreihen durchgeführt. Deshalb sind die Zahlen etwas unsicherer und der Anteilwerttest hat weniger Power. Auffällig ist lediglich, dass das Perzentil-Konfidenzintervall schwächer abschneidet. Auch die BCA-Methode ist für das 99%-Konfidenzintervall nachweislich zu liberal.

$n = 60, B = 2000$				
	$Tr_X, R = 20000$		$\Delta Tr_{XY}, R = 10000$	
KI-Niveau	95 %	99 %	95 %	99 %
SD-NV	95,59 %*** (50%)	99,41 %*** (47%)	95,08 % (47%)	99,22 %* (58%)
SD-DF	95,45 %** (50%)	99,28 %*** (53%)	95,17 % (46%)	99,29 %** (59%)
TKI	95,52 %*** (51%)	99,25 %*** (53%)	95,76 %*** (46%)	99,30 %** (57%)
TBias	95,75 %*** (50%)	99,31 %*** (50%)	95,88 %*** (46%)	99,32 %** (56%)
Student	95,19 % (51%)	99,22 %** (52%)	94,90 % (47%)	99,15 % (55%)
Perzent	95,20 % (52%)	98,66 %*** (56%)	94,21 %*** (47%)	98,83 %* (56%)
BCA	94,44 %*** (52%)	98,88 % (55%)	94,05 %*** (47%)	98,61 %*** (55%)

Zweiseitige Anteilswerttests auf Einhaltung des KI-Niveaus, Ablehnung von  $H_0$ :  
\* zum Niveau  $\alpha = 5\%$ ; \*\* zum Niveau  $\alpha = 1\%$ ; \*\*\* zum Niveau  $\alpha = 0,1\%$   
In Klammern:  $q_{\theta < KI}$  (vgl. Formel 5.2, S. 169)

Tabelle 5.23: Überdeckungshäufigkeiten verschiedener Konfidenzintervalle für  $Tr_X$  und  $\Delta Tr_{XY}$

Für das **Treynor-Ratio** (siehe Tab. 5.23) liegen ebenfalls alle Überdeckungshäufigkeiten innerhalb eines Prozentpunkts um das theoretische Niveau 95 % bzw. innerhalb eines halben Prozentpunkts um 99 %. Damit wich keine der Methoden grob ab, aber zuverlässige Ergebnisse im Sinne des Hypothesentests konnte auch keine Methode liefern. Für das 95%-Konfidenzintervall lagen nur die Überdeckungshäufigkeiten des studentisierten Bootstrap-Intervalls und des Perzentil-Intervalls innerhalb des Annahmebereiches eines Anteilswerttests zum Niveau 1 %. Für das 99%-Konfidenzintervall galt dies nur für das BCA-Intervall.

Bei der Differenz der Treynor-Ratios zweier Portfolios schnitten die Nicht-Bootstrap-Intervalle zufriedenstellend ab, vor allem bei den 95%-Konfidenzintervallen. Die eindeutig besten Ergebnisse lieferte jedoch der studentisierte Bootstrap. Hier unterschieden sich beide Überdeckungshäufigkeiten auf 5%-Testniveau nicht signifikant vom theoretischen KI-Niveau.

An dieser Stelle kann man die Frage stellen, wie sich die verschiedenen Konfidenzintervall-Methoden in jenen Situationen schlagen, für die in Abschnitt 5.3.2 festgestellt wurde, dass sie für  $n = 60$  noch ungenügend konvergieren (vgl. Abb. 5.9, S. 138), insbesondere bei schwacher Korrelation zwischen Portfolio und Marktindex. Dazu wurde die Simulation noch einmal mit dem abweichenden Parameterwert  $\rho_{M,X} = 0,25$  (statt 0,75) durchgeführt. Die Ergebnisse (mit  $R = 10000$ ) stehen in Tabelle 5.24.

Wie zu erwarten war, erweisen sich die beiden auf Basis der Stichprobenmomente errechneten Konfidenzintervalle als unbrauchbar. In ähnlicher Weise gilt dies auch für die beiden Konfidenzintervalle auf Basis von Bootstrap-Standardabweichung und  $t$ -Verteilung. Auch das Perzentil-Intervall war weitaus zu konservativ. Lediglich das studentisierte Bootstrap-Intervall lieferte relativ exakte Überdeckungshäufigkeiten, wobei dieses Intervall nicht ganz überde-

$n = 60, B = 2000, \rho_{M,X} = 0,25$		
$Tr_X, R = 10\,000$		
KI-Niveau	95 %	99 %
SD-NV	99,32 % <sup>***</sup> (16%)	99,99 % <sup>***</sup> (0%)
SD-DF	99,10 % <sup>***</sup> (21%)	99,93 % <sup>***</sup> (0%)
TKI	99,79 % <sup>***</sup> (19%)	99,91 % <sup>***</sup> (33%)
TBias	99,67 % <sup>***</sup> (24%)	99,84 % <sup>***</sup> (31%)
Student	94,42 % <sup>**</sup> (35%)	99,13 % (39%)
Perzent	98,13 % <sup>***</sup> (35%)	99,78 % <sup>***</sup> (41%)
BCA	– (teilweise Abbruch der Berechnungen)	
Zweiseitige Anteilwerttests auf Einhaltung des KI-Niveaus, Ablehnung von $H_0$ :		
* zum Niveau $\alpha = 5\%$ ; ** zum Niveau $\alpha = 1\%$ ; *** zum Niveau $\alpha = 0,1\%$		
In Klammern: $q_{\theta < KI}$ (vgl. Formel 5.2, S. 169)		

Tabelle 5.24: Überdeckungshäufigkeiten verschiedener Konfidenzintervalle für  $Tr_X$  bei  $\rho_{M,X} = 0,25$

ckungssymmetrisch ist. In jeweils über 60 % aller Fehlschätzungen lag das Konfidenzintervall unter dem wahren Wert. Bei der Berechnung des BCA-Intervalls gab es bei manchen Zeitreihen Probleme bei der Berechnung der Quantile, was zum Programmabbruch führte. Deshalb liegt für diese Methode kein Wert für die Überdeckungshäufigkeit vor.

Im Vergleich der Konfidenzintervalle für den Schätzer von **Jensens Alpha** schnitten die Nicht-Bootstrap-Konfidenzintervalle am besten ab (Tab. 5.25). Die Tests lehnten die Nullhypothese, dass das theoretische Niveau eingehalten wird, bei einer Fehlerwahrscheinlichkeit 1. Art von 5% in allen vier Fällen nicht ab. Die Bootstrap-Konfidenzintervalle verfehlten das theoretische Niveau jeweils mehr oder weniger deutlich. Vor allem beim BCA-Konfidenzintervall sind die Abweichungen mit  $-1,07$  bzw  $-0,53$  Prozentpunkten erheblich.

Auch bei der Differenz der Alphas zweier Portfolios gab es bei den  $\hat{\sigma}_{\hat{\theta}}$ -Konfidenzintervallen die geringsten Abweichungen. Die Überdeckungshäufigkeiten der Bootstrap-Konfidenzintervalle bei den 95%-Konfidenzintervallen sind etwa einen Prozentpunkt zu niedrig, beim BCA-Konfidenzintervall sogar gut 1,5 Prozentpunkte.

Für das Performancemaß nach Modigliani/Modigliani (**MM**) unterscheiden wir wieder die beiden Versionen  $MM_{ori}$  und  $MM_{alt}$ , siehe Tabelle 5.26.

Für  $MM_{ori}$  sind die Ergebnisse der Bootstrap- $t$ -Intervalle am besten. Ihre Überdeckungshäufigkeiten unterscheiden sich nicht signifikant vom den vorgegebenen Niveaus. Auch das 95 %- $\hat{\sigma}_{\hat{\theta}}$ -Konfidenzintervall bei Normalverteilungsannahme weicht nur geringfügig ab. Die Abweichung beim 99%-Konfidenzintervall ist zwar



$n = 60, B = 2000$				
	$\alpha_X, R = 20\,000$		$\Delta\alpha_{XY}, R = 10\,000$	
KI-Niveau	95 %	99 %	95 %	99 %
SD-NV	94,80 % <sup>(49%)</sup>	98,91 % <sup>(53%)</sup>	94,49 % <sup>** (49%)</sup>	98,94 % <sup>(49%)</sup>
SD-DF	94,91 % <sup>(49%)</sup>	98,99 % <sup>(51%)</sup>	94,54 % <sup>* (50%)</sup>	98,97 % <sup>(51%)</sup>
TKI	94,44 % <sup>*** (49%)</sup>	98,77 % <sup>** (52%)</sup>	93,92 % <sup>*** (48%)</sup>	98,67 % <sup>** (49%)</sup>
TBias	94,31 % <sup>*** (49%)</sup>	98,75 % <sup>*** (50%)</sup>	93,90 % <sup>*** (48%)</sup>	98,61 % <sup>*** (50%)</sup>
Student	94,52 % <sup>** (49%)</sup>	98,85 % <sup>* (51%)</sup>	94,04 % <sup>*** (48%)</sup>	98,74 % <sup>* (48%)</sup>
Perzent	94,35 % <sup>*** (48%)</sup>	98,62 % <sup>*** (49%)</sup>	93,93 % <sup>*** (48%)</sup>	98,68 % <sup>** (52%)</sup>
BCA	93,93 % <sup>*** (48%)</sup>	98,47 % <sup>*** (48%)</sup>	93,43 % <sup>*** (48%)</sup>	98,45 % <sup>*** (52%)</sup>

Zweiseitige Anteilswerttests auf Einhaltung des KI-Niveaus, Ablehnung von  $H_0$ :  
\* zum Niveau  $\alpha = 5\%$ ; \*\* zum Niveau  $\alpha = 1\%$ ; \*\*\* zum Niveau  $\alpha = 0,1\%$   
In Klammern:  $q_{\theta < KI}$  (vgl. Formel 5.2, S. 169)

Tabelle 5.25: Überdeckungshäufigkeiten verschiedener Konfidenzintervalle für  $\alpha_X$  und  $\Delta\alpha_{XY}$

moderat, aber bereits statistisch signifikant. Bei den vier anderen Methoden gibt es deutliche Abweichungen. Die  $\hat{\sigma}_{\hat{\theta}}$ -Konfidenzintervalle auf Basis der Varianzformel ohne Verteilungsannahme sind deutlich zu konservativ. In Abschnitt 5.3.4 hatten wir bereits gesehen (vgl. Abb. 5.17b, S. 152), dass betragsmäßig große Schätzwerte für  $MM_{ori}$  mit deutlich erhöhten Schätzwerten für die Schätzvarianz einhergehen, was eine mögliche Ursache für die überhöhten Überdeckungshäufigkeiten ist. Die drei weiteren Bootstrap-Konfidenzintervalle (Student, Perzent und BCA) sind dagegen teils deutlich zu liberal.

Bei  $MM_{alt}$  sind die Ergebnisse ähnlich. Die Überdeckungshäufigkeiten der Bootstrap- $t$ -Intervalle weichen zwar für das Konfidenzniveau 95 % signifikant ab, aber die Abweichungen halten sich mit gut 0,5 Prozentpunkten im Rahmen. Auch die  $\hat{\sigma}_{\hat{\theta}}$ -Konfidenzintervalle liefern, vor allem beim 95 %-Konfidenzniveau, gute Ergebnisse. Für die drei anderen Berechnungsmethoden sind die Abweichungen nach unten deutlicher.

Bei den Konfidenzintervallen für die Differenz der  $MM$  zweier Portfolios liefern die  $\hat{\sigma}_{\hat{\theta}}$ -i.i.d.-NV-Konfidenzintervalle die besten Überdeckungshäufigkeiten. Wie bei den Ergebnissen für  $MM_{ori}$  und  $MM_{alt}$  sind die  $\hat{\sigma}_{\hat{\theta}}$ -i.i.d.-Konfidenzintervalle anhand der Formel ohne Normalverteilungssannahme etwas zu konservativ und alle Bootstrap-Intervalle in unterschiedlichem Maße zu liberal. Bei den Bootstrap- $t$ -Intervallen ist die Abweichung eher gering, beim BCA-Intervall dagegen erheblich: Das Konfidenzintervall verfehlt den wahren Wert 1,36- bzw. 1,56mal häufiger als gefordert.

Bei allen bisher untersuchten Maßen – Sharpe-Ratio, Treynor-Ratio<sup>38</sup>, Jensens Alpha und  $MM$  – lagen die Konfidenzintervalle mit allen Schätzmethoden etwa gleich häufig über wie unter dem theoretischen Wert  $\theta$ . Die Werte von  $q_{\theta < KI}$

<sup>38</sup>Mit Ausnahme der Ergebnisse in Tabelle 5.24 zum Szenario mit geringer Korrelation zwischen Portfolio und Markt

$n = 60, B = 2000$				
	$MM_{oriX}, R = 20000$		$MM_{altX}, R = 20000$	
KI-Niveau	95 %	99 %	95 %	99 %
SD-NV	95,24 % (50%)	99,24 %*** (54%)	94,83 % (49%)	99,09 % (44%)
SD-DF	96,80 %*** (48%)	99,68 %*** (49%)	95,79 %*** (47%)	99,33 %*** (41%)
TKI	94,77 % (51%)	98,92 % (54%)	94,44 %*** (49%)	98,88 % (46%)
TBias	94,82 % (51%)	98,92 % (53%)	94,41 %*** (50%)	98,84 %* (47%)
Student	93,59 %*** (49%)	98,45 %*** (53%)	94,03 %*** (50%)	98,63 %*** (46%)
Perzent	94,27 %*** (51%)	98,68 %*** (55%)	94,23 %*** (50%)	98,70 %*** (47%)
BCA	93,99 %*** (50%)	98,57 %*** (53%)	93,79 %*** (51%)	98,43 %*** (48%)
$\Delta_{MM_{XY}}, R = 10000$				
KI-Niveau	95 %		99 %	
SD-NV	94,62 % (47%)		99,10 % (53%)	
SD-DF	95,75 %*** (43%)		99,38 %*** (45%)	
TKI	94,29 %** (47%)		98,94 % (53%)	
TBias	94,21 %*** (47%)		98,97 % (52%)	
Student	94,20 %*** (47%)		98,92 % (53%)	
Perzent	93,60 %*** (48%)		98,60 %*** (51%)	
BCA	93,19 %*** (46%)		98,44 %*** (51%)	
Zweiseitige Anteilwerttests auf Einhaltung des KI-Niveaus, Ablehnung von $H_0$ : * zum Niveau $\alpha = 5\%$ ; ** zum Niveau $\alpha = 1\%$ ; *** zum Niveau $\alpha = 0,1\%$ In Klammern: $q_{\theta < KI}$ (vgl. Formel 5.2, S. 169)				

Tabelle 5.26: Überdeckungshäufigkeiten verschiedener Konfidenzintervalle für Schätzer des Performancemaßes von Modigliani-Modigliani

lagen stets innerhalb des Bereichs von 40 % bis 60 %.

Bereits in Abschnitt 5.3.5 hatten wir gesehen, dass die Verteilung der Schätzer für das verallgemeinerte Performancemaß **Kappa**( $\alpha$ ) deutlich von einer Normalverteilung abweicht (vgl. Abb. 5.21, S. 159), insbesondere für  $\alpha = 1$ . Das schlägt sich in Simulationsergebnissen (siehe Tab. 5.27) nieder. Die Überdeckungshäufigkeiten der 99%-Konfidenzintervalle bei  $\alpha = 1$  sind – bis auf Perzentil- und BCA-Intervalle – deutlich zu niedrig. Die Konfidenzintervalle verfehlen mehr als doppelt so häufig den wahren Wert wie verlangt (2,31 % bis 2,84 % statt 1 %). Außerdem kommen diese Abweichungen nicht gleichmäßig durch Über- oder Unterschätzen des wahren Wertes zustande. Der Anteil  $q_{\theta < KI}$  ist stets null, es kam also überhaupt nicht vor, dass das Konfidenzintervall vollständig über dem wahren Wert lag.<sup>39</sup> Lediglich die Perzentil- und BCA-Konfidenzintervalle sind stets annähernd überdeckungssymmetrisch. Bei den 95%-Konfidenzintervallen sind die Abweichungen deutlich weniger ausgeprägt. Für  $\kappa_{\alpha=1}$  (= Omega) sollte man gerade für höhere Konfidenzniveaus die Perzentil- oder die BCA-Methode wählen.

Bei höheren Werten für  $\alpha$  ( $\alpha = 2 \rightarrow$  Sortino-Ratio) sind die Abweichungen der Überdeckungshäufigkeiten vom vorgegebenen Konfidenzniveau zwar zumeist signifikant, bleiben aber in einem vertretbaren Rahmen. Bei  $\alpha = 3$  etwa liegen die Abweichungen innerhalb eines (95%-KI) bzw. eines halben (99%-KI) Prozentpunkts.

