

Mehmke, Fabian; Cremers, Heinz; Packham, Natalie

Working Paper

Validierung von Konzepten zur Messung des Marktrisikos: Insbesondere des Value at Risk und des Expected Shortfall

Frankfurt School - Working Paper Series, No. 192

Provided in Cooperation with:

Frankfurt School of Finance and Management

Suggested Citation: Mehmke, Fabian; Cremers, Heinz; Packham, Natalie (2012) : Validierung von Konzepten zur Messung des Marktrisikos: Insbesondere des Value at Risk und des Expected Shortfall, Frankfurt School - Working Paper Series, No. 192, Frankfurt School of Finance & Management, Frankfurt a. M.

This Version is available at:

<https://hdl.handle.net/10419/60503>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Frankfurt School - Working Paper Series

No. 192

**Validierung von Konzepten zur Messung des
Marktrisikos - insbesondere des Value at
Risk und des Expected Shortfall**

by Fabian Mehmke, Heinz Cremers, Natalie Packham

June 2012



Frankfurt School of
Finance & Management
Bankakademie | HfB

Sonnemannstr. 9-11 60314 Frankfurt am Main, Germany

Phone: +49 (0)69 154 008 0 Fax: +49 (0)69 154 008 728

Internet: www.frankfurt-school.de

Abstract

Market risk management is one of the key factors to success in managing financial institutions. Underestimated risk can have disastrous consequences for individual companies and even whole economies, not least as could be seen during the recent crises. Overestimated risk, on the other side, may have negative effects on a company's capital requirements. Companies as well as national authorities thus have a strong interest in developing market risk models that correctly quantify certain key figures such as Value at Risk or Expected Shortfall.

This paper presents several state of the art methods to evaluate the adequacy of almost any given market risk model. Existing models are enhanced by in-depth analysis and simulations of statistical properties revealing some previously unknown effects, most notably inconsistent behaviour of alpha and beta errors. Furthermore, some new market risk validation models are introduced. In the end, a simulation with various market patterns demonstrates strengths and weaknesses of each of the models presented under realistic conditions.

Key words: Backtesting, Market Risk, Value at Risk, Expected Shortfall, Validation, Alpha Error, Beta Error, Time Until First Failure, Proportion of Failure, Traffic Light Approach, Magnitude of Loss Function, Markow-Test, Gauss-Test, Rosenblatt, Kuiper, Kolmogorov-Smirnov, Jarque-Bera, Regression, Likelihood Ratio, Truncated Distribution, Censored Distribution, Simulation

JEL classification: C01, C02, C12, C13, C14, C15, C32, G32, G38

ISSN: 14369753

Contact:

Fabian Mehmke
Bergweg 1
61440 Oberursel
F.Mehmke@googlemail.com

Prof. Dr. Heinz Cremers
Centre for Practical Quantitative Finance
Frankfurt School of Finance & Management
Sonnemannstraße 9-11
60314 Frankfurt am Main
H.Cremers@frankfurt-school.de

Prof. Dr. Natalie Packham
Centre for Practical Quantitative Finance
Frankfurt School of Finance & Management
Sonnemannstraße 9-11
60314 Frankfurt am Main
N.Packham@frankfurt-school.de