Auch der Schätzer für den **Excess Return on Value-at-Risk** ist (vgl. Abb. 5.24a, S. 165) für  $n = 60$  eindeutig nicht normalverteilt. Die besten Ergebnisse liefert hier das BCA-Intervall, das bei beiden Konfidenzniveaus und für beide Parameterwerte von  $\alpha$  (0,05 und 0,02) Überdeckungshäufigkeiten sehr nahe den theoretischen Niveaus aufweist. Alle anderen Methoden weisen jeweils in zumindest einem Fall zum Niveau 0,1 % signifikante Überdeckungshäufigkeiten auf. Wiederum sind nur BCA- und Perzentilintervall einigermaßen überdeckungssymmetrisch.

Die in Kapitel 4 geäußerten Bedenken<sup>40</sup> gegenüber dem Einsatz von Bootstrap-Verfahren für Statistiken, die Quantile beinhalten, bestätigen sich mit diesen Ergebnissen nicht. Die Ergebnisse einiger Bootstrap-Varianten (TKI, TBias, Student) sind zwar nicht gut, aber im Vergleich zu den Nicht-Bootstrap-Konfidenzintervallen und den Ergebnissen für andere Performancemaße keineswegs außergewöhnlich.

<sup>39</sup>Beispielsweise beinhaltet bei  $\alpha = 1$  das studentisierte 99%-KI den wahren Wert in 97,69 % der Fälle, lag nie über dem wahren Wert, aber in 2,31 % der Fälle unter dem wahren Wert.

<sup>40</sup>Siehe Fußnote 35, S. 105.

$n = 60, B = 2000, R = 20000$				
	$\kappa_X(\alpha = 1)$		$\kappa_X(\alpha = 2)$	
KI-Niveau	95 %	99 %	95 %	99 %
SD-DF	93,87 % <sup>***</sup> (1%)	97,16 % <sup>***</sup> (0%)	94,90 % <sup>(9%)</sup>	98,29 % <sup>***</sup> (0%)
TKI	94,87 % <sup>(0%)</sup>	97,59 % <sup>***</sup> (0%)	95,55 % <sup>***</sup> (3%)	98,36 % <sup>***</sup> (0%)
TBias	93,76 % <sup>***</sup> (0%)	97,19 % <sup>***</sup> (0%)	95,19 % <sup>(1%)</sup>	98,20 % <sup>***</sup> (0%)
Student	94,35 % <sup>***</sup> (1%)	97,69 % <sup>***</sup> (0%)	95,39 % <sup>*</sup> (4%)	98,43 % <sup>***</sup> (0%)
Perzent	94,57 % <sup>**</sup> (50%)	98,69 % <sup>***</sup> (54%)	94,54 % <sup>**</sup> (52%)	98,68 % <sup>***</sup> (55%)
BCA	95,44 % <sup>**</sup> (50%)	99,11 % <sup>(54%)</sup>	95,73 % <sup>**</sup> (50%)	99,20 % <sup>**</sup> (45%)
$\kappa_X(\alpha = 3)$				
KI-Niveau	95 %		99 %	
SD-DF	95,19 % <sup>(16%)</sup>		98,65 % <sup>***</sup> (1%)	
TKI	95,85 % <sup>***</sup> (8%)		98,72 % <sup>***</sup> (4%)	
TBias	95,84 % <sup>***</sup> (3%)		98,65 % <sup>***</sup> (0%)	
Student	95,80 % <sup>***</sup> (9%)		98,78 % <sup>**</sup> (0%)	
Perzent	94,46 % <sup>***</sup> (52%)		98,69 % <sup>***</sup> (54%)	
BCA	95,98 % <sup>***</sup> (52%)		99,26 % <sup>***</sup> (50%)	
Zweiseitige Anteilswerttests auf Einhaltung des KI-Niveaus, Ablehnung von $H_0$ : * zum Niveau $\alpha = 5\%$ ; ** zum Niveau $\alpha = 1\%$ ; *** zum Niveau $\alpha = 0,1\%$ In Klammern: $q_{\theta < KI}$ (vgl. Formel 5.2, S. 169)				

Tabelle 5.27: Überdeckungshäufigkeiten verschiedener Konfidenzintervalle für  $\kappa_X(\alpha = 1, 2, 3)$ 

$n = 60, B = 2000, R = 20000$				
	$ERVaR(\alpha = 0,05)$		$ERVaR(\alpha = 0,02)$	
KI-Niveau	95 %	99 %	95 %	99 %
SD-param.	96,01 % <sup>***</sup> (4%)	98,90 % <sup>(0%)</sup>	96,52 % <sup>***</sup> (5%)	99,19 % <sup>**</sup> (0%)
SD-nonpa.	95,53 % <sup>***</sup> (12%)	98,89 % <sup>(3%)</sup>	95,84 % <sup>***</sup> (23%)	99,17 % <sup>*</sup> (8%)
TKI	96,18 % <sup>***</sup> (5%)	98,83 % <sup>*</sup> (1%)	96,47 % <sup>***</sup> (11%)	99,17 % <sup>*</sup> (2%)
TBias	95,86 % <sup>***</sup> (3%)	98,72 % <sup>***</sup> (1%)	96,42 % <sup>***</sup> (7%)	99,11 % <sup>(1%)</sup>
Student	94,86 % <sup>(18%)</sup>	98,56 % <sup>***</sup> (5%)	94,94 % <sup>(23%)</sup>	98,65 % <sup>***</sup> (12%)
Perzent	94,53 % <sup>**</sup> (50%)	98,62 % <sup>***</sup> (52%)	94,59 % <sup>**</sup> (51%)	98,67 % <sup>***</sup> (52%)
BCA	94,94 % <sup>(45%)</sup>	98,88 % <sup>(42%)</sup>	95,13 % <sup>(45%)</sup>	99,03 % <sup>(32%)</sup>
Zweiseitige Anteilswerttests auf Einhaltung des KI-Niveaus, Ablehnung von $H_0$ : * zum Niveau $\alpha = 5\%$ ; ** zum Niveau $\alpha = 1\%$ ; *** zum Niveau $\alpha = 0,1\%$ In Klammern: $q_{\theta < KI}$ (vgl. Formel 5.2, S. 169)				

Tabelle 5.28: Überdeckungshäufigkeiten verschiedener Konfidenzintervalle für  $ERVaR_X(\alpha = 0,05; 0,02)$

	R	SD-NV <sup>1</sup>	SD-DF <sup>1</sup>	TKI	TBias	Student	Perzent	BCA
$Sh_X$	100 000	o	o	o	+	o	o	o
$\Delta Sh_{X,Y}$	10 000	++	++	++	++	++	-	o
$Tr_X$	20 000	-	o	o	o	+	o	o
$\Delta Tr_{X,Y}$	10 000	++	+	o	o	++	o	o
Jensens $\alpha_X$	20 000	++	++	o	o	+	o	-
$\Delta \alpha_{X,Y}$	10 000	+	++	-	-	o	-	-
$MM_{ori,X}$	20 000	o	-	++	++	-	o	-
$MM_{alt,X}$	20 000	++	o	o	o	o	o	-
$\Delta MM_{X,Y}$	10 000	++	o	+	o	o	-	-
$\kappa_{1,X}$ (Omega)	20 000	- - <sup>2</sup>		--	--	--	o	+
$\kappa_{2,X}$ (Sortino)	20 000	- <sup>2</sup>		-	-	-	o	+
$\kappa_{3,X}$	20 000	o <sup>2</sup>		o	o	o	o	o
ERVaR <sub>0,05;X</sub>	20 000	par.: -	np.: o	-	o	-	o	++
ERVaR <sub>0,02;X</sub>	20 000	par.: -	np.: o	-	-	o	o	++
<b>Legende:</b> Überdeckungshäufigkeiten im Vergleich zu Konfidenzniveau...								
++: ... nie signifikant abweichend zum Niveau 1 %								
+: ... nie signifikant abweichend zum Niveau 0,1 %								
o: ... mind. einmal signifikant abweichend zum Niveau 0,1 %								
-: Abweichung mind. einmal > 1 Prozentpunkt (95%-KI) oder > 0,4 Prozentp. (99%-KI)								
--: Abweichung mind. einmal > 2 Prozentpunkte (95%-KI) oder > 0,8 Prozentp. (99%-KI)								
<sup>1</sup> Bei ERVaR: Spalte SD-NV: SD-parametrisch, Spalte SD-DF: SD-nonparametrisch								
<sup>2</sup> keine vereinfachte Formel für normalverteilte Daten vorhanden								

Tabelle 5.29: Vergleich der Schätzmethoden für Konfidenzintervalle

In Tabelle 5.29 sind die Resultate für die einzelnen Maße zusammengefasst.<sup>41</sup> Man erkennt, dass keine der sieben Schätzmethoden eindeutig überlegen oder generell zuverlässig ist. Jede Methode liefert mindestens dreimal und bei mindestens zwei verschiedenen Performancemaßen Bewertungen von „–“ oder schlechter. In gewissem Sinne zuverlässig ist allein die einfache Bootstrap-Perzentil-Methode und zwar dahingehend, dass sie konstant schwache, aber nicht irreführend schlechte Ergebnisse hervorbringt.

Eher sollte man die Wahl der KI-Berechnungsmethode nach dem zu untersuchenden Performancemaß richten. Die Intervallschätzung für Performancemaße klassischen Zuschnitts (Sharpe, Treynor, Jensen, MM), deren Schätzer i. d. R. annähernd normalverteilt sind, lässt sich gut mit jenen Methoden bewerkstelligen, die auf geschätzten Standardabweichungen und einem Prüfverteilungs-Quantil beruhen (SD-NV, SD-DF, TKI, TBias, Student). Ihre Ergebnisse sind häufig erkennbar besser und nur selten schlechter als jene der Konfidenzintervalle auf Basis von Quantilen der Bootstrapverteilung (Perzent, BCA). Das studentisierte Bootstrap-Intervall erscheint robuster bei Schätzern mit extrem leptokurtischer Verteilung, wenn man die Ergebnisse für das Treynor-Ratio bei  $\rho_{M,X} = 0,25$  (siehe Tab. 5.24) heranzieht. Bei Downside-Risk-Performancemaßen (Kappa, ERVaR), deren Schätzer eine schiefe, von der Normalverteilung stark abweichende Verteilung haben, sind die Resultate der  $\hat{\sigma}$ -Prüfverteilungs-Konfidenzintervalle vor allem bei 99%-Konfidenzintervallen sehr schlecht. Hier ist das BCA-Konfidenzintervall die verlässlichere Alternative.

Die Ergebnisse in Tabelle 5.29 sind mit Vorsicht zu betrachten, da hier rechenzeitbedingt im Gegensatz zu den Simulationen in Abschnitt 5.3 nicht systematisch verschiedene Parameterwerte durchgeprüft wurden. Im vorigen Abschnitt konnte man sehen, dass die Verteilungseigenschaften der Schätzer, die Qualität der Varianzschätzung und die Überdeckungshäufigkeiten der SD-NV/SD-DF-Konfidenzintervalle teilweise erheblich von den Verteilungsparametern der simulierten Renditen abhängen. Auch in diesem Abschnitt sind die Ergebnisse in den Tabellen für das Treynor-Ratio (Tab. 5.23 ( $\rho_{M,X} = 0,75$ ) und Tab. 5.24 ( $\rho_{M,X} = 0,25$ )) deutlich unterschiedlich. In diesem Sinne ist die Tabelle 5.29 lediglich als zusammenfassende Übersicht über die Ergebnisse dieses Unterabschnitts zu sehen, nicht als abschließende Bewertung der Eignung bestimmter Berechnungsmethoden für die verschiedenen Performancemaße.

Zum Abschluss noch ein Blick auf die Anzahl der Bootstrap-Replikationen  $B$ . Wir haben stets mit den von Efron und Tibshirani vorgeschlagenen  $B = 2000$  Replikationen gearbeitet und nicht die Frage gestellt, ob diese Anzahl wirklich notwendig ist. Deshalb berechnen wir nun exemplarisch Überdeckungshäufigkeiten bei verschiedenen Kombinationen von  $n$  und  $B$  für das Sharpe-Ratio (Tab. 5.30). Für die Daten in der linken Hälfte der Tabelle wurden die bekannten Berechnungen auf Basis von  $R = 10000$  Zeitreihen zu  $n = 60$  durchgeführt. Die ersten beiden Datenzeilen enthalten die direkt aus der Stichprobe berechneten

<sup>41</sup>Die in Tab. 5.24 dargestellten Ergebnisse für  $\widehat{TR}_X$  bei  $\rho_{M,X} = 0,25$  wurden nicht berücksichtigt. Man beachte, dass die Annahmehereiche der Tests je nach Anzahl simulierter Zeitreihen  $R$  unterschiedlich breit sind. Bei hohem  $R$  reichen geringere Abweichungen vom theoretischen Niveau, damit „o“ anstatt „+“ oder „++“ vergeben wird.

$\hat{\sigma}_{\hat{\theta}}$ -Intervalle. In den darauffolgenden Abschnitten der Tabelle stehen die Ergebnisse für die Bootstrap-Intervalle mit  $B = 100$ ,  $B = 500$  bzw.  $B = 2000$  Replikationen. Für die Kombination ( $n = 60$ ,  $B = 100$ ) erkennt man, dass die beiden Bootstrap- $t$ -Intervalle bereits gute Ergebnisse liefern. Die Student-, Perzentil- und BCA-Intervalle sind dagegen gerade für das 99%-Konfidenzniveau deutlich schlechter. Erhöht man die Anzahl der Bootstrapreplikationen auf 500, fallen die Ergebnisse deutlich besser, aber nicht immer ganz zufriedenstellend aus. Mit dem Schritt auf  $B = 2000$  sieht man keine weitere Änderung. Offensichtlich lassen sich die Ergebnisse für eine relative kleine Stichprobe von  $n = 60$  ab einem gewissen Punkt nicht mehr durch Erhöhung der Anzahl der Bootstrap-Ziehungen erhöhen. Führt man dieselben Berechnungen für Zeitreihen der Länge  $n = 200$  durch, sind die Ergebnisse bei  $B = 100$  Replikationen nur wenig besser als bei  $n = 60$ .<sup>42</sup> Bei  $B = 500$  Replikationen sind die Überdeckungshäufigkeiten bereits fast exakt. Lediglich bei den einfachen Perzentil-Intervallen sind die Abweichungen hier (und auch bei  $B = 2000$ ) noch nennenswert.

---

<sup>42</sup>BCA-Intervalle lassen sich für  $n = 200$ ,  $B = 100$  nicht berechnen.

i.i.d- $t_{10}$ , $\mu_X = 0,002$ , $\sigma_X = 0,1$ , $R = 10\,000$				
	$Sh_X, n = 60$		$Sh_X, n = 200$	
KI-Niveau	95 %	99 %	95 %	99 %
SD-NV	95,28 %	99,13 %	95,20 %	99,13 %
SD-DF	94,87 %	98,82 %	95,04 %	99,04 %
$B = 100$				
TKI	95,13 %	98,84 %	94,73 %	98,81 %
TBias	95,26 %	98,84 %	94,90 %	98,82 %
Student	95,59 %**	98,05 %***	95,01 %	97,83 %***
Perzent	94,56 %*	97,63 %***	94,82 %	97,87 %***
BCA	95,16 %	97,60 %***	-	-
$B = 500$				
TKI	95,13 %	98,89 %	94,67 %	98,84 %
TBias	95,46 %*	98,95 %	94,80 %	98,89 %
Student	95,41 %	99,06 %	94,91 %	99,00 %
Perzent	94,48 %*	98,71 %**	94,51 %*	98,84 %
BCA	95,57 %**	99,10 %	94,82 %	98,94 %
$B = 2\,000$				
TKI	95,08 %	98,83 %	94,73 %	98,88 %
TBias	95,45 %	98,95 %	94,84 %	98,97 %
Student	95,63 %**	99,09 %	94,84 %	99,02 %
Perzent	94,44 %*	98,60 %***	94,53 %*	98,78 %*
BCA	95,56 %*	99,08 %	94,88 %	98,96 %
Zweiseitige Anteilswerttests auf Einhaltung des KI-Niveaus, Ablehnung von $H_0$ : * zum Niveau $\alpha = 5\%$ ; ** zum Niveau $\alpha = 1\%$ ; *** zum Niveau $\alpha = 0,1\%$				

Tabelle 5.30: Überdeckungshäufigkeiten für Konfidenzintervalle des Sharpe-Ratios bei Variation von  $n$  und  $B$



# Kapitel 6

## Schlussfolgerungen

### 6.1 Zusammenfassung

In dieser Arbeit wurden mittels der Delta-Methode asymptotische Varianzen der Schätzer für die Performancemaße Sharpe-Ratio, Treynor-Ratio, Jensens Alpha, Risk-Adjusted Performance (Modigliani-Modigliani), Kappa (darunter Omega und Sortino-Ratio) und Excess Return on Value-at-Risk (ERVaR) hergeleitet. Zur Herleitung wurden die Annahmen unabhängig und identisch verteilter Renditen und der Existenz der ersten (i.d.R. vier) Momente der Renditeverteilungen benötigt. Für alle Performancemaße außer ERVaR wurden mit derselben Methode asymptotische Varianzen für die Differenz der Performanceschätzungen zweier Portfolios bestimmt. Für alle Varianzformeln außer ERVaR und Kappa wurden darüber hinaus vereinfachte Versionen für (ggf. multivariat) normalverteilte Renditen vorgestellt. Alle Varianzformeln sind in Anhang A.2 aufgelistet. Tabelle 6.1 gibt einen Überblick, für welche Performancemaße asymptotische Varianzformeln existieren und welche davon in dieser Arbeit erstmals veröffentlicht sind.