Inhaltsverzeichnis

1. Einleitung	4
2. Validierung von Marktrisikomodellen	6
2.1. Statistische Verfahren und Testtheorie	6
2.1.1. Maximum Likelihood-Parameterschätzung	6
2.1.2. Hypothesentests	8
2.1.3. Likelihood-Quotienten-Tests	11
2.1.4. α - und β -Fehler sowie statistische Macht	13
2.2. Prognose von Risiken	18
2.2.1. Risikomanagement & Marktrisikomodelle	18
2.2.2. Vorhersage von Wahrscheinlichkeitsverteilungen	18
2.2.3. Value at Risk	21
2.2.4. Überschreitungsfunktion	22
2.2.5. Expected Shortfall	23
2.2.6. Regulatorische Auswirkungen der Risikoprognose	23
2.2.7. Überprüfung der Modelle	24
2.3. Validierung basierend auf Überschreitungen	26
2.3.1. Time Until First Failure	26
2.3.2. Proportion of Failure	37
2.3.3. Baseler Ampelansatz / Traffic Light Approach	43
2.3.4. Magnitude of Loss Function	47
2.3.5. Markow-Test auf Unabhängigkeit	49
2.4. Validierung basierend auf der gesamten Verteilung	56
2.4.1. Rosenblatt-Transformation zur Gleichverteilung	56
2.4.2. CD-Verfahren: Kuiper-Statistik mit Worry Function	58
2.4.3. Kolmogorov-Smirnov-Test mit Worry Function	62
2.4.4. Wahrscheinlichkeitsintegral zur Transformation in eine Normalverteilung	65
2.4.5. Jarque-Bera-Test auf Normalverteilung	68
2.4.6. Regression auf Mittelwert, Varianz und Autokorrelation	70
2.4.7. Fazit zu ganzheitlichen Validierungsmodellen	74
2.5. Validierung basierend auf einem Teilbereich der Verteilung	76
2.5.1. Gestutzter Likelihood-Quotienten-Test	76

2.5.2.	Gauss-Test des Expected Shortfall bei gestutzter Verteilung	79
2.5.3.	Sattelpunktapproximation mittels Lugannani Rice-Formel zum Überprüfen des Expected Shortfall	82
3.	Simulation	85
3.1.	Korrekt spezifiziertes Risikomodell	86
3.2.	Risikomodell mit rollierenden Varianzen	90
3.3.	Risikomodell mit exponentiell geglätteten Varianzen	91
3.4.	GARCH-Risikomodell	92
3.5.	Konstante Risikounterschätzung	93
3.6.	Korrekter Value at Risk, unterschätzter Expected Shortfall	94
3.7.	Zusammenfassung der Simulationsergebnisse	96
4.	Fazit	98
	Literaturverzeichnis	100
	Nomenklatur	102
A.	Anhang	107
A.1.	Tabellen	107

1. Einleitung

Zahlreiche Finanzinstitute sind einem beträchtlichen finanziellen Risiko durch Veränderung von Marktpreisen ausgesetzt. Aus diesem Grund wurde es immer bedeutender, daraus resultierende potentielle Folgen möglichst präzise vorhersagen zu können. Zur konkreten Quantifizierung derartiger Risiken setzten sich in der Praxis vorrangig zwei Kennzahlen durch: Der Value at Risk sowie der Expected Shortfall. Ersterer soll dabei eine Aussage über eine Verlustgrenze treffen, die höchstens mit einer bestimmten, geringen Wahrscheinlichkeit - meist 1% oder 5% - überschritten wird. Der Expected Shortfall hingegen soll den erwarteten Verlust unter der Bedingung, dass der Value at Risk überschritten wird, angeben.

Beide Kennzahlen sind allerdings nur Kondensationen eines ganzheitlichen Risikomodells, welches nach unterschiedlichsten Annahmen und Kriterien strukturiert werden kann. Je nach Art und konkreter Anwendung verschiedener Modelle können dabei stark unterschiedliche Werte für Risikokenngrößen ermittelt werden.

Eine bedeutende Aufgabe ist es deshalb, Marktrisikomodelle auf ihre Güte zu überprüfen. Es muss festgestellt werden können, ob ein Modell verlässliche Aussagen über das reale Marktrisiko trifft. Dieser Aspekt wird im Folgenden den Schwerpunkt unserer Arbeit darstellen: Wir werden Verfahren beschreiben und entwickeln, die Marktrisikomodelle anhand ihrer Prognosequalität beurteilen können.

Im ersten Abschnitt des folgenden Kapitels werden wir dabei zuerst auf einige Grundlagen aus der Statistik eingehen, welche eine zentrale Rolle im Rest dieser Arbeit einnehmen werden. Anschließend werden wir detaillierter auf die Fragen eingehen, in welchem Zusammenhang Marktrisikomodelle zum allgemeinen Risikomanagement stehen, wie Marktrisikomanagement in der Praxis durchgeführt wird und welche Resultate es liefert.