In Kapitel 5 wurden die Eigenschaften der Performancemaßschätzer anhand von Simulationen untersucht. In erster Linie wurde dort überprüft, ob die asymptotische Varianz bzw. deren Schätzer die empirische Varianz gut nähern und inwieweit die Verteilung des Performancemaßschätzers gegen die Normalverteilung konvergiert. Dazu wurden verschiedene Zeitreihen-/ Verteilungsmodelle und Parameterkombinationen systematisch getestet. Außerdem wurden verschiedene Ansätze zur Schätzung von Konfidenzintervallen hinsichtlich der Einhaltung ihres Konfidenzniveaus verglichen.

Bei langen Zeitreihen (hier:  $n = 500$ ) kann man zumeist eine gute bis sehr gute Konvergenz der Verteilung der Performancemaß-Schätzer zur Normalverteilung feststellen und die empirische Varianz der Schätzwerte weicht, gemessen mit einem  $\chi^2$ -Test, in der Regel nicht signifikant von der aus den Parametern berechneten asymptotischen Varianz ab. Das gilt sowohl für i.i.d.-normalverteilte

Performancemaß	Maß $\hat{\theta}_X$		Differenz $\hat{\theta}_X - \hat{\theta}_Y$	
	i.i.d.	i.i.d.-normal	i.i.d.	i.i.d.-normal
Sharpe-Ratio <sup>a</sup>	(X)	(X)	(X)	(X)
Treynor-Ratio	(X)	(X) <sup>b</sup>	(X)	(X)
Jensens Alpha	X	(X)	X	X
MM	X	X	X	X
Kappa	X <sup>c</sup>	–	X	–
ERVaR	X	–	–	–

(X): Formel bereits aus der existierenden Literatur bekannt

X : Formel in dieser Arbeit erstmals veröffentlicht

– : keine Formel vorhanden

<sup>a</sup>Für das Sharpe-Ratio existieren zudem Formeln für nicht unabhängig und identisch verteilte Renditen, siehe Schmid und Schmidt (2010)

<sup>b</sup>Die Formel war in Jobson und Korkie (1981) fehlerhaft angegeben.

<sup>c</sup>Die Varianzformel für Omega ist bereits in Schmid und Schmidt (2008) beschrieben.

Tabelle 6.1: Übersicht über vorhandene Varianzformeln für Performancemaßschätzer

als auch für i.i.d.- $t_5$ -verteilte Simulationswerte. Ausnahmen hiervon sind die Kappa-Performancemaße sowie das Treynor-Ratio für bestimmte Parameterwerte. Hier lässt sich zwar beim Vergleich unterschiedlicher Zeitreihenlängen eine Konvergenz erkennen, aber sie erfolgt sehr langsam.<sup>1</sup>

Der realistischere Fall von  $n = 60$ , entsprechend den Monatsrenditen von fünf Jahren, wurde deutlich umfangreicher simuliert. Wie zu erwarten, ist die Annäherung der Schätzer an die Asymptotik schlechter. In der Mehrzahl der Fälle wurde in den Tests die jeweilige Nullhypothese (Normalverteilung des Schätzers bzw. Schätzvarianz gleich asymptotische Varianz gemäß Berechnung aus den Parametern) abgelehnt, ohne dass die tatsächliche Abweichung erheblich war. Solange die Renditemodelle die Konvergenzbedingungen einhalten (unabhängig und identisch verteilte Renditen, Existenz der ersten (i.d.R. vier) Momente), kann man folgende grobe Einteilung vornehmen:

- Jensens Alpha ist das einzige Maß, bei dem die Konvergenz schon bei  $n = 60$  für alle getesteten Parameterkombinationen annähernd erreicht war. Mit leichten Abstrichen gilt das auch für den ökonomisch relevanten Bereich des Sharpe-Ratios.
- Inferenzstatistik für das Treynor-Ratio ist nur dann sinnvoll, wenn die Korrelation  $\rho_{M,X}$  zwischen Portfolio und Markt so ausgeprägt ist, dass Schätzwerte von  $\hat{\rho}_{M,X}$  in der Umgebung von null praktisch nicht vorkom-

<sup>1</sup>Einziger nicht konvergierender Fall ist das Treynor-Ratio bei unkorrelierten  $X$  und  $M$ , hier ist der theoretische Wert des Performancemaßes nicht definiert.

men. Bei der Stichprobenlänge  $n = 60$  sollte mindestens  $|\rho_{M,X}| \geq 0,5$  gegeben sein, besser  $|\rho_{M,X}| \geq 0,75$ .

- Die Verteilung der Schätzer für die Kappa-Maße weicht insbesondere bei niedrigen Werten des Parameters  $\alpha$  ( $\alpha = 1$ : Omega;  $\alpha = 2$ : Sortino-Ratio) deutlich von der Normalverteilung ab. Die Verteilung des Schätzers ist stark rechtsschief und der Schätzer ist deutlich verzerrt.
- Bei den restlichen Maßen sind die Ergebnisse gemischt. Je nach Parameterkombination liegen einzelne Merkmale der Schätzerverteilung mehr oder weniger nahe an den asymptotischen Zielwerten. Wer sich über die Konvergenzeigenschaften eines bestimmten dieser Maße informieren will, sei auf die ausführlichen Darstellungen in Abschnitt 5.3 verwiesen.

Für die meisten Maße ist eine Formel für unabhängig und identisch verteilte Renditen ohne Annahme eines Verteilungstyps angegeben und eine Formel für den Spezialfall unabhängig und identisch normalverteilter Renditen. Im Falle nicht normalverteilter Renditen bietet letztere Formel nicht mehr die korrekte asymptotische Varianz. Andererseits müssen für die allgemeine Formel mehr höhere Momente geschätzt werden. Deren Schätzfehler und Ausreißeranfälligkeit erhöhen den Schätzfehler im Rechenergebnis. Mit steigendem  $n$  und entsprechend genauerer Schätzung gewinnt der erstgenannte Effekt gegenüber dem zweitgenannten an Bedeutung. Für  $n = 60$  kann man feststellen, dass die Formeln mit Normalverteilungsannahme stabilere Schätzwerte liefern und hinsichtlich der Varianztests und der Überdeckungshäufigkeiten auch bei  $t$ - und GARCH-Zeitreihen keine schlechteren Ergebnisse liefern als die Formeln ohne Verteilungsannahme. Bei der Schätzung der Varianzen aus den Daten kann man häufig beobachten, dass im Mittel die Schätzwerte über die NV-Formeln etwas zu klein und jene aus der verteilungsfreien Formel etwas zu groß sind.

Hinsichtlich der untersuchten Verteilungsmodelle konvergieren Performancemaßschätzer bei i.i.d.-normalverteilten Renditen erwartungsgemäß besser als bei Renditeverteilungen mit höherer Kurtosis. Aber die Varianzschätzer sind relativ wenig empfindlich gegenüber Verletzungen der Verteilungs- oder Konvergenzannahmen. Die Varianzen auf Basis der i.i.d.-Formeln mit und ohne Normalverteilungsannahme waren auch dann brauchbar, wenn sie aus GARCH-Prozessen geschätzt wurden, deren unbedingtes viertes Moment nicht existiert. Sind die Renditen dagegen autokorreliert, sind die i.i.d.-Formeln ungeeignet. Für das Sharpe-Ratio hat sich gezeigt, dass die verallgemeinerte Varianzformel nach Schmid und Schmidt (2010) in der Lage ist, die Autokorrelation von AR(1)-Prozessen adäquat in die Varianzberechnung einzubeziehen.

Hinsichtlich der Konfidenzintervalle für Schätzer von Performancemaßen gibt es keinen generell überlegenen Ansatz. Für die Performancemaße klassischen Zugschnitts (Sharpe, Treynor, Jensen, MM), deren Schätzer zumeist in guter Näherung normalverteilt sind, zeigten Gauß- bzw.  $t$ -Intervalle mit aus der Stichprobe oder per Bootstrap geschätzten Varianzen für  $n = 60$  zufriedenstellende bis gute Überdeckungshäufigkeiten. Für Kappa und ERVaR, deren Schätzerverteilung erheblich von der Normalverteilung abweicht, sollten Konfidenzintervalle auf Ba-

sis von Bootstrap-Perzentilen (z.B. BCA-Intervalle) bevorzugt werden. Ein Teil der beobachteten Abweichungen von den theoretisch-asymptotischen Vorgaben ist auf die geringe Zeitreihenlänge von  $n = 60$  zurückzuführen. Wann immer längere Zeitreihen simuliert wurden, erhielt man eine sichtbare Verbesserung der Konvergenz. Aus einer schmalen Datenbasis lässt sich mit ausgefeilten asymptotischen Ansätzen kein genaues Ergebnis schätzen – im Gegenteil: je mehr Größen für einen Ansatz geschätzt werden müssen, desto mehr Daten sind nötig, damit die größeren zufälligen Fehler so weit reduziert sind, dass die besseren Konvergenzeigenschaften sichtbar werden. Dasselbe gilt auch für Bootstrap-Ansätze. Verfeinerte Methoden und eine höhere Anzahl Bootstrap-Replikationen können nur Informationen auswerten, die in den Ausgangsdaten tatsächlich vorhanden sind.

Wer also Performancemaße und ihre Varianzen schätzen möchte, findet in Abschnitt 3.3 und Kapitel 4 bzw. Anhang A.2 die entsprechenden Formeln und kann sich in Kapitel 5 ein Bild von den statistischen Eigenschaften der einzelnen Schätzer bei kleinen Stichprobengrößen machen.

Wie gesehen ist es in der Regel auch bei Renditezeitreihen moderater Länge möglich, Standardabweichungen und Konfidenzintervalle mit guten bis befriedigenden Ergebnissen zu schätzen. Es bleibt nun die Frage, ob und wann die Schätzung von Performancemaßen auch zu ökonomisch aussagekräftigen Resultaten führt.

## 6.2 Kann man Performancemaße aussagekräftig schätzen?

Eine Schätzung ist ökonomisch aussagekräftig, wenn mit ihr eine ökonomische Größe zuverlässig so genau bestimmt werden kann, dass sie Entscheidungen ermöglicht oder zumindest Tatsachenaussagen substantiell stützt oder entkräftet. Zu Beginn von Kapitel 3 wurden drei Problemstellungen genannt, bei denen Performancemessung potentiell hilfreich sein könnte. Welche Voraussetzungen müssen für die Performancemaßschätzung erfüllt sein, damit sie aussagekräftige Antworten liefern kann?

- a) *Leistungsbewertung*: Hat sich ein Portfolio in der Vergangenheit besser entwickelt als ein Vergleichsmaßstab (z.B. risikoloser Zinssatz)?

Um diese Frage mittels der Schätzung eines Performancemaßes beantworten zu können, muss ein „gut gemanagtes“ Portfolio zuverlässig ein besseres Schätzergebnis erhalten als ein Vergleichsportfolio, das für einen uninformierten Investor erreichbar ist. Gute Portfolios müssen also von schlechten Portfolios unterscheidbar sein.

- b) *Anlageentscheidung*: Ist es lohnend, in eine bestimmte Anlage zu investieren oder nicht?

Damit aufgrund der Schätzung von Performancemaßen aus Renditen vergangener Anlageperioden die Unterstützung einer Entscheidung über die

zukünftige Anlage von Vermögenswerten möglich ist, müssen gute Portfolios im Sinne von a) identifizierbar sein und zusätzlich muss die bessere Performance dauerhafter Natur (persistent) sein. Da empirische Studien eine solche Persistenz nur selten und wenn, in sehr geringem Maße identifizieren konnten<sup>2</sup>, kann man davon ausgehen, dass die Schätzung von Performancemaßen aus Vergangenheitsdaten nicht zur Unterstützung von Anlageentscheidungen geeignet ist.<sup>3</sup>

- c) *Portfoliovergleich*: War die Wertentwicklung eines Portfolios besser als diejenige eines anderen Portfolios?

Hierzu sollte die geschätzte Performancemaß-Differenz  $\widehat{\Delta}_\theta$  zwischen einem besseren und einem schlechteren Portfolio zuverlässig positiv sein. Wenn wir den Vergleich von mehr als zwei Portfolios ausklammern, können wir diese Aufgabe analog zu a) beurteilen, nur dass anstatt des Performancemaßes für ein Portfolio die Differenz des Performancemaßes zweier Portfolios betrachtet wird.

Zur Beantwortung dieser Frage beschränken wir uns hier auf das einfachste und am besten beherrschbare Szenario, nämlich unabhängig und identisch normalverteilte Renditen. Wenn ein Performancemaß unter dieser Annahme aussagekräftig schätzbar ist, dann kann man der Frage nachgehen, ob dies auch für Modelle mit leptokurtischen Verteilungen und abhängigen Renditen der Fall ist. Kommt man dagegen zum Schluss, dass ein Maß bei i.i.d.-normalverteilten Renditen nur sehr ungenau schätzbar ist, dürfte es schwer fallen, für ein realistischeres Verteilungsmodell das Gegenteil nachzuweisen.

Schon den Tabellen und Grafiken in Kapitel 5 ist auffällig, dass der Wert für die asymptotische Standardabweichung des Performancemaßschätzers  $\sigma_0(\widehat{\theta})$  fast immer deutlich größer ist als der Wert des Performancemaßes  $\theta$  bzw. dessen Betrag. Das bedeutet, dass die Schätzvarianz so groß ist, dass der Schätzer mit einer erheblichen Wahrscheinlichkeit so weit vom wahren Wert abweicht, dass er das falsche Vorzeichen hat. Wenn z.B. der wahre Wert des Maßes  $\theta = 0,1$  ist und die Standardabweichung des Schätzers  $\sigma_{\widehat{\theta}} = 0,5$ , dann ist die Wahrscheinlichkeit, dass ein normalverteilter Schätzer fälschlicherweise ein negatives Vorzeichen hat, gleich

$$\Phi\left(\frac{0-\theta}{\sigma_{\widehat{\theta}}}\right) = \Phi\left(\frac{0-0,1}{0,5}\right) = \Phi(-0,2) = 0,4207$$

Entsprechend kann man die Wahrscheinlichkeit für ein falsches Vorzeichen eines normalverteilten Schätzers in Abhängigkeit vom Kehrwert des Variationskoeffizienten  $1/VC_{\widehat{\theta}} = \theta/\sigma_{\widehat{\theta}}$  aus der Verteilungstabelle der Standardnormalverteilung ablesen. Zur Orientierung sind in Tabelle 6.2 einige Werte genannt.

Um zumindest das Vorzeichen des Performancemaßes einigermaßen zuverlässig

<sup>2</sup>Vgl. Abschnitt 3.5, S. 67 f.

<sup>3</sup>Das schließt nicht aus, dass Performancemaße für Anlageentscheidungen nützlich sein können, wenn sie auf anderem Wege als über die vergangenen Renditen des betreffenden Portfolios ermittelt werden.

$\theta/\sigma_{\hat{\theta}}$	$P(\hat{\theta} < 0)$	$\theta/\sigma_{\hat{\theta}}$	$P(\hat{\theta} < 0)$
0	50%	1,0	15,87 %
0,05	48,01%	1,25	10,56 %
0,1	46,02%	1,5	6,68 %
0,2	42,07%	1,75	4,01 %
0,3	38,21%	2	2,28 %
0,4	34,46%	2,5	0,62 %
0,5	30,85%	3	0,13 %
0,6	27,43%	3,5	0,02 %
0,8	21,19%	4	0,003%

Tabelle 6.2: Wahrscheinlichkeiten für falsches Vorzeichen von  $\hat{\theta}$ 

schätzen zu können, sollte die Standardabweichung des Schätzers  $\sigma_{\hat{\theta}}$  möglichst kleiner als der Betrag des wahren Wertes sein.

Die konkrete Größe des Verhältnisses  $\theta/\sigma_{\hat{\theta}}$  ist bei Simulationen natürlich das Resultat der gewählten Werte für die Modellparameter. Die Standardabweichung  $\sigma_{\hat{\theta}}$  kann durch Erhöhung der Zeitreihenlänge  $n$  beliebig verkleinert werden. Der Wert des Performancemaßes hängt bei gegebenem  $\sigma_X$  in erster Linie von der erwarteten Portfoliorendite  $\mu_X$  ab, die bei allen hier untersuchten Performancemaßen entweder ein Summand oder der Zähler eines Bruchs ist. Tabelle 6.3 enthält die asymptotischen Standardabweichungen  $\sigma_0(\hat{\theta})$  aller untersuchten Performancemaße für verschiedene Kombinationen von erwarteter Portfoliorendite  $\mu_X$  und Zeitreihenlänge  $n$ . Alle anderen Parameter sind konstant:  $\sigma_X = 0,1$  sowie für Performancemaße mit Marktportfolio  $\mu_M = 0,002$ ,  $\sigma_M = 0,1$  und  $\rho_{X,M} = 0,75$ . Werte von  $\sigma_0(\hat{\theta})$  kleiner als der wahre Wert des Performancemaßes  $\theta$  sind kursiv gedruckt; Werte, die um mindestens den Faktor 1,96 kleiner als  $\theta$  sind, sind fett gedruckt. Bei ersteren ist die Wahrscheinlichkeit für ein fälschlicherweise negatives Vorzeichen von  $\hat{\theta}$  kleiner als 15,87 %, bei letzterem kleiner als 2,5 %.

Um das Vorzeichen von  $\hat{\theta}$  mit einiger Sicherheit richtig zu schätzen, sind relativ hohe Werte für  $\mu_X$  nötig. Bei  $n = 60$  muss  $\mu_X$  stets größer als 0,01 sein, damit das Vorzeichen von  $\hat{\theta}$  in mindestens ca. 85 % der Fälle positiv ist. Wenn man die sonstige Parameterkombination als geeignet für Monatsrenditen ansieht, muss die monatliche Überrendite also über 1% betragen (entsprechend  $e^{12 \cdot 0,01} - 1 = 12,7\%$  jährlicher diskreter Überrendite). Das ist ökonomisch gesehen ausgesprochen viel.<sup>4</sup>

<sup>4</sup>Man kann hier einwenden, dass eine Standardabweichung der Monatsrenditen von  $\sigma_X = 0,1$  für diversifizierte Portfolios überhöht erscheint. Dann sollte man eine geeignetere Annahme über  $\sigma_X$  machen und sich an den in der zweiten Zeile von Tabelle 6.3 angegebenen Variationskoeffizienten orientieren. Für  $\sigma_X = 0,04$  z.B. entspräche ein Variationskoeffizient von 10

$\mu_X$	0	0,001	0,002	0,003	0,005	0,01	0,02	0,03	0,05
$VC_X$	$\infty$	<b>100</b>	<b>50</b>	<b>33</b>	<b>20</b>	<b>10</b>	<b>5</b>	<b>3,33</b>	<b>2</b>
$Sh_X$	0	0,01	0,02	0,03	0,05	0,1	0,2	0,3	0,5
$n = 36$	0,1667	0,1667	0,1667	0,1667	0,1668	0,1671	<i>0,1683</i>	<i>0,1704</i>	<b>0,1768</b>
$n = 60$	0,1291	0,1291	0,1291	0,1291	0,1292	0,1294	<i>0,1304</i>	<b>0,1320</b>	<b>0,1369</b>
$n = 120$	0,0913	0,0913	0,09136	0,0913	0,0913	<i>0,0915</i>	<b>0,0922</b>	<b>0,0933</b>	<b>0,0968</b>
$n = 500$	0,0447	0,0447	0,0447	0,0447	<i>0,0447</i>	<b>0,0448</b>	<b>0,0452</b>	<b>0,0457</b>	<b>0,0474</b>
$Tr_X$	0	0,0013	0,0027	0,004	0,0067	0,0133	0,0267	0,04	0,0667
$n = 36$	0,0222	0,0222	0,0222	0,0222	0,0222	0,0223	<i>0,0226</i>	<i>0,023</i>	<b>0,0243</b>
$n = 60$	0,0172	0,0172	0,0172	0,0172	0,0172	0,0173	<i>0,0175</i>	<b>0,0178</b>	<b>0,0188</b>
$n = 120$	0,0122	0,0122	0,0122	0,0122	0,0122	<i>0,0122</i>	<b>0,0124</b>	<b>0,0126</b>	<b>0,0133</b>
$n = 500$	0,006	0,006	0,006	0,006	<i>0,006</i>	<b>0,006</b>	<b>0,0061</b>	<b>0,0062</b>	<b>0,0065</b>
Jen. $\alpha_X$	-0,0015	-0,0005	0,0005	0,0015	0,0035	0,0085	0,0185	0,0285	0,0485
$n = 36$	0,011	0,011	0,011	0,011	0,011	0,0111	<i>0,0112</i>	<b>0,0115</b>	<b>0,0123</b>
$n = 60$	0,0085	0,0085	0,0085	0,0085	0,0085	0,0086	<b>0,0087</b>	<b>0,0089</b>	<b>0,0095</b>
$n = 120$	0,006	0,006	0,006	0,006	0,006	<i>0,0061</i>	<b>0,0062</b>	<b>0,0063</b>	<b>0,0068</b>
$n = 500$	0,003	0,003	0,003	0,003	<i>0,003</i>	<b>0,003</b>	<b>0,003</b>	<b>0,0031</b>	<b>0,0033</b>
$MM_{X,ori}$	0	0,001	0,002	0,003	0,005	0,01	0,02	0,03	0,05
$n = 36$	0,0167	0,0167	0,0167	0,0167	0,0167	0,0167	<i>0,0168</i>	<i>0,017</i>	<b>0,0176</b>
$n = 60$	0,0129	0,0129	0,0129	0,0129	0,0129	0,0129	<i>0,013</i>	<b>0,0132</b>	<b>0,0136</b>
$n = 120$	0,0091	0,0091	0,0091	0,0091	0,0091	<i>0,0091</i>	<b>0,0092</b>	<b>0,0093</b>	<b>0,0096</b>
$n = 500$	0,0045	0,0045	0,0045	0,0045	<i>0,0045</i>	<b>0,0045</b>	<b>0,0045</b>	<b>0,0046</b>	<b>0,0047</b>
$MM_{X,alt}$	-0,002	-0,001	0	0,001	0,003	0,008	0,018	0,028	0,048
$n = 36$	0,0118	0,0118	0,0118	0,0118	0,0118	0,01186	<i>0,012</i>	<b>0,0122</b>	<b>0,013</b>
$n = 60$	0,0091	0,0091	0,0091	0,0091	0,0091	0,0092	<i>0,0093</i>	<b>0,0095</b>	<b>0,0101</b>
$n = 120$	0,0065	0,0065	0,0065	0,0065	0,0065	<i>0,0065</i>	<b>0,0066</b>	<b>0,0067</b>	<b>0,0071</b>
$n = 500$	0,0032	0,0032	0,0032	0,0032	0,0032	<b>0,0032</b>	<b>0,0032</b>	<b>0,0033</b>	<b>0,0035</b>
$\kappa(1)_X$	0	0,0254	0,0514	0,0781	0,1335	0,285	0,6517	1,1246	2,5278
$n = 36$	0,4176	0,4178	0,418	0,4183	0,4193	0,4254	<i>0,4622</i>	<b>0,5562</b>	<b>1,0422</b>
$n = 60$	0,3235	0,3236	0,3238	0,324	0,3248	0,3295	<i>0,358</i>	<b>0,4309</b>	<b>0,8073</b>
$n = 120$	0,2287	0,2289	0,2289	0,2291	0,2297	<i>0,233</i>	<b>0,2531</b>	<b>0,3047</b>	<b>0,5708</b>
$n = 500$	0,1121	0,1121	0,1122	0,1122	<i>0,1125</i>	<b>0,1141</b>	<b>0,124</b>	<b>0,1493</b>	<b>0,2796</b>
$\kappa(2)_X$	0	0,0143	0,0287	0,0435	0,0736	0,1534	0,3336	0,5459	1,092
$n = 36$	0,2357	0,2357	0,2358	0,2359	0,2363	0,2383	<i>0,2492</i>	<b>0,2743</b>	<b>0,3957</b>
$n = 60$	0,1826	0,1826	0,1826	0,1827	0,183	0,1846	<i>0,193</i>	<b>0,2125</b>	<b>0,3065</b>
$n = 120$	0,1291	0,1291	0,1291	0,1292	0,1294	<i>0,1305</i>	<b>0,1365</b>	<b>0,1503</b>	<b>0,2167</b>
$n = 500$	0,0632	0,0633	0,0633	0,0633	<i>0,0634</i>	<b>0,0639</b>	<b>0,0669</b>	<b>0,0736</b>	<b>0,1062</b>
$ERV_{aR}$	0	0,0061	0,0123	0,01866	0,03146	0,0647	0,1384	0,2231	0,4367
$n = 36$	0,1013	0,1026	0,1039	0,1052	0,108	0,1156	<i>0,1346</i>	<i>0,1601</i>	<i>0,2403</i>
$n = 60$	0,0785	0,0795	0,0805	0,0815	0,0836	0,0895	<i>0,1043</i>	<i>0,124</i>	<b>0,1862</b>
$n = 120$	0,0555	0,0562	0,0569	0,0576	0,0591	<i>0,0633</i>	<i>0,0737</i>	<b>0,0877</b>	<b>0,1316</b>
$n = 500$	0,0272	0,0275	0,0279	0,0282	<i>0,029</i>	<b>0,031</b>	<b>0,0361</b>	<b>0,043</b>	<b>0,0645</b>

Asymptotische Varianzen verschiedener Performancemaße für i.i.d.-normalverteilte Daten mit  $\sigma_X = 0,1$  sowie ggf.  $\mu_M = 0,002$ ,  $\sigma_M = 0,1$  und  $\rho_{M,X} = 0,75$  bei Variation von  $\mu_X$  und Stichprobenlänge  $n$ .

Performancemaß  $\kappa(\alpha, z)$  mit Parameter  $z = 0$ ;  $ERV_{aR}(\alpha)$  mit Parameter  $\alpha = 0,05$ .

Erste Zeile der Tabelle: Spaltenwerte für  $\mu_X$ , darunter der Variationskoeffizient  $\sigma_X/\mu_X$ .

Je Performancemaß:

Erste Zeile: Theoretischer Wert  $\theta$  des Performancemaßes bei jeweiligem  $\mu_X$ .

Zeilen darunter: Asymptotische Standardabweichungen  $\sigma_0(\hat{\theta})$  für verschiedene  $n$  und  $\mu_X$ . Werte *kursiv*, wenn  $\sigma_0(\hat{\theta}) \leq \theta < 1,96 \sigma_0(\hat{\theta})$ ; Werte **fett**, wenn  $\theta \geq 1,96 \sigma_0(\hat{\theta})$ .

Tabelle 6.3: Asymptotische Varianzen von Performancemaßen bei verschiedenen erwarteten Renditen und Stichprobenlängen

In Kapitel 5 werden auch andere Parameter als  $\mu_X$  und  $n$  variiert. Aber auch dort zeigt eine Durchsicht der Tabellen und Grafiken, dass die asymptotische Standardabweichung meist größer ist als der Betrag des Schätzwertes. Die  $1,96 - \sigma_0(\hat{\theta})$ -Umgebung um den wahren Wert schließt den Wert Null wird fast immer mit ein.<sup>5</sup>

Man könnte zur Erhöhung der Schätzgenauigkeit nun längere Zeitreihen heranziehen, etwa die Monatsrenditen von zehn statt von fünf Jahren ( $n = 120$ ). Wenn man die damit verbundenen Probleme außer acht lässt (z.B. Annahme der Konstanz von  $\theta$  über einen langen Zeitraum), ist damit nicht allzu viel gewonnen, da die Standardabweichungen der Schätzer durch Verdopplung der Zeitreihenlänge nur um den Faktor  $\sqrt{2}$  sinken.

Wenn man die Frequenz der Renditen erhöht, etwa von Monats- auf Wochen- oder Tagesrenditen übergeht, erhöht man ebenfalls die Zeitreihenlänge. Auch diese Maßnahme bringt Probleme mit sich, etwa stärkere Leptokurtosis und ausgeprägtere intertemporale Abhängigkeit der Renditen (vgl. Abschnitt 2.3.1.1). Im Gegenzug wird die Konvergenz der Verteilung des Schätzers zur Normalverteilung durch die deutlich höhere Stichprobenlänge verbessert, wie in den Abschnitten 5.2 und 5.3 des vorigen Kapitels an einigen Stellen aufgezeigt wird. Das ursprüngliche Problem wird damit aber nicht gelöst. Bei der Aufspaltung von Periodenrenditen in  $k$  Teilperioden sinkt zwar einerseits die Standardabweichung der Rendite je Periode, bei i.i.d.-Renditen um den Faktor  $\sqrt{k}$ . Ebenso führt die Erhöhung der Stichprobenlänge zu einer zusätzlichen Reduktion des Schätzfehlers um den Faktor  $\sqrt{k}$ . Die erwartete Rendite der Teilperiode sinkt aber ebenfalls auf  $1/k$  der erwarteten Rendite der Ausgangsperiode, was den Effekt vollständig aufhebt.

Wenn man Zeitreihen mit einer Länge von maximal fünf oder zehn Jahren einsetzt, lassen sich Performancemaße also nur sehr ungenau schätzen. Die wahre Überrendite muss ausgesprochen hoch sein, um überhaupt mit einiger Sicherheit das Vorzeichen des Performancemaßes richtig zu schätzen.

Im Großen und Ganzen gelten diese Ergebnisse auch für die Schätzung der Differenz der Performancemaße zweier Portfolios. Wenn die zwei verglichenen Portfolios allerdings sehr stark korreliert sind, können auch geringere Renditeunterschiede relativ sicher zum richtigen Vorzeichen der Differenz und somit zur richtigen Reihung der beiden Portfolios führen.

Der Grund für diese schlechte Aussagekraft von Performancemaß-Schätzungen liegt nicht an den statistischen Eigenschaften – die Schätzer sind schon bei überschaubarer Stichprobengröße in ausreichender Näherung normalverteilt und auch die asymptotischen Varianzen sowie ihre Schätzer sind schon bei  $n = 60$  zumeist gut brauchbar.

Das Problem ist, dass die auf den Finanzmärkten realistisch erzielbaren Über-

---

einer Überrendite von 0,4 % pro Monat. Diese Umrechnung ist bei Maßen mit Marktportfolio nur möglich, wenn man die Parameter des Marktportfolios ebenfalls entsprechend ändert.

<sup>5</sup>Ausnahme: Jensens Alpha bei sehr geringer Renditevarianz  $\sigma_X$  (vgl. Abb. 5.13, S. 145)



$n$	12	36	60	120	250	500	1 000
$VC(= \sqrt{n})$	3,46	6	7,75	10,95	15,81	22,36	31,62
$\sigma_X$	$\sigma_X/\sqrt{n} \hat{=} \text{Minimum für } \mu_X, \text{ damit } P(\bar{X} > 0) \geq 84\%$						
0,01	0,0029	0,0017	0,0013	0,0009	0,0006	0,0004	0,0003
0,02	0,0058	0,0033	0,0026	0,0018	0,0013	0,0009	0,0006
0,05	0,0144	0,0083	0,0065	0,0046	0,0032	0,0022	0,0016
0,1	0,0289	0,0167	0,0129	0,0091	0,0063	0,0045	0,0032
entspricht diskreten Jahresrenditen /-standardabweichungen							
$\sigma_X, \text{ Jahr, diskret}$	$\mu_X, \text{ Jahr, diskret}$						
3,52%	3,52%	2,02%	1,56%	1,10%	0,76%	0,54%	0,38%
7,17%	7,17%	4,08%	3,15%	2,22%	1,53%	1,08%	0,76%
18,91%	18,91%	10,52%	8,05%	5,63%	3,87%	2,72%	1,92%
41,40%	41,40%	22,14%	16,76%	11,58%	7,88%	5,51%	3,87%

Tabelle 6.4: Minimale erwartete Portfoliorendite pro Monat bzw. Jahr für eine mit mind. 84% Wahrscheinlichkeit korrekte Schätzung des Vorzeichens von  $\bar{X}$

renditen im Verhältnis zu ihrer Variabilität einfach zu gering sind. Das lässt sich ganz einfach anhand der Schätzvarianz der (Über-)Rendite zeigen:

Die Überrendite eines Portfolios  $\mu_X$  kann effizient mit dem Mittelwert der empirischen Renditen  $\bar{X}$  geschätzt werden. Bei normalverteilten Renditen – und ansonsten, wenn der Zentrale Grenzwertsatz angewendet werden kann, asymptotisch – ist die Standardabweichung dieses Schätzers

$$\sigma_{\bar{X}} = \frac{\sigma_X}{\sqrt{n}}.$$

Hiermit kann man zum Beispiel berechnen, wie groß  $\mu_X$  sein muss, damit  $\mu_X/\sigma_{\bar{X}}$  größer als eins ist, so dass die Wahrscheinlichkeit, dass der Schätzer  $\bar{X}$  das richtige Vorzeichen hat (vgl. Tabelle 6.2) zumindest kleiner als ca. 16% ist.

$$\begin{aligned} \frac{\mu_X}{\sigma_{\bar{X}}} &= \frac{\mu_X}{\frac{\sigma_X}{\sqrt{n}}} \geq 1 \\ \text{oder} \quad \mu_X &\geq \frac{\sigma_X}{\sqrt{n}} \\ \text{oder} \quad \frac{\sigma_X}{\mu_X} &= VC_X \leq \sqrt{n} \end{aligned}$$

Damit also der Schätzer der Rendite mit wenigstens 84% Wahrscheinlichkeit überhaupt das richtige Vorzeichen hat – geschweige denn, dass er die wahre Überrendite einigermaßen genau bestimmt –, darf der Variationskoeffizient der Rendite  $VC_X$  nicht größer als  $\sqrt{n}$  sein.