In Abschnitt 2.3 werden wir schließlich erste Verfahren zur Validierung von Marktrisikomodellen betrachten. Dabei wird vor allem die Frage, ob der Value at Risk korrekt modelliert wird, im Vordergrund stehen. Mit der Magnitude of Loss Function kommen wir dabei im Abschnitt auch auf ein Verfahren zu sprechen, das in der Literatur bislang wenig Beachtung fand.

Anschließend werden wir Verfahren betrachten, die sich von dem reinen Fokus auf bestimmte Kennzahlen lösen und Marktrisikomodelle in ganzheitlichen Ansätzen

validieren. Vorweggenommen sei an dieser Stelle, dass damit jedoch automatisch auch die Prüfung auf Korrektheit des Value at Risk und des Expected Shortfall mit inbegriffen ist, wenngleich auch nicht mehr als einziges Kriterium. Im Rahmen des Kolmogorov-Smirnov-Tests mit Worry Function kommen wir dabei auf ein bislang noch nicht näher erforschtes Verfahren zu sprechen.

Danach werden wir auf Konzepte eingehen, welche ganz bestimmte Teilbereiche eines Risikomodells überprüfen. Damit wird es vor allem möglich sein, Spektralmaße wie den Expected Shortfall zu testen. Mit dem gestutzten Likelihood-Quotienten-Test modifizieren wir dabei ein bekanntes Verfahren. Mit dem Gauss-Test des Expected Shortfall bei gestutzter Verteilung führen wir ein gänzlich neues Testverfahren ein.

Zuletzt werden wir die bis dahin entwickelten theoretischen Verfahren in einer umfangreichen Simulation auf statistische Eigenschaften überprüfen, um die Qualität der einzelnen Testverfahren beurteilen zu können. Ein anschließendes Fazit wird die im Laufe dieser Arbeit gewonnenen Erkenntnisse abschließend zusammenfassen.

2. Validierung von Marktrisikomodellen

2.1. Statistische Verfahren und Testtheorie

Im Folgenden werden wir auf einige statistische Verfahren eingehen, die eine wichtige Basis für den weiteren Verlauf dieser Arbeit darstellen und unverzichtbar für jedes der später dargelegten Validierungskonzepte sein werden:

1. Die Schätzung unbekannter Verteilungsparameter
2. Hypothesentests zum Überprüfen bestimmter Aussagen, etwa Aussagen über die geschätzten Parameter
3. Likelihood-Quotienten-Tests als spezielle Form eines Hypothesentests
4. Fehler, die im Rahmen von Hypothesentests begangen werden können

2.1.1. Maximum Likelihood-Parameterschätzung

Die Maximum Likelihood-Parameterschätzung dient, wie auch schon von Ruppert (2010) erläutert, zur Schätzung unbekannter Verteilungsparameter. Die Frage im Rahmen des Maximum Likelihood-Verfahrens ist: Wie sollten unbekannte Parameter gewählt werden um die Wahrscheinlichkeit zu maximieren, eine konkrete Stichprobe aus der spezifizierten Verteilungsform zu realisieren.

Ausgangspunkt ist eine Zufallsvariable mit Dichtefunktion, welche von einem Parametervektor θ mit in der Regel einem oder mehreren unbekanntem Komponenten abhängt.

Beispielsweise könnte eine Zufallsvariable X vorliegen, von welcher wir wissen, dass sie normalverteilt mit bekannter Varianz 1, jedoch unbekanntem Parameter ϑ für den Erwartungswert ist. Dann ist $\theta = (\vartheta, 1)$, $X \sim \mathcal{N}(\vartheta, 1)$. Wir wollen nun basierend auf einer Stichprobe ermitteln, was der wahrscheinlichste Wert für ϑ ist.

Technisch gesehen liefert die Dichtefunktion einer kontinuierlich verteilten Zufallsvariable an einer einzigen Stelle keine sinnvoll interpretierbare Aussage über Wahrscheinlichkeiten, da im Kontinuierlichen die Wahrscheinlichkeit für den Erhalt irgendeines bestimmten Wertes stets exakt 0 ist. Eine konkrete Wahrscheinlichkeit erhält man lediglich durch Integration der Dichtefunktion auf einem bestimmten Intervall.

Die hier betrachteten stetigen Dichtefunktionen stehen dennoch in enger Verbindung zu den Wahrscheinlichkeiten. Nicht ganz präzise könnte man interpretieren: Je höher das Ergebnis der Dichtefunktion an einer bestimmten Stelle x ist, desto höher ist die Wahrscheinlichkeit, eine Realisation der Zufallsvariable zu erhalten, die *ungefähr* x entspricht. Wir werden zur Lesbarkeit im Folgenden auf diese präzise Unterscheidung verzichten, sprechen implizit jedoch stets vom hier erwähnten Zusammenhang.

Um auf unser Beispiel zurückzukommen: Wir wissen, dass die Dichtefunktion einer Normalverteilung ihr globales Maximum an der Stelle ihres Erwartungswerts hat. Eine Standardnormalverteilung hat also ihr Maximum an der Stelle 0, eine $\mathbb{N}(1, 1)$ -Verteilung an der Stelle 1 usw.

Nehmen wir an, wir beobachten nur eine Realisation $x_1 = 2$ unserer Zufallsvariablen X . Unser Ziel ist es nach wie vor, den Parameter ϑ zu finden, für den der Erhalt unserer Stichprobe - hier also nur des Werts $x_1 = 2$ - am wahrscheinlichsten ist. Dies können wir ganz intuitiv lösen: Wir wissen, dass die Dichtefunktion einer $\mathbb{N}(2, 1)$ -Verteilung ihr globales Maximum an der Stelle 2 hat. Bei dieser speziellen Verteilung wäre es also am wahrscheinlichsten, den Wert 2 zu erhalten. Wir wissen auch, dass jede andere Dichtefunktion $\mathbb{N}(\vartheta, 1)$, $\vartheta \in \mathbb{R} \setminus \{2\}$ ihr Maximum an einer anderen Stelle hat und somit die Wahrscheinlichkeit, aus irgendeiner anderen der angegebenen Normalverteilungen den Wert 2 zu erhalten, kleiner als bei der $\mathbb{N}(2, 1)$ -Verteilung ist.

Dann haben wir unsere Fragestellung bereits gelöst: Die Verteilung, die mit höchster Wahrscheinlichkeit für die Generierung unserer Stichprobe verantwortlich ist, ist die $\mathbb{N}(2, 1)$ -Verteilung.

Die Maximum Likelihood-Parameterschätzung erfolgt immer nach dem hier vorgestellten Prinzip. Wir suchen stets die Verteilung, die die größte Wahrscheinlichkeit für die Generierung einer konkreten Stichprobe liefert. Natürlich muss dieses Verfahren für realitätsnähere Einsatzzwecke formalisiert werden.

Wir definieren die allgemeine Dichtefunktion mit $f(x_1, \dots, x_N; \theta)$, wobei x_1, \dots, x_N die Realisationen einer Folge unabhängiger Wiederholungen der Zufallsvariable darstellen und θ wie gehabt der Parametervektor ist.

Wir definieren des Weiteren die Likelihoodfunktion L , welche die Wahrscheinlichkeit für den Erhalt der gesamten Stichprobe in Abhängigkeit vom Parametervektor θ angibt:

$$L(\theta) = f(x_1, \dots, x_N; \theta) = \prod_{n=1}^N f_{X_n}(x_n; \theta)$$

Wir wollen also das θ finden, für das die Likelihood-Funktion am größten wird. Dies erreichen wir, indem wir die Funktion ableiten, gleich Null setzen und entsprechend nach den unbekanntem Parametern auflösen. Das Ableiten stellt sich für große Produkte jedoch als relativ umständlich heraus. Aus diesem Grund ist es meist sinnvoll, die Likelihood-Funktion vorher zu logarithmieren. Da der \ln eine monotone Funktion ist, verändert er nicht die Lage der Maximalstellen.

Die so erhaltenen Maximum Likelihood-Schätzer haben hervorragende Eigenschaften. Sie sind konsistent, d.h. konvergieren gegen den unbekanntem Parameter und sind so asymptotisch erwartungstreu. Zudem sind sie effizient, d.h. es gibt keinen anderen Schätzer mit vergleichbaren Eigenschaften, der eine asymptotisch kleinere Varianz aufweist. Auch ist es möglich, Hypothesentests mit ML-Schätzern zu konstruieren, da ihre Verteilung gegen eine Normalverteilung konvergiert.