Tabelle 6.4 zeigt die Portfolio(über)renditen, die bei gegebener Stichprobenlänge  $n$  nötig sind, damit  $VC_X < \sqrt{n}$  gilt. Beispielsweise ist bei  $n = 60$

und einer monatlichen Standardabweichung von 5% eine erwartete Monatsrendite von 0,65 % erforderlich. Annualisiert und in diskrete Renditen übersetzt heißt das, dass bei einer jährlichen Standardabweichung von 18,91 % und der Schätzung mit Monatsrenditen aus fünf Jahren die jährliche erwartete Rendite mindestens 8,05 Prozentpunkte über dem risikolosen Zinssatz liegen muss, nur damit das Vorzeichen in ca. fünf von sechs Fällen korrekt geschätzt wird.

Auch wenn man eine gewisse Risikoprämie auf den risikolosen Zinssatz mit einkalkuliert, dürften dauerhafte Überrenditen in dieser Höhe zumindest durch klassische Finanzmarktprodukte nur schwer zu erreichen sein.

Alle in dieser Arbeit untersuchten Performancemaße bestehen aus der nur ungenau schätzbaren erwarteten Portfolioüberrendite, die durch eine andere zu schätzende Größe geteilt wird und/oder von der eine solche subtrahiert wird. Bei der Schätzung derartiger Performancemaße bleibt der durch  $\hat{\mu}_X$  eingebrachte Schätzfehler in der Regel bestehen.<sup>6</sup> Deshalb ist es schwierig, durch Schätzungen von Performancemaßen aus Vergangenheitsdaten ökonomisch aussagekräftige Erkenntnisse zu erlangen. Zudem können die Renditedaten selber problembehaftet sein, wie es in Abschnitt 2.3.3 für Hedgefonds dargestellt wurde.

Vom theoretischen Standpunkt sind Performancemaße dazu geeignet, Eigenschaften von Finanzmarktrenditen für das Nutzenkalkül von Anlegern nutzbar zu machen. Man kann zum Beispiel realistische Annahmen über die langfristigen Renditeeigenschaften von Anlageklassen (Festgeld, Rentenfonds, Aktien) machen und Performancemaße zu Rate ziehen, um die Entscheidung über eine bestimmte Vermögensstruktur (z.B. mittels des Sharpe-Ratios) oder über eine Neuanlage kleinerer Vermögensposten (z.B. mittels des Treynor-Ratios) quantitativ zu fundieren.

Eine Ex-post-Bewertung von Portfolios oder Informationsobjekten anhand von Performancemaßen, die aus den realisierten Renditen geschätzt werden, erscheint dagegen angesichts der Streuung der Schätzer von zweifelhaftem Nutzen.

---

<sup>6</sup>Eine Ausnahme sind Jensens Alpha und  $MM_{alt}$ , wenn das untersuchte Portfolio und das Marktportfolio sehr stark korreliert sind. Die Schätzfehler von  $\bar{X}$  und  $\bar{M}$  heben sich dann teilweise gegeneinander auf.

# Anhang A

## Anhang

### A.1 Umformungen höherer Momente

#### Umformungen höherer Momente bei beliebig verteilten Daten

$$\begin{aligned}\sigma_{A,(A-\mu_A)^2} &= E(A(A-\mu_A)^2) - \mu_A \sigma_A^2 & (A.1) \\ &= E(A(A-\mu_A)^2 - \mu_A(A-\mu_A)^2 + \mu_A(A-\mu_A)^2) - \mu_A \sigma_A^2 \\ &= E((A-\mu_A)^3) + E(\mu_A(A-\mu_A)^2) - \mu_A \sigma_A^2 \\ &= \sigma_A^3 \frac{E((A-\mu_A)^3)}{\sigma_A^3} + \mu_A \sigma_A^2 - \mu_A \sigma_A^2 \\ &= \sigma_A^3 \gamma_{1,A}\end{aligned}$$

$$\begin{aligned}\sigma_{(A-\mu_A)^2}^2 &= E(((A-\mu_A)^2)^2 - (E((A-\mu_A)^2))^2) & (A.2) \\ &= E((A-\mu_A)^4 - (\sigma_A^2)^2) \\ &= E((A-\mu_A)^4) - \sigma_A^4 = \frac{\sigma_A^4 E((A-\mu_A)^4)}{\sigma_A^4} - \sigma_A^4 \\ &= \sigma_A^4 (\gamma_{2,A} - 1)\end{aligned}$$

#### Umformungen höherer Momente bei normalverteilten Daten

Es sei  $(A, B, C, D) \sim N(\mathbf{0}, \mathbf{\Gamma})$ . Dann gilt nach dem Satz von Isserlis (siehe S. 75):

$$E(ABC) = E(A^2 B) = E(A^3) = 0 \quad (A.3)$$

$$E(ABCD) = \sigma_{A,B} \sigma_{C,D} + \sigma_{A,C} \sigma_{B,D} + \sigma_{A,D} \sigma_{B,C} \quad (A.4)$$

$$E(A^2 BC) = \sigma_A^2 \sigma_{B,C} + 2\sigma_{A,B} \sigma_{A,C} \quad (A.5)$$

$$E(A^2 B^2) = \sigma_A^2 \sigma_B^2 + 2\sigma_{A,B}^2 \quad (A.6)$$

$$E(A^3 B) = 3\sigma_A^2 \sigma_{A,B} \quad (A.7)$$

$$E(A^4) = 3\sigma_A^4 \quad (A.8)$$

Seien  $(A', B') \sim N(\boldsymbol{\mu}, \boldsymbol{\Gamma})$ , so dass  $((A' - \mu_{A'}), (B' - \mu_{B'})) \sim N(\mathbf{0}, \boldsymbol{\Gamma})$ . Für Momente dritter Ordnung gilt

$$\sigma_{A', (A' - \mu_{A'})^2} \stackrel{(A.1)}{=} \sigma_A^3 \gamma_{1, A'} \stackrel{(\gamma_{1, A'} = 0)}{=} 0 \quad (A.9)$$

$$\begin{aligned} \sigma_{B', (A' - \mu_{A'})^2} &= E(B'(A' - \mu_{A'})^2) - \mu_{B'} \sigma_{A'}^2 & (A.10) \\ &= E((B' - \mu_{B'})(A' - \mu_{A'})^2 + \mu_{B'}(A' - \mu_{A'})^2) - \mu_{B'} \sigma_{A'}^2 \\ &= E((B' - \mu_{B'})(A' - \mu_{A'})^2) + \mu_{B'} \sigma_{A'}^2 - \mu_{B'} \sigma_{A'}^2 \\ &= E((B' - \mu_{B'})(A' - \mu_{A'})^2) \\ &\stackrel{(A.3)}{=} 0 \end{aligned}$$

Mit (A.3) bis (A.8) können die Momente vierter Ordnung durch Momente zweiter Ordnung ausgedrückt werden.

$$\sigma_{A^2}^2 = E(A^4) - (\sigma_A^2)^2 = 2\sigma_A^4 \quad (A.11)$$

$$\sigma_{AB}^2 = E(A^2 B^2) - \sigma_{A,B}^2 = \sigma_A^2 \sigma_B^2 + \sigma_{A,B}^2 \quad (A.12)$$

$$\sigma_{A^2, B^2} = E(A^2 B^2) - \sigma_A^2 \sigma_B^2 = 2\sigma_{A,B}^2 \quad (A.13)$$

$$\sigma_{A^2, BC} = E(A^2 BC) - \sigma_A^2 \sigma_{B,C} = 2\sigma_{A,B} \sigma_{A,C} \quad (A.14)$$

$$\sigma_{AB, AC} = E(A^2 BC) - \sigma_{A,B} \sigma_{A,C} = \sigma_A^2 \sigma_{B,C} + \sigma_{A,B} \sigma_{A,C} \quad (A.15)$$

$$\sigma_{A^2, AB} = E(A^3 B) - \sigma_A^2 \sigma_{A,B} = 2\sigma_A^2 \sigma_{A,B} \quad (A.16)$$

## A.2 Varianzformeln für Performancemaße

### A.2.1 Sharpe-Ratio

Varianz von  $\widehat{Sh}_X$ ,  $X_t$  i.i.d.

$$n \sigma_{\widehat{Sh}}^2 = 1 - Sh_X \gamma_{1,X} + \frac{1}{4} Sh_X^2 (\gamma_{2,X} - 1) \quad (\text{A.17})$$

Varianz von  $\widehat{Sh}_X$ ,  $X_t$  i.i.d. normalverteilt

$$n \sigma_{\widehat{Sh},norm.}^2 = 1 + \frac{1}{2} Sh_X^2 \quad (\text{A.18})$$

Varianz von  $\widehat{Sh}_X$ ,  $X_t$  nach Schmid und Schmidt (2010)

$$n \sigma_{\widehat{Sh}}^2 = \frac{\sigma_{11}}{\sigma_X^2} - \frac{\sigma_{12}}{\sigma_X^3} Sh_X + \frac{1}{4} Sh_X^2 \frac{\sigma_{22}}{\sigma_X^4} \quad \text{mit} \quad \mathbf{\Gamma} := \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix} \quad (\text{A.19})$$

mit dem Schätzer für  $\mathbf{\Gamma}$

$$\widehat{\mathbf{\Gamma}} = \sum_{j=-l}^l \frac{1}{n - |j|} \sum_{t=1-j+\max(0,j)}^{n-j+\min(0,j)} \begin{pmatrix} (X_t - \bar{X}_n)(X_{t+j} - \bar{X}_n) & (X_t - \bar{X}_n)(X_{t+j} - \bar{X}_n)^2 \\ (X_t - \bar{X}_n)(X_{t+j} - \bar{X}_n)^2 & (X_t - \bar{X}_n)^2(X_{t+j} - \bar{X}_n)^2 - (S_n^2)^2 \end{pmatrix}$$

Varianz von  $\widehat{\Delta}_{Sh}$  zweier Portfolios  $X$  und  $Y$ ,  $(X, Y)_t$  i.i.d.

$$\begin{aligned} n \sigma_{\widehat{\Delta}_{Sh}}^2 &= 2 - 2\rho_{X,Y} - Sh_X \gamma_{1,X} - Sh_Y \gamma_{1,Y} + \frac{1}{4} Sh_X^2 (\gamma_{2,X} - 1) \quad (\text{A.20}) \\ &+ \frac{1}{4} Sh_Y^2 (\gamma_{2,Y} - 1) + Sh_X \frac{\sigma_{Y,(X-\mu_X)}^2}{\sigma_X^2 \sigma_Y} + Sh_Y \frac{\sigma_{X,(Y-\mu_Y)}^2}{\sigma_X \sigma_Y^2} \\ &- Sh_X Sh_Y \frac{\sigma_{(X-\mu_X)^2,(Y-\mu_Y)^2}}{2\sigma_X^2 \sigma_Y^2} \end{aligned}$$

Varianz von  $\widehat{\Delta}_{Sh}$  zweier Portfolios  $X$  und  $Y$ ,  $(X, Y)_t$  i.i.d. normalverteilt

$$n \sigma_{\widehat{\Delta}_{Sh},norm.}^2 = 2(1 - \rho_{X,Y}) + \frac{1}{2} (Sh_X^2 + Sh_Y^2 - 2Sh_X Sh_Y \rho_{X,Y}^2) \quad (\text{A.21})$$

### A.2.2 Treynor-Ratio

Varianz von  $\widehat{Tr}_X$ ,  $(X, M)_t$  i.i.d.

$$\begin{aligned} n \sigma_{\widehat{Tr}}^2 &= Tr_X^2 \left[ \frac{\sigma_X^2}{\mu_X^2} + (\gamma_{2,M} - 1) + \frac{\sigma_{(X-\mu_X)(M-\mu_M)}^2}{\sigma_{X,M}^2} + 2 \frac{\sigma_{X,(M-\mu_M)}^2}{\sigma_M^2 \mu_X} \right. \\ &\quad \left. - 2 \frac{\sigma_{X,(X-\mu_X)(M-\mu_M)}}{\sigma_{X,M} \mu_X} - 2 \frac{\sigma_{(M-\mu_M)^2,(X-\mu_X)(M-\mu_M)}}{\sigma_{X,M} \sigma_M^2} \right] \quad (\text{A.22}) \end{aligned}$$

Varianz von  $\widehat{Tr}_X$ ,  $(X, M)_t$  i.i.d. normalverteilt

$$n \sigma_{\widehat{Tr}, norm.}^2 = Tr_X^2 \left( VC_X^2 + \frac{1}{\rho_{X,M}^2} - 1 \right) \quad (A.23)$$

Varianz von  $\widehat{\Delta}_{Tr}$  zweier Portfolios  $X$  und  $Y$ ,  $(X, Y, M)_t$  i.i.d.

$$\begin{aligned} n \sigma_{\widehat{\Delta}_{Tr}}^2 &= -2 \frac{\sigma_M^2 \sigma_{X,(M-\mu_M)^2} \mu_Y}{\sigma_{X,M} \sigma_{Y,M}} + 2 \frac{\sigma_M^4 \mu_Y \sigma_{X,(M-\mu_M)(Y-\mu_Y)}}{\sigma_{X,M} \sigma_{Y,M}^2} \quad (A.24) \\ &-2 \frac{\sigma_M^2 \sigma_{Y,(M-\mu_M)^2} \mu_X}{\sigma_{X,M} \sigma_{Y,M}} + 2 \frac{\sigma_M^4 \mu_X \sigma_{Y,(M-\mu_M)(X-\mu_X)}}{\sigma_{Y,M} \sigma_{X,M}^2} \\ &-2 \frac{\mu_X \sigma_{(M-\mu_M)^2}^2 \mu_Y}{\sigma_{X,M} \sigma_{Y,M}} + \frac{\sigma_M^4 \sigma_X^2}{\sigma_{X,M}^2} + \frac{\sigma_M^4 \sigma_Y^2}{\sigma_{Y,M}^2} \\ &+ \frac{\mu_X^2 \sigma_{(M-\mu_M)^2}^2}{\sigma_{X,M}^2} + \frac{\mu_Y^2 \sigma_{(M-\mu_M)^2}^2}{\sigma_{Y,M}^2} - 2 \frac{\sigma_M^4 \sigma_{X,Y}}{\sigma_{X,M} \sigma_{Y,M}} \\ &+ 2 \frac{\sigma_M^2 \sigma_{X,(M-\mu_M)^2} \mu_X}{\sigma_{X,M}^2} - 2 \frac{\sigma_M^4 \mu_X \sigma_{X,(M-\mu_M)(X-\mu_X)}}{\sigma_{X,M}^3} \\ &+ 2 \frac{\sigma_M^2 \sigma_{Y,(M-\mu_M)^2} \mu_Y}{\sigma_{Y,M}^2} - 2 \frac{\sigma_M^4 \mu_Y \sigma_{Y,(M-\mu_M)(Y-\mu_Y)}}{\sigma_{Y,M}^3} \\ &-2 \frac{\mu_X^2 \sigma_M^2 \sigma_{(M-\mu_M)^2, (M-\mu_M)(X-\mu_X)}}{\sigma_{X,M}^3} \\ &-2 \frac{\mu_Y^2 \sigma_M^2 \sigma_{(M-\mu_M)^2, (M-\mu_M)(Y-\mu_Y)}}{\sigma_{Y,M}^3} \\ &+ \frac{\mu_X^2 \sigma_M^4 \sigma_{(M-\mu_M)(X-\mu_X)}^2}{\sigma_{X,M}^4} + \frac{\mu_Y^2 \sigma_M^4 \sigma_{(M-\mu_M)(Y-\mu_Y)}^2}{\sigma_{Y,M}^4} \\ &+ 2 \frac{\mu_X \mu_Y \sigma_M^2 \sigma_{M-\mu_M^2, (M-\mu_M)(Y-\mu_Y)}}{\sigma_{X,M} \sigma_{Y,M}^2} \\ &+ 2 \frac{\mu_X \mu_Y \sigma_M^2 \sigma_{(M-\mu_M)^2, (M-\mu_M)(X-\mu_X)}}{\sigma_{Y,M} \sigma_{X,M}^2} \\ &-2 \frac{\sigma_M^4 \mu_X \mu_Y \sigma_{(M-\mu_M)(X-\mu_X), (M-\mu_M)(Y-\mu_Y)}}{\sigma_{X,M}^2 \sigma_{Y,M}^2} \end{aligned}$$