2.1.2. Hypothesentests

Die Grundlage eines jeden Hypothesentests sind zwei konkurrierende Hypothesen. Man unterscheidet bei Hypothesentests zwischen parametrischen und nicht-parametrischen Testverfahren.

Parametrische Tests sind solche, bei denen Aussagen über einen unbekanntem Parameter einer gegebenen Verteilung getestet werden können. Wir führten im vorherigen Abschnitt eine normalverteilten Zufallsvariable mit unbekanntem Erwartungswert ein, $X \sim \mathbb{N}(\vartheta, 1)$. Nehmen wir an, wir treffen die Vermutung, dass der Erwartungswert eigentlich Null entsprechen müsste - wir können uns aber nicht sicher sein, ob dies auch wirklich stimmt. Also wollen wir zwei konkurrierende Hypothesen testen:

$$H_0 : \theta = (0, 1) \text{ gegen } H_1 : \theta = (\vartheta, 1), \vartheta \neq 0$$

Wir testen also ganz konkrete numerische Werte für die Parameter. Auch wenn in diesem Beispiel nur ein Parameter, der Erwartungswert, variabel ist wurden im Beispiel die Hypothesen explizit in Abhängigkeit des gesamten Parametervektors

inklusive der hier konstanten Varianz gewählt. Dies soll unterstreichen, dass parametrische Tests keineswegs auf nur einen einzigen Parameter beschränkt sind. Man kann simultan beliebige Parameterkombinationen testen.

Die Hypothesen können dabei weitestgehend beliebig aufgebaut sein. Wir könnten in der Nullhypothese auch über verschiedene Parameterintervalle testen. Zudem muss die Alternative nicht auf Ungleichheit testen, sondern kann auch prüfen, ob die Parameter alternativ einen ganz konkreten Wert haben, kleiner als ein bestimmter Wert oder größer als ein bestimmter Wert sind. Wichtig ist jedoch, dass darauf geachtet werden sollte, dass stets eine der beiden Hypothesen wahr ist. Allgemeiner können wir also eine Hypothese wie folgt formulieren:

$$H_0 : \theta \in \Theta_0 \text{ gegen } H_1 : \theta \in \Theta_1 \\ \Theta_0 \cup \Theta_1 = \Theta$$

Nicht-parametrische Tests sind solche, bei denen wir nicht direkt numerische Parameter testen können. Wir haben im vorherigen Beispiel das Wissen vorausgesetzt, dass die Zufallsvariable X normalverteilt ist und nur den Mittelwert der Normalverteilung als unbekannt angenommen. Doch in der Realität können wir uns in den seltensten Fällen wirklich sicher sein, dass die Daten auch wirklich einer bekannten Verteilung entstammen. Könnte es im obigen Beispiel nicht etwa genauso gut sein, dass X t-verteilt ist? In dem Fall würde der oben präsentierte Hypothesentest trügerisch sein, da unter Umständen die Daten einer t-Verteilung den gleichen Mittelwert und die gleiche Varianz wie normalverteilte Daten aufweisen könnten.

Speziell die Verteilungsform ist aber kein einfach quantifizierbarer Zusammenhang. Wir können strikt gesehen niemals wirklich testen, ob X tatsächlich normalverteilt ist oder nicht. Wir können uns allerdings sehr wohl quantifizierbare Zusammenhänge zu Nutze machen. Wie wir später sehen werden, könnten wir beispielsweise testen, ob unsere Schätzung für Schiefe und Wölbung mit einer Normalverteilung vereinbar sind - dieser quantifizierte Zusammenhang wäre nun wieder überprüfbar. Doch genau betrachtet testen wir nun nicht mehr unsere wirkliche Hypothesen, sondern völlig neue Hypothesen, von denen wir vermuten, dass sie uns zumindest Indizien für die Beantwortung unserer ursprünglichen Problemstellung liefern.