Varianz von  $\widehat{\Delta}_{Tr}$  zweier Portfolios  $X$  und  $Y$ ,  $(X, Y, M)_t$  i.i.d. normalverteilt

$$\begin{aligned} n \sigma_{\widehat{\Delta}_{Tr}, norm.}^2 &= Tr_X^2 \left( VC_X^2 + \frac{1}{\rho_{X,M}^2} - 1 \right) \quad (A.25) \\ &+ Tr_Y^2 \left( VC_Y^2 + \frac{1}{\rho_{Y,M}^2} - 1 \right) \\ &- 2 \rho_{X,Y} Tr_X Tr_Y \left( VC_X VC_Y + \frac{1}{\rho_{X,M} \rho_{Y,M}} - \frac{1}{\rho_{X,Y}} \right) \end{aligned}$$

### A.2.3 Jensens Alpha

Varianz von  $\hat{\alpha}_X, (X, M)_t$  i.i.d.

$$\begin{aligned}
 n \sigma_{\hat{\alpha}}^2 &= \sigma_X^2 - \frac{\sigma_{X,M}^2}{\sigma_M^2} + \frac{\mu_M \sigma_{X,M} \sigma_{X,(M-\mu_M)^2}}{\sigma_M^4} & (A.26) \\
 &- \frac{2 \mu_M \sigma_{X,(X-\mu_X)(M-\mu_M)}}{\sigma_M^2} - \frac{2 \sigma_{X,M}^2 \mu_M \sigma_{M,(M-\mu_M)^2}}{\sigma_M^6} \\
 &+ \frac{2 \mu_M \sigma_{X,M} \sigma_{M,(X-\mu_X)(M-\mu_M)}}{\sigma_M^4} + \frac{\mu_M^2 \sigma_{X,M}^2 \sigma_{(M-\mu_M)^2}^2}{\sigma_M^8} \\
 &- \frac{2 \mu_M^2 \sigma_{X,M} \sigma_{(M-\mu_M)^2, (X-\mu_X)(M-\mu_M)}}{\sigma_M^6} + \frac{\mu_M^2 \sigma_{(X-\mu_X)(M-\mu_M)}^2}{\sigma_M^4}.
 \end{aligned}$$

Varianz von  $\hat{\alpha}_X, (X, M)_t$  i.i.d. normalverteilt

$$n \sigma_{\hat{\alpha},norm.}^2 = \sigma_X^2 (1 - \rho_{X,M}^2) \left( 1 + \frac{\mu_M^2}{\sigma_M^2} \right) \quad (A.27)$$

Varianz von  $\widehat{\Delta}_\alpha$  zweier Portfolios  $X$  und  $Y$ ,  $(X, Y, M)_t$  i.i.d.

$$\begin{aligned}
n \sigma_{\widehat{\Delta}_\alpha}^2 &= \sigma_Y^2 + \sigma_X^2 - 2\sigma_{X,Y} & (A.28) \\
&+ \frac{2\sigma_{Y,M}\mu_M\sigma_{M,(Y-\mu_Y)(M-\mu_M)}}{\sigma_M^4} - \frac{2\mu_M\sigma_{X,M}^2\sigma_{M,(M-\mu_M)^2}}{\sigma_M^6} \\
&- \frac{2\mu_M\sigma_{X,M}\sigma_{Y,(M-\mu_M)^2}}{\sigma_M^4} - \frac{2\mu_M\sigma_{Y,M}^2\sigma_{M,(M-\mu_M)^2}}{\sigma_M^6} \\
&- \frac{2\sigma_{Y,M}\mu_M\sigma_{M,(X-\mu_X)(M-\mu_M)}}{\sigma_M^4} - \frac{2\mu_M\sigma_{Y,M}\sigma_{X,(M-\mu_M)^2}}{\sigma_M^4} \\
&+ \frac{2\mu_M\sigma_{X,M}\sigma_{X,(M-\mu_M)^2}}{\sigma_M^4} + \frac{2\mu_M\sigma_{Y,M}\sigma_{Y,(M-\mu_M)^2}}{\sigma_M^4} \\
&+ \frac{4\sigma_{X,M}\sigma_{Y,M}\mu_M\sigma_{M,(M-\mu_M)^2}}{\sigma_M^6} - \frac{2\mu_M^2\sigma_{X,M}\sigma_{Y,M}\sigma_{(M-\mu_M)^2}^2}{\sigma_M^8} \\
&- \frac{2\mu_M\sigma_{X,(X-\mu_X)(M-\mu_M)}}{\sigma_M^2} + \frac{2\mu_M\sigma_{X,(Y-\mu_Y)(M-\mu_M)}}{\sigma_M^2} \\
&+ \frac{2\mu_M\sigma_{Y,(X-\mu_X)(M-\mu_M)}}{\sigma_M^2} - \frac{2\mu_M\sigma_{Y,(Y-\mu_Y)(M-\mu_M)}}{\sigma_M^2} \\
&+ \frac{2\sigma_{X,M}\sigma_{Y,M}}{\sigma_M^2} + \frac{\mu_M^2\sigma_{(X-\mu_X)(M-\mu_M)}^2}{\sigma_M^4} \\
&- \frac{2\mu_M^2\sigma_{(X-\mu_X)(M-\mu_M),(Y-\mu_Y)(M-\mu_M)}}{\sigma_M^4} + \frac{\mu_M^2\sigma_{(Y-\mu_Y)(M-\mu_M)}^2}{\sigma_M^4} \\
&- \frac{\sigma_{X,M}^2}{\sigma_M^2} - \frac{\sigma_{Y,M}^2}{\sigma_M^2} + \frac{2\mu_M\sigma_{X,M}\sigma_{M,(X-\mu_X)(M-\mu_M)}}{\sigma_M^4} \\
&- \frac{2\mu_M\sigma_{X,M}\sigma_{M,(Y-\mu_Y)(M-\mu_M)}}{\sigma_M^4} + \frac{\mu_M^2\sigma_{(M-\mu_M)^2}^2\sigma_{Y,M}^2}{\sigma_M^8} \\
&+ \frac{2\mu_M^2\sigma_{Y,M}\sigma_{(M-\mu_M)^2,(X-\mu_X)(M-\mu_M)}}{\sigma_M^6} \\
&- \frac{2\mu_M^2\sigma_{Y,M}\sigma_{(M-\mu_M)^2,(Y-\mu_Y)(M-\mu_M)}}{\sigma_M^6} \\
&+ \frac{\mu_M^2\sigma_{(M-\mu_M)^2}^2\sigma_{X,M}^2}{\sigma_M^8} - \frac{2\mu_M^2\sigma_{X,M}\sigma_{(M-\mu_M)^2,(X-\mu_X)(M-\mu_M)}}{\sigma_M^6} \\
&+ \frac{2\mu_M^2\sigma_{X,M}\sigma_{(M-\mu_M)^2,(Y-\mu_Y)(M-\mu_M)}}{\sigma_M^6}
\end{aligned}$$

Varianz von  $\widehat{\Delta}_\alpha$  zweier Portfolios  $X$  und  $Y$ ,  $(X, Y, M)_t$  i.i.d. normalverteilt

$$\begin{aligned}
n \sigma_{\widehat{\Delta}_\alpha, norm.}^2 &= \left(1 + \frac{\mu_M^2}{\sigma_M^2}\right) [\sigma_X^2(1 - \rho_{X,M}^2) + \sigma_Y^2(1 - \rho_{Y,M}^2)] & (A.29) \\
&- 2\sigma_X\sigma_Y(\rho_{X,Y} - \rho_{X,M}\rho_{Y,M})
\end{aligned}$$



### A.2.4 Modigliani-Modigliani

Varianz von  $\widehat{MM}_{X,alt}$ ,  $(X, M)_t$  i.i.d. (alternative Formulierung von MM)

$$\begin{aligned} n \sigma_{\widehat{MM}_{alt}}^2 &= 2 \sigma_M^2 \left( 1 - \frac{\rho_{X,M}}{\sigma_X^2} \right) \\ &+ \frac{\mu_X}{\sigma_X} \left( \frac{\sigma_{X,(M-\mu_M)}^2}{\sigma_X} + \frac{\sigma_M \sigma_{M,(X-\mu_X)}^2}{\sigma_X^2} - \sigma_M^2 (\gamma_{1,X} + \gamma_{1,M}) \right) \\ &+ \frac{\mu_X^2}{\sigma_X^2} \left( \frac{\sigma_M^2}{4} (\gamma_{2,X} + \gamma_{2,M} - 2) - \frac{\sigma_{(X-\mu_X)^2, (M-\mu_M)^2}}{2 \sigma_X^2} \right) \end{aligned} \quad (\text{A.30})$$

Varianz von  $\widehat{MM}_{X,alt}$ ,  $(X, M)_t$  i.i.d. normalverteilt

$$n \sigma_{\widehat{MM}_{alt, norm.}}^2 = \sigma_M^2 \left[ 2 (1 - \rho_{X,M}) + \frac{\mu_X^2}{\sigma_X^2} (1 - \rho_{X,M}^2) \right] \quad (\text{A.31})$$

Varianz von  $\widehat{MM}_{X,ori}$ ,  $(X, M)_t$  i.i.d. (Originalformulierung von MM)

$$\begin{aligned} n \sigma_{\widehat{MM}_{ori}}^2 &= \frac{\sigma_M^2}{\sigma_X^2} \left( \sigma_X^2 - \mu_X \sigma_X \gamma_{1,X} + \frac{\mu_X \sigma_{X,(M-\mu_M)}^2}{\sigma_M^2} \right) \\ &- \frac{\mu_X^2 \sigma_{(X-\mu_X)^2, (M-\mu_M)^2}}{2 \sigma_M^2 \sigma_X^2} + \frac{1}{4} \mu_X^2 (\gamma_{2,X} + \gamma_{2,M} - 2) \end{aligned} \quad (\text{A.32})$$

Varianz von  $\widehat{MM}_{X,ori}$ ,  $(X, M)_t$  i.i.d. normalverteilt

$$n \sigma_{\widehat{MM}_{ori, norm.}}^2 = \sigma_M^2 + \mu_X^2 \frac{\sigma_M^2}{\sigma_X^2} (1 - \rho_{X,M}^2) \quad (\text{A.33})$$

Varianz von  $\widehat{\Delta}_{MM}$  zweier Portfolios  $X$  und  $Y$ ,  $(X, Y, M)_t$  i.i.d.

$$\begin{aligned} n \sigma_{\widehat{\Delta}_{MM}}^2 &= 2 \sigma_M^2 + \frac{\sigma_M^2 \mu_Y \sigma_{X,(Y-\mu_Y)}^2}{\sigma_X \sigma_Y^3} - \frac{\mu_Y \sigma_{X,(M-\mu_M)}^2}{\sigma_X \sigma_Y} \\ &+ \frac{\mu_X \sigma_M^2 \sigma_{Y,(X-\mu_X)}^2}{\sigma_X^3 \sigma_Y} - \frac{\mu_X \sigma_{Y,(M-\mu_M)}^2}{\sigma_X \sigma_Y} - \frac{\mu_X \mu_Y \sigma_M^2 \sigma_{(X-\mu_X)^2, (Y-\mu_Y)^2}}{2 \sigma_X^3 \sigma_Y^3} \\ &+ \frac{\mu_X \mu_Y \sigma_{(X-\mu_X)^2, (M-\mu_M)^2}}{2 \sigma_Y \sigma_X^3} + \frac{\mu_X \mu_Y \sigma_{(Y-\mu_Y)^2, (M-\mu_M)^2}}{2 \sigma_X \sigma_Y^3} \\ &- \frac{\mu_X \mu_Y \sigma_{(M-\mu_M)^2}^2}{2 \sigma_X \sigma_Y \sigma_M^2} - \frac{2 \sigma_M^2 \sigma_{X,Y}}{\sigma_X \sigma_Y} - \frac{\mu_X \sigma_M^2 \sigma_{X,(X-\mu_X)}^2}{\sigma_X^4} \\ &+ \frac{\mu_X \sigma_{X,(M-\mu_M)}^2}{\sigma_X^2} - \frac{\mu_Y \sigma_M^2 \sigma_{Y,(Y-\mu_Y)}^2}{\sigma_Y^4} + \frac{\mu_Y \sigma_{Y,(M-\mu_M)}^2}{\sigma_Y^2} \\ &+ \frac{\mu_X^2 \sigma_M^2 \sigma_{(X-\mu_X)}^2}{4 \sigma_X^6} - \frac{\mu_X^2 \sigma_{(X-\mu_X)^2, (M-\mu_M)^2}}{2 \sigma_X^4} + \frac{\mu_Y^2 \sigma_M^2 \sigma_{(Y-\mu_Y)}^2}{4 \sigma_Y^6} \\ &- \frac{\mu_Y^2 \sigma_{(Y-\mu_Y)^2, (M-\mu_M)^2}}{2 \sigma_Y^4} + \frac{\mu_X^2 \sigma_{(M-\mu_M)}^2}{4 \sigma_X^2 \sigma_M^2} + \frac{\mu_Y^2 \sigma_{(M-\mu_M)}^2}{4 \sigma_Y^2 \sigma_M^2} \end{aligned} \quad (\text{A.34})$$

Varianz von  $\widehat{\Delta}_{MM}$  zweier Portfolios  $X$  und  $Y$ ,  $(X, Y, M)_t$  i.i.d. normalverteilt

$$\begin{aligned} n \sigma_{\widehat{\Delta}_{MM}, norm.}^2 &= \sigma_M^2 \left( 2(1 - \rho_{X,Y}) + \frac{\mu_X^2}{\sigma_X^2} (1 - \rho_{X,M}^2) \right. \\ &\quad \left. + \frac{\mu_Y^2}{\sigma_Y^2} (1 - \rho_{Y,M}^2) - \frac{\mu_X \mu_Y}{\sigma_X \sigma_Y} (1 + \rho_{X,Y}^2 - \rho_{X,M}^2 - \rho_{Y,M}^2) \right) \end{aligned} \quad (\text{A.35})$$

### A.2.5 Kappa

Varianz von  $\widehat{\kappa}_{\alpha,z}$

$$\begin{aligned} n \sigma_{\widehat{\kappa}_{\alpha,z}}^2 &= \kappa_{\alpha,z}^2 \left( \frac{E((X-z)^2)}{(E(X-z))^2} + \frac{2}{\alpha} \frac{E((z-X)_+^{\alpha+1})}{E(X-z)E((z-X)_+^\alpha)} \right. \\ &\quad \left. + \frac{1}{\alpha^2} \frac{E((z-X)_+^{2\alpha})}{(E((z-X)_+^\alpha))^2} - \left(1 - \frac{1}{\alpha}\right)^2 \right) \end{aligned} \quad (\text{A.36})$$

Darunter  $\widehat{\kappa}_{1,z}$  (identisch mit der Varianz von  $\Omega_{\alpha,z}$ , siehe Abschnitt A.3.1)

$$\begin{aligned} n \sigma_{\widehat{\kappa}_{1,z}}^2 &= \frac{E((X-z)^2)}{(E((z-X)_+))^2} + \frac{2E(X-z)E((z-X)_+^2)}{(E((z-X)_+))^3} \\ &\quad + \frac{(E(X-z))^2 E((z-X)_+^2)}{(E((z-X)_+))^4} \end{aligned} \quad (\text{A.37})$$

Darunter  $\widehat{\kappa}_{2,z}$  (Sortino-Ratio)

$$\begin{aligned} n \sigma_{\widehat{\kappa}_{2,z}}^2 &= \frac{E((X-z)^2)}{E((z-X)_+^2)} + \frac{E(X-z)E((z-X)_+^3)}{(E((z-X)_+^2))^2} \\ &\quad + \frac{(E(X-z))^2 E((z-X)_+^4)}{4(E((z-X)_+^2))^3} - \frac{(E(X-z))^2}{4E((z-X)_+^2)} \end{aligned} \quad (\text{A.38})$$

Darunter  $\widehat{\kappa}_{3,z}$

$$\begin{aligned} n \sigma_{\widehat{\kappa}_{3,z}}^2 &= \kappa_{3,z}^2 \left( \frac{E((X-z)^2)}{(E(X-z))^2} + \frac{2}{3} \frac{E((z-X)_+^4)}{E(X-z)E((z-X)_+^3)} \right. \\ &\quad \left. + \frac{1}{6} \frac{E((z-X)_+^6)}{(E((z-X)_+^3))^2} - \frac{4}{9} \right) \end{aligned} \quad (\text{A.39})$$

### A.2.6 ERVaR

Varianz von  $\widehat{ERVaR}_X$ :

$$\begin{aligned} n \sigma_{\widehat{ERVaR}_X}^2 &= ERVaR_X^2 \left( \frac{\sigma_X^2}{\mu_X^2} - \frac{E(|X - x_\alpha|) + (2\alpha - 1)(\mu_X - x_\alpha)}{\mu_X x_\alpha f(x_\alpha)} \right. \\ &\quad \left. + \frac{\alpha(1-\alpha)}{(x_\alpha f(x_\alpha))^2} \right) \end{aligned} \quad (\text{A.40})$$

### A.3 Weitere, ergänzende Berechnungen

#### A.3.1 Überleitung von $\sigma_{\widehat{\kappa}_{1,z}}^2$ zu $\sigma_{\widehat{\Omega}_z}^2$

Nachweis, dass Formel 4.9 auf S. 90 über die Varianz von  $\widehat{\kappa}_{1,z}$  in die aus Schmid und Schmidt (2008) stammende Formel 4.10 über die Varianz von  $\widehat{\Omega}_z$  überführbar ist. Formel 4.9 lautet:

$$n\sigma_{\widehat{\kappa}_{1,z}}^2 = \frac{(E(X-z))^2}{(E((z-X)_+))^2} \left( \frac{E((X-z)^2)}{(E(X-z))^2} + \frac{2E((z-X)_+^2)}{E(X-z)E((z-X)_+)} + \frac{E((z-X)_+^2)}{(E((z-X)_+))^2} \right)$$

Hieraus ergibt sich durch Umformungen:

$$\begin{aligned} &= \frac{E((X-z)^2)}{(E((z-X)_+))^2} + \frac{2E(X-z)E((z-X)_+^2)}{(E((z-X)_+))^3} + \frac{(E(X-z))^2E((z-X)_+^2)}{(E((z-X)_+))^4} \\ &= \underbrace{\left( \frac{E(X-z)_+}{E(z-X)_+} \right)^2}_{\Omega^2} \cdot \left( \frac{E((X-z)^2)}{(E(X-z)_+)^2} + 2 \frac{E(X-z)E((z-X)_+^2)}{(E(X-z)_+)^2(E(z-X)_+)} \right. \\ &\quad \left. + \frac{(E((X-z)^2)E((z-X)_+^2))}{(E((X-z)_+))^2(E((z-X)_+))^2} \right) \\ &= \Omega^2 \left( \frac{E(((X-z)_+ - (z-X)_+)^2)}{(E(X-z)_+)^2} + \frac{2E((X-z)_+ - (z-X)_+)E((z-X)_+^2)}{(E(X-z)_+)^2E(z-X)_+} \right. \\ &\quad \left. + \frac{E((X-z)_+ - (z-X)_+)^2E((z-X)_+^2)}{E((X-z)_+)^2E((z-X)_+)^2} \right) \\ &= \Omega^2 \left( \frac{E((X-z)_+^2) + E((z-X)_+^2)}{(E(X-z)_+)^2} + \frac{2E((z-X)_+^2)}{(E(X-z)_+)^2} \cdot \frac{\overbrace{E(X-z)_+ - E(z-X)_+}^{\Omega-1}}{E(z-X)_+} \right. \\ &\quad \left. + \frac{E((z-X)_+^2)}{E((X-z)_+)^2} \cdot \frac{(E(X-z)_+)^2 - 2E(X-z)_+E(z-X)_+ + (E(z-X)_+)^2}{\underbrace{(E(z-X)_+)^2}_{\Omega^2 - 2\Omega + 1}} \right) \\ &= \Omega^2 \left( \frac{E((X-z)_+^2)}{(E(X-z)_+)^2} + \frac{E((z-X)_+^2)}{(E(X-z)_+)^2} + \frac{E(z-X)_+^2}{(E(X-z)_+)^2} \underbrace{(\Omega^2 + 2\Omega - 2\Omega - 2 + 1)}_{\Omega^2 - 1} \right) \end{aligned}$$

$$\begin{aligned}
&= \Omega^2 \left( \frac{E((X-z)_+^2)}{(E(X-z)_+)^2} + \frac{E((z-X)_+^2)}{(E(X-z)_+)^2} + \frac{E((z-X)_+^2)(E(X-z)_+)^2}{(E(X-z)_+)^4} \right. \\
&\quad \left. - \frac{E((z-X)_+^2)}{E((X-z)_+)^2} \right) \\
&= \Omega^2 \left( \frac{E(X-z)_+^2}{(E(X-z)_+)^2} + \frac{E((z-X)_+^2)}{(E(X-z)_+)^2} \right) = \sigma_{\Omega_z}^2 \quad (\text{siehe Formel 4.10, S. 90})
\end{aligned}$$

### A.3.2 Erläuterung von Formel 5.1, S. 113

Sei  $F(x)$  eine (stetige) Prüfverteilung mit Quantilfunktion  $F^{-1}(p)$  und  $T$  eine Testgröße, für die unter  $H_0$  gilt:

$$T \sim F.$$

Wird ein zweiseitiger Test zum Niveau  $\alpha$  durchgeführt, lautet die Entscheidungsanweisung:

$$\text{Lehne } H_0 \text{ ab, wenn } T \notin \left[ F^{-1}\left(\frac{\alpha}{2}\right); F^{-1}\left(1 - \frac{\alpha}{2}\right) \right].$$

Der  $p$ -Wert  $p_w(T)$  ist der Wert für  $\alpha$  in obiger Formel, bei dem  $T$  gerade auf der Grenze des kritischen Bereiches liegen würde. Bei zweiseitigem Test:

$$T = \begin{cases} F^{-1}\left(\frac{p_w}{2}\right) & \text{falls } F(T) \leq 0,5 \\ F^{-1}\left(1 - \frac{p_w}{2}\right) & \text{falls } F(T) > 0,5 \end{cases} \quad | \quad F(\dots)$$

Nun wird nur noch umgeformt:

$$\begin{aligned}
F(T) &= \begin{cases} \frac{p_w}{2} & \text{falls } F(T) \leq 0,5 \\ 1 - \frac{p_w}{2} & \text{falls } F(T) > 0,5 \end{cases} \\
p_w &= \begin{cases} 2 \cdot F(T) & \text{falls } F(T) \leq 0,5 \\ 2 - 2 \cdot F(T) & \text{falls } F(T) > 0,5 \end{cases} \\
p_w &= \begin{cases} (1 - 2 \cdot 0,5) + 2 \cdot F(T) & \text{falls } F(T) \leq 0,5 \\ (1 + 2 \cdot 0,5) - 2 \cdot F(T) & \text{falls } F(T) > 0,5 \end{cases} \\
p_w &= \begin{cases} 1 - 2 \cdot (0,5 - F(T)) & \text{falls } F(T) \leq 0,5 \\ 1 - 2 \cdot (F(T) - 0,5) & \text{falls } F(T) > 0,5 \end{cases} \\
p_w &= 1 - 2 \cdot |0,5 - F(T)|
\end{aligned}$$

## Literaturverzeichnis

- Agarwal, V., Daniel, N., Naik, N. (2007): Why is Santa so kind to hedge funds? The December return puzzle!, CFR Working Paper No. 07-09, Köln.
- Agarwal, V., Naik, N. (2004): Risks and portfolio decisions involving hedge funds., *Review of Financial Studies*, 17(1), S. 63–98.
- Aggarwal, R. K., Jorion, P. (2010): The performance of emerging hedge funds and managers, *Journal of Financial Economics*, 96(2), S. 238–256.
- Aiken, A. L., Clifford, C. P., Ellis, J. (2012): Out of the dark: Hedge fund reporting bias and commercial databases, *Review of Financial Studies*, 26(1), S. 208–243.
- Albrecht, P. (2003): Zur Messung von Finanzrisiken, *Mannheimer Manuskripte zu Risikotheorie, Portfolio Management und Versicherungswirtschaft*, (1/2003).
- Anderson, T. W. (1984): *An introduction to multivariate statistical analysis*, Wiley, New York, 2. Auflage.
- Anderson, T. W., Darling, D. A. (1954): A test of goodness of fit, *Journal of the American Statistical Association*, 49(268), S. 765–769.
- Andrews, D. W. K. (1991): Heteroskedasticity and autocorrelation consistent covariance matrix estimation, *Econometrica*, 59(3), S. 817–858.
- Andrews, D. W. K., Monahan, J. C. (1992): An improved heteroskedasticity and autocorrelation consistent covariance matrix estimator, *Econometrica*, 60(4), S. 953–966.
- Artzner, P., Delbaen, F., Eber, J.-M., Heath, D. (1999): Coherent measures of risk, *Mathematical Finance*, 9(3), S. 203–228.
- Auer, L. v. (2007): *Ökonometrie*, Springer, Berlin, 4. Auflage.
- Bacon, C. (2004): *Practical portfolio performance measurement and attribution*, Wiley, Chichester.
- Bacon, C. (2013): *Practical risk-adjusted performance measurement*, Wiley, Hoboken, NJ.
- BaFin (2011): Liste der genehmigten inländischen Sondervermögen mit zusätzlichen Risiken und Investmentaktiengesellschaften mit § 112 InvG vergleichbaren Teilgesellschaftsvermögen, Bundesamt für Finanzdienstleistungsaufsicht, Frankfurt am Main.
- Barras, L., Scaillet, O., Wermers, R. (2010): False discoveries in mutual fund performance: Measuring luck in estimated alphas, *The Journal of Finance*, 65, S. 179–216.

- Bauwens, L., Laurent, S., Rombouts, J. V. K. (2006): Multivariate GARCH models, *Journal of Applied Econometrics*, 21(1), S. 79–109.
- Bohrnstedt, G. W., Goldberger, A. S. (1969): On the exact covariance of products of random variables, *Journal of the American Statistical Association*, 64(328), S. 1439–1442.
- Bollen, N., Pool, V. (2009): Do hedge fund managers misreport returns? Evidence from the pooled distribution, *The Journal of Finance*, 64(5), S. 2257–2288.
- Bollerslev, T. (1986): Generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics*, 31(3), S. 307–327.
- Bollerslev, T. (1990): Modelling the coherence in short-run nominal exchange rates: A multivariate generalized ARCH model, *The Review of Economics and Statistics*, 72(3), S. 498–505.
- Bollerslev, T., Engle, R. F., Wooldridge, J. M. (1988): A capital asset pricing model with time-varying covariances., *Journal of Political Economy*, 96(1), S. 116–131.
- Brailsford, T. J., Faff, R. W. (1996): An evaluation of volatility forecasting techniques, *Journal of Banking & Finance*, 20(3), S. 419–438.
- Brooks, C., Burke, S. P. (2003): Information criteria for GARCH model selection., *European Journal of Finance*, 9(6), S. 557–580.
- Brooks, C., Kat, H. M. (2002): The statistical properties of hedge fund index returns and their implications for investors, *Journal of Alternative Investments*, 5(2), S. 26.
- Brown, S. J., Goetzmann, W. N. (1995): Performance persistence, *Journal of Finance*, 50(2), S. 679–698.
- Cadsby, C. B. (1986): Performance hypothesis testing with the Sharpe and Treynor measures: A comment, *The Journal of Finance*, 41(5), S. 1175–1176.
- Carhart, M. M. (1997): On persistence in mutual fund performance, *Journal of Finance*, 52(1), S. 57–82.
- Cherubini, U., Luciano, E., Vecchiato, W. (2004): *Copula methods in finance*, Wiley, Hoboken, NJ.
- Christie, A. A. (1982): The stochastic behavior of common stock variances: Value, leverage and interest rate effects, *Journal of Financial Economics*, 10(4), S. 407–432.
- Christopherson, J., Carino, D., Ferson, W. (2009): *Portfolio performance measurement and benchmarking*, McGraw-Hill, New York.

- Cici, G., Kempf, A., Pütz, A. (2010): Caught in the act: How hedge funds manipulate their equity positions, CFR Working Paper No 10-15, Köln.
- Clark, E., Kassimatis, K. (2011): An alternative measure of the "world market portfolio": Determinants, efficiency, and information content, *Journal of International Money and Finance*, 30(5), S. 724–748.
- Cogneau, P., Hübner, G. (2009a): The (more than) 100 ways to measure portfolio performance. Part 1: Standardized risk-adjusted measures, *Journal of Performance Measurement*, 13(4), S. 56–71.
- Cogneau, P., Hübner, G. (2009b): The (more than) 100 ways to measure portfolio performance. Part 2: Special measures and comparison, *Journal of Performance Measurement*, 14(1), S. 56–69.
- Cont, R. (2001): Empirical properties of asset returns: Stylized facts and statistical issues., *Quantitative Finance*, 1(2), S. 223–236.
- Cryer, J., Chan, K.-S. (2008): *Time series analysis*, Springer, Berlin, 2. Auflage.
- Cumby, R. E., Glen, J. D. (1990): Evaluating the performance of international mutual funds, *The Journal of Finance*, 45(2), S. 497–521.
- Curto, J. D., Pinto, J. C., Tavares, G. N. (2009): Modeling stock markets' volatility using GARCH models with normal, Student's t and stable paretian distributions., *Statistical Papers*, 50(2), S. 311–321.
- Davison, A. C., Hinkley, D. V. (1997): *Bootstrap methods and their application*, Cambridge University Press, Cambridge.
- Deutsch, H.-P. (2008): *Derivate und interne Modelle*, Schäffer-Poeschel, Stuttgart, 4. Auflage.
- Dichev, I. D., Yu, G. (2011): Higher risk, lower returns: What hedge fund investors really earn, *Journal of Financial Economics*, 100(2), S. 248–263.
- Dickey, D. A., Fuller, W. A. (1981): Likelihood ratio statistics for autoregressive time series with a unit root, *Econometrica*, 49(4), S. 1057–1072.
- Diebold, F. X., Nerlove, M. (1989): The dynamics of exchange rate volatility: A multivariate latent factor ARCH model, *Journal of Applied Econometrics*, 4(1), S. 1–21.
- Dowd, K. (2000): Adjusting for risk: An improved Sharpe ratio., *International Review of Economics & Finance*, 9(3), S. 209–222.
- Drobny, S. (2010): *The invisible hands: Hedge funds off the record*, Wiley, Hoboken.
- DuMouchel, W. H. (1983): Estimating the stable index alpha in order to measure tail thickness: A critique, *The Annals of Statistics*, 11(4), S. 1019–1031.

- Eberlein, E., Keller, U. (1995): Hyperbolic distributions in finance, *Bernoulli*, 1(3), S. 281–299.
- Edwards, F. B., Caglayan, M. O. (2001): Hedge fund performance and manager skill, *Journal of Futures Markets*, 21(11), S. 1003–1028.
- Efron, B. (1979): Bootstrap methods: Another look at the jackknife, *The Annals of Statistics*, 7(1), S. 1–26.
- Efron, B. (1987): Better bootstrap confidence intervals, *Journal of the American Statistical Association*, 82(397), S. 171–185.
- Efron, B., Tibshirani, R. (1993): *An introduction to the bootstrap*, Chapman & Hall, New York.
- Eisenführ, F., Weber, M., Langer, T. (2010): *Rationales Entscheiden*, Springer, Berlin, 5. Auflage.
- Eling, M. (2008): Does the measure matter in the mutual fund industry?, *Financial Analysts Journal*, 64(3), S. 54–66.
- Eling, M. (2009): Does hedge fund performance persist? Overview and new empirical evidence, *European Financial Management*, 15(2), S. 362–401.
- Eling, M., Schuhmacher, F. (2007): Does the choice of performance measure influence the evaluation of hedge funds?, *Journal of Banking & Finance*, 31(9), S. 2632–2647.
- Engle, R. F. (1982): Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation, *Econometrica*, 50(4), S. 987–1007.
- Engle, R. F. (1987): Multivariate ARCH with factor structures: Cointegration in variance, mimeo, University of California Press.
- Engle, R. F., Bollerslev, T. (1986): Modelling the persistence of conditional variances, *Econometric Reviews*, 5(1), S. 1–50.
- Engle, R. F., Kroner, K. F. (1995): Multivariate simultaneous generalized ARCH, *Econometric Theory*, 11(1), S. 122–150.
- Engle, R. F., Ng, V. K., Rothschild, M. (1990): Asset pricing with a factor-ARCH covariance structure: Empirical estimates for treasury bills, *Journal of Econometrics*, 45(1-2), S. 213–237.
- Everitt, B. S. (1982): Mixture distributions, in: *Encyclopedia of Statistical Sciences*, vol. 5, S. 559–569, Wiley, New York.
- Fabozzi, F. J., Mann, S. V. (Hrsg.) (2005): *The handbook of fixed income securities*, McGraw-Hill, New York, 7. Auflage.



- Fama, E. F., French, K. R. (2010): Luck versus skill in the cross-section of mutual fund returns, *Journal of Finance*, 65(5), S. 1915–1947.
- Favre, L., Galeano, J.-A. (2002): Mean-modified value-at-risk optimization with hedge funds., *Journal of Alternative Investments*, 5(2), S. 21–25.
- Ferguson, T. S. (1999): Asymptotic joint distribution of sample mean and a sample quantile, URL <http://www.math.ucla.edu/~tom/papers/unpublished/meanmed.pdf>.
- Fischer, B. (2010): *Performanceanalyse in der Praxis*, Oldenbourg, München, 3. Auflage.
- Franke, J., Härdle, W., Hafner, C. (2001): *Einführung in die Statistik der Finanzmärkte*, Springer, Berlin.
- Frees, E., Valdez, E. (1998): Understanding relationships using copulas, *North American Actuarial Journal*, 2(1), S. 1–25.
- French, K. R., Schwert, G. W., Stambaugh, R. F. (1987): Expected stock returns and volatility, *Journal of Financial Economics*, 19(1), S. 3–29.
- Fung, H.-G., Xu, X. E., Yau, J. (2002): Global hedge funds: Risk, return, and market timing, *Financial Analysts Journal*, 58(6), S. 19–30.
- Fung, W., Hsieh, D. A. (2000): Performance characteristics of hedge funds and commodity funds: Natural vs. spurious biases, *Journal of Financial & Quantitative Analysis*, 35(3), S. 291–307.
- Fung, W., Hsieh, D. A. (2009): Measurement biases in hedge fund performance data: An update, *Financial Analysts Journal*, 65(3), S. 1–3.
- Fung, W. K. H., Hsieh, D. A. (2006): Hedge funds: An industry in its adolescence, *Economic Review*, 91(4), S. 1–34.
- Géhin, W. (2004): A survey of the literature of hedge fund performance, URL <http://ssrn.com/abstract=626441>, EDHEC Business School.
- Getmansky, M., Lo, A. W., Makarov, I. (2004): An econometric model of serial correlation and illiquidity in hedge fund returns, *Journal of Financial Economics*, 74(3), S. 529–609.
- Glosten, L., Jagannathan, R., Runkle, D. E. (1993): On the relation between the expected value and the volatility of the nominal excess return on stocks, *Journal of Finance*, 48(5), S. 1779–1801.
- Gourieroux, C., Jasiak, J. (2001): *Financial Econometrics: Problems, Models, and Methods*, Princeton University Press, Princeton.
- Greuc, A., Malkiel, B. G., Saha, A. (2007): Why do hedge funds stop reporting performance?, *Journal of Portfolio Management*, 34(1), S. 119–126.

- Greene, W. (2008): *Econometric analysis*, Pearson Prentice Hall, Upper Saddle River, NJ, 6. Auflage.
- Gregoriou, G. N., Gueyie, J.-P. (2003): Risk-adjusted performance of funds of hedge funds using a modified Sharpe ratio, *Journal of Wealth Management*, 6(3), S. 77–83.
- Gumbel, E. J. (1939): La probabilité des hypothèses, *Comptes rendus hebdomadaires des séances de l'Académie des Sciences*, 209, S. 645–647.
- Hansen, P. R., Lunde, A. (2005): A forecast comparison of volatility models: Does anything beat a GARCH(1,1)?, *Journal of Applied Econometrics*, 20(7), S. 873–889.
- Hendry, D., Nielsen, B. (2007): *Econometric modeling. A Likelihood Approach*, Princeton University Press, Princeton.
- HFR (2009): HFR industry report Q1 2009.
- Hull, J. C. (2011): *Options, Futures and other Derivatives*, Prentice Hall, Upper Saddle River, NJ, 8. Auflage.
- Hyndman, R. J., Fan, Y. (1996): Sample quantiles in statistical packages, *The American Statistician*, 50(4), S. 361–365.
- Ibbotson, R. G., Chen, P., Zhu, K. X. (2011): The ABCs of hedge funds: Alphas, betas, and costs, *Financial Analysts Journal*, 67(1), S. 15–25.
- Ingersoll, J., Spiegel, M., Goetzmann, W., Welch, I. (2007): Portfolio performance manipulation and manipulation-proof performance measures, *The Review of Financial Studies*, 20(5), S. 1503–1546.
- Isserlis, L. (1918): On a formula for the product-moment coefficient of any order of a normal frequency distribution in any number of variables, *Biometrika*, 12(1/2), S. 134–139.
- Jarque, C. M., Bera, A. K. (1980): Efficient tests for normality, homoscedasticity and serial independence of regression residuals, *Economics Letters*, 6(3), S. 255–259.
- Jensen, M. C. (1968): The performance of mutual funds in the period 1945-1964, *The Journal of Finance*, 23(2), S. 389–416.
- Jobson, J. D., Korkie, B. M. (1981): Performance hypothesis testing with the Sharpe and Treynor measures, *The Journal of Finance*, 36(4), S. 889–908.
- Joe, H. (1997): *Multivariate Models and Dependence Concepts*, Chapman & Hall, London.
- Jondeau, E., Poon, S.-H., Rockinger, M. (2007): *Financial modeling under non-Gaussian distributions*, Springer, London.

- Jorion, P. (2007): *Value at risk*, McGraw-Hill, New York, 3. Auflage.
- Kaiser, D. (2009): *Hedgefonds - Entmystifizierung einer Anlagekategorie*, Gabler, Wiesbaden, 2. Auflage.
- Kaplan, P., Knowles, J. (2004): Kappa: A generalized downside risk-adjusted performance measure, *Journal of Performance Measurement*, 8(3), S. 42–54.
- Karian, Z., Dudewicz, E. (2000): *Fitting Statistical Distributions*, Chapman & Hall, Boca Raton.
- Kat, H. M. (2003): 10 things that investors should know about hedge funds, *Journal of Wealth Management*, 5(4), S. 72–81.
- Keating, C., Shadwick, W. (2002): A universal performance measure, *Journal of Performance Measurement*, 6(3), S. 59–84.
- Kijima, M., Ohnishi, M. (1993): Mean-risk analysis of risk aversion and wealth effects on optimal portfolios with multiple investment opportunities, *Annals of Operations Research*, 45(1-4), S. 147–163.
- Kosowski, R., Naik, N. Y., Teo, M. (2007): Do hedge funds deliver alpha? A bayesian and bootstrap analysis, *Journal of Financial Economics*, 84(1), S. 229–264.
- Kosowski, R., Timmermann, A., Wermers, R., White, H. (2006): Can mutual fund “Stars” really pick stocks? New evidence from a bootstrap analysis, *Journal of Finance*, 61(6), S. 2551–2595.
- Kotz, S., Nadarajah, S. (2004): *Multivariate t distributions and their applications*, Cambridge University Press, Cambridge.
- Le Sourd, V. (2007): Performance measurement for traditional investment - literature survey, EDHEC risk and asset management centre.
- Ledoit, O., Wolf, M. (2008): Robust performance hypothesis testing with the Sharpe ratio, *Journal of Empirical Finance*, 15(5), S. 850–859.
- Lehmann, E. L. (1999): *Elements of Large-Sample Theory*, Springer, New York.
- Lhabitant, F.-S. (2002): *Hedge funds*, Wiley, Chichester.
- Ljung, G. M., Box, G. (1979): The likelihood function of stationary autoregressive-moving average models, *Biometrika*, 66(2), S. 265–270.
- Lo, A. W. (2008): *Hedge funds: An analytic perspective*, Princeton University Press, Princeton.
- Loretan, M., Phillips, P. C. B. (1994): Testing the covariance stationarity of heavy-tailed time series: An overview of the theory with applications to several financial datasets, *Journal of Empirical Finance*, 1(2), S. 211–248.

- Lütkepohl, H. (2005): *New Introduction to Multiple Time Series Analysis*, Springer, Berlin.
- Malkiel, B. G., Saha, A. (2005): Hedge funds: Risk and return, *Financial Analysts Journal*, 61(6), S. 80–88.
- Mammen, E. (1992): *When does bootstrap work? Asymptotic results and simulations*, Springer, New York.
- Manner, H. (2010): *Modeling Asymmetric and Time-Varying Dependence*, Ph.D. thesis, Maastricht University, Maastricht.
- McLachlan, G., Basford, K. (1988): *Mixture models: inference and applications to clustering*, Dekker, New York.
- McNeil, A., Frey, R., Embrechts, P. (2005): *Quantitative risk management*, Princeton University Press, Princeton.
- Memmel, C. (2003): Performance hypothesis testing with the Sharpe ratio, *Finance Letters*, 1(1), S. 21–23.
- Meyer, J. (1987): Two-moment decision models and expected utility maximization, *The American Economic Review*, 77(3), S. 421–430.
- Meyer, J., Rasche, R. H. (1992): Sufficient conditions for expected utility to imply mean-standard deviation rankings: Empirical evidence concerning the location and scale condition, *Economic Journal*, 102(410), S. 91–106.
- Miller, R. E., Gehr, A. K. (1978): Sample size bias and Sharpe's performance measure: A note, *Journal of Financial & Quantitative Analysis*, 13(5), S. 943–946.
- Mills, T., Markellos, R. (2008): *The econometric modelling of financial time series*, Cambridge University Press, Cambridge, 3. Auflage.
- Modigliani, F., Modigliani, L. (1997): Risk-adjusted performance, *Journal of Portfolio Management*, 23(2), S. 45–54.
- Morey, M. R., Morey, R. C. (2000): An analytical confidence interval for the Treynor index: Formula, conditions and properties, *Journal of Business Finance & Accounting*, 27(1/2), S. 127–154.
- Nelsen, R. (2006): *An introduction to copulas*, Springer, New York, 2. Auflage.
- Nelson, D. B. (1991): Conditional heteroskedasticity in asset returns: A new approach, *Econometrica*, 59(2), S. 347–370.
- Pagan, A. (1996): The econometrics of financial markets, *Journal of Empirical Finance*, 3(1), S. 15–102.

- Patton, A. (2009): Copula-Based models for financial time series, in: Andersen, T., Davis, R., Kreiß, J.-P., Mikosch, T. (Hrsg.), *Handbook of Financial Time Series*, S. 767–785, Springer, Berlin.
- Pedersen, C. S., Satchell, S. E. (1998): An extended family of financial risk measures, *Geneva Papers on Risk and Insurance*, 23, S. 89–117.
- Pedersen, C. S., Satchell, S. E. (2000): Small sample analysis of performance measures in the asymmetric response model, *The Journal of Financial and Quantitative Analysis*, 35(3), S. 425–450.
- Posthuma, N., van der Sluis, P. J. (2003): A reality check on hedge funds returns, URL <http://ssrn.com/paper=438840>, Working Paper.
- Rachev, S., Mittnik, S. (2000): *Stable paretian models in finance*, Wiley, Chichester.
- Rachev, S., Stoyanov, S., Biglova, A., Fabozzi, F. (2005): An empirical examination of daily stock return distributions for U.S. stocks, in: Baier, D., Decker, R., Schmidt-Thieme, L. (Hrsg.), *Data Analysis and Decision Support*, S. 269–281, Technical Reports, Karlsruher Institut für Technologie.
- Read, O. (1998): *Parametrische Modelle zur Ermittlung des Value-at-Risk*, Ph.D. thesis, Universität zu Köln, Köln.
- Rockafellar, T., Uryasev, S., Zabarankin, M. (2002): Deviation measures in risk analysis and optimization, URL <http://ssrn.com/paper=365640>.
- Rydberg, T. H. (2000): Realistic statistical modelling of financial data, *International Statistical Review / Revue Internationale de Statistique*, 68(3), S. 233–258.
- Sachs, L., Hedderich, J. (2006): *Angewandte Statistik*, Springer, Berlin, 12. Auflage.
- Scherer, B. (2004): An alternative route to performance hypothesis testing, *Journal of Asset Management*, 5(1), S. 5–12.
- Schlittgen, R. (1996): *Einführung in die Statistik*, Oldenbourg, München, 6. Auflage.
- Schlittgen, R., Streitberg, B. (2001): *Zeitreihenanalyse*, Oldenbourg, München, 9. Auflage.
- Schmid, F., Schmidt, A. (2006): Nonparametric estimation of the coefficient of overlapping - theory and empirical application, *Computational Statistics & Data Analysis*, 50(6), S. 1583 – 1596.
- Schmid, F., Schmidt, R. (2008): Statistical inference for performance measure Omega, Unveröffentlichtes Manuskript, Köln.

- Schmid, F., Schmidt, R. (2010): Statistical inference for Sharpe's ratio, in: Berkelaar, A. B., Coche, J., Nyholm, K. (Hrsg.), *Interest Rate Models, Asset Allocation and Quantitative Techniques for Central Banks and Sovereign Wealth Funds*, S. 337–357, Palgrave Macmillan, Basingstoke.
- Schmid, F., Trede, M. (2006): *Finanzmarktstatistik*, Springer, Berlin.
- Schuhmacher, F., Eling, M. (2012): A decision-theoretic foundation for reward-to-risk performance measures, *Journal of Banking & Finance*, 36(7), S. 2077–2082.
- Scott, R., Horvath, P. (1980): On the direction of preference for moments of higher order than the variance, *Journal of Finance*, 35(4), S. 915–919.
- Sentana, E. (1998): The relation between conditionally heteroskedastic factor models and factor GARCH models, *Econometrics Journal*, 1(2), S. 1–9, URL <http://dx.doi.org/10.1111/1368-423X.12014>.
- Sewell, M. (2011a): Characterization of financial time series, Research Note RN/11/01, University College London.
- Sewell, M. (2011b): Fund performance, Research Note RN/11/03, University College London.
- Sharpe, W. F. (1966): Mutual fund performance, *Journal of Business*, 39(1), S. 119.
- Sharpe, W. F. (1994): The Sharpe ratio, *Journal of Portfolio Management*, 21(1), S. 49–58.
- Silvennoinen, A., Teräsvirta, T. (2009): Multivariate GARCH models, in: Andersen, T. (ed.), *Handbook of financial time series*, S. 201–232, Springer, Berlin.
- Silverman, B. W. (1998): *Density estimation for statistics and data analysis*, Chapman & Hall, London, Neudruck der 1. Auflage.
- Sinn, H.-W. (1989): *Economic Decisions under Uncertainty*, Physica, Heidelberg, 2. Auflage.
- Sklar, A. (1959): Fonctions de répartition à n dimensions et leurs marges, *Publications de l'Institut de Statistique de l'Université de Paris*, 8.
- Sortino, F. A., van der Meer, R. (1991): Downside risk, *Journal of Portfolio Management*, 17(4), S. 27–31.
- Spanos, A. (1999): *Probability Theory and Statistical Inference*, Cambridge University Press, Cambridge.
- Steiner, P., Uhler, H. (2001): *Wertpapieranalyse*, Physica, Heidelberg, 4. Auflage.
- Straumann, D. (2009): Measuring the quality of hedge fund data, *Journal of Alternative Investments*, 12(2), S. 26–40.

- Stuart, A., Ord, J. K. (1987): *Kendall's Advanced Theory of Statistics, Volume 1: Distribution Theory*, Charles Griffin, London, 5. Auflage.
- Taylor, S. (2005): *Asset price dynamics, volatility, and prediction*, Princeton University Press, Princeton.
- Teräsvirta, T. (2009): An introduction to univariate GARCH models, in: Andersen, T., Davis, R., Kreiß, J.-P., Mikosch, T. (Hrsg.), *Handbook of financial time series*, S. 17–42, Springer, Berlin.
- Thode, H. C. (2002): *Testing for Normality*, Dekker, New York.
- Treynor, J. L. (1965): How to rate management of investment funds, *Harvard Business Review*, 43(1), S. 63–75.
- Vinod, H., Morey, M. R. (2000): Confidence intervals and hypothesis testing for the Sharpe and Treynor performance measures: A bootstrap approach, in: Abu-Mostafa, Y., LeBaron, B., Lo, A., Weigend, A. (Hrsg.), *Computational Finance 1999*, S. 25–39, MIT Press, Cambridge.
- Vrontos, I. D., Meligkotsidou, L., Vrontos, S. D. (2011): Performance evaluation of mutual fund investments: The impact of non-normality and time-varying volatility, *Journal of Asset Management*, 12(4), S. 292–307.
- Wand, M. P., Jones, M. C. (1995): *Kernel Smoothing*, Chapman & Hall, London.
- Wilkins, M., Scholz, H. (1999): Von der Treynor-Ratio zur Market Risk-Adjusted Performance, *Finanz-Betrieb*, 10/1999, S. 308–315.
- Wittrock, C. (2000): *Messung und Analyse der Performance von Wertpapierportfolios*, Uhlenbruch, Bad Soden, 3. Auflage.
- Xu, X. E., Liu, J., Loviscek, A. (2010): Hedge fund attrition, survivorship bias, and performance: Perspectives from the global financial crisis, URL <http://ssrn.com/paper=1572116>, Working Paper.
- Zakoian, J.-M. (1994): Threshold heteroskedastic models, *Journal of Economic Dynamics and Control*, 18(5), S. 931–955.

**Für diese Arbeit genutzte R-Pakete:**

- *boot*, Version 1.2-43, Autoren: Angelo Canty, Brian Ripley
- *ellipse*, Version 0.3-7 Autoren: Duncan Murdoch, E. D. Chow
- *fGarch*, Version 2110.80, Autoren: Diethelm Wuertz, Yohan Chalabi, Michal Miklovic, Chris Boudt, Pierre Chausse et al.
- *fUtilities*, Version 2110.78, Autoren: Diethelm Wuertz et al.
- *micEcon*, Version 0.6-7, Autor: Arne Henningsen
- *moments*, Version 0.11, Autor: Lukasz Komsta, Frederick Novomestky
- *mvtnorm*, Version 0.9-92, Autoren: Alan Genz, Frank Bretz, Tetsuhisa Miwa, Xuefei Mi, Friedrich Leisch, Fabian Scheipl, Torsten Hothorn
- *nortest*, Version 1.0, Autor: Juergen Gross