Der eigentliche Ablauf eines Hypothesentests lässt sich wie folgt darstellen:

1. Testproblem

Im Verlauf dieser Arbeit werden wir stets Hypothesentests der folgenden Form verwenden:

$$H_0 : \theta \in \Theta_0 \text{ gegen } H_1 : \theta \in \Theta_1$$

dabei gilt für unsere Zwecke stets $\Theta_0 \subset \mathbb{R}^n$, $\Theta_1 = \mathbb{R}^n \setminus \Theta_0$, wobei n hier die Dimension des Parametervektors θ bezeichnet. Θ sei im Folgenden der gesamte Parameterraum, $\Theta = \Theta_0 \cup \Theta_1 = \mathbb{R}^n$.

Wir testen in dieser Arbeit stets zweiseitig, wobei die Nullhypothese eine einfache und die Alternativhypothese eine zusammengesetzte Hypothese ist.

2. Testfunktion

Zum Überprüfen der im ersten Schritt aufgestellten Hypothesen benötigen wir nun eine Testfunktion, die uns anhand der für das Testproblem relevanten Daten einen überprüfaren Wert liefert. In der Praxis bedeutet dies meist, dass man anhand aller Realisationen der relevanten Zufallsvariable(n) einen einzigen Wert berechnet, welcher einen Bezug zu unserem Testproblem hat. Bei einer suffizienten Testfunktion sollte die Komprimierung der komplexen Stichprobe auf eine Zahl dabei möglichst keinen Informationsverlust haben.

Für unsere Zwecke genügt folgende Definition der Testfunktion T :

$$T : \mathbb{R}^n \rightarrow \mathbb{R}$$

3. Ablehnungsbereich

Der Wert der Testfunktion alleine hilft uns bislang noch nicht weiter, denn wir wissen noch nicht, für welche Werte die Nullhypothese abgelehnt werden muss und für welche Werte sie nicht verworfen werden kann.

Dazu definieren wir nun mit $\mathbf{X} = (X_1, \dots, X_N)^T$ einen Ablehnungsbereich A ,

$$A \subset \mathbb{R}$$
$$\sup_{\theta \in \Theta_0} \mathcal{P}(T(\mathbf{X}) \in A | \theta) = \alpha$$

Die Wahrscheinlichkeit, dass die Testfunktion für eine richtige Hypothese in den Ablehnungsbereich fällt darf also nur α betragen. In der Regel kann α frei gewählt werden.

4. Entscheidungsregel

Basierend auf den ersten drei Schritten kann man nun eine Regel aufstellen, wann die Nullhypothese verworfen werden muss, und wann sie nicht abgelehnt werden kann.

Wir werden stets ablehnen, wenn die Testfunktion in den Ablehnungsbereich fällt, ansonsten werden wir die Nullhypothese nicht ablehnen:

$$\begin{aligned} &\text{Ablehnung für } T(\mathbf{x}) \in A \\ &\text{Nicht-Ablehnung für } T(\mathbf{x}) \notin A \end{aligned}$$

Anzumerken ist, dass eine Nicht-Ablehnung der Nullhypothese keiner expliziten Annahme derselbigen darstellt. Der Grund hierfür liegt in der Asymmetrie der Signifikanztests, auf welche wir im Abschnitt 2.1.4 genauer eingehen werden.

2.1.3. Likelihood-Quotienten-Tests

Ein Verfahren, welches oft zur Bestimmung der Testfunktion zum Einsatz kommt, ist das sogenannte allgemeine Likelihood-Quotienten Testverfahren. Die Testfunktion ist hierbei ein Quotient aus dem Maximum der Likelihood-Funktion unter Einschränkung der Parameter auf die Bereiche der Nullhypothese zu dem maximalen Wert der uneingeschränkten Likelihood-Funktion,

$$\lambda(\mathbf{x}) = \frac{\sup_{\theta \in \Theta_0} L(\mathbf{x}; \theta)}{\sup_{\theta \in \Theta} L(\mathbf{x}; \theta)}$$

$\lambda(\mathbf{x})$ kann hierbei offensichtlich einen Wert zwischen 0 und 1 inklusive annehmen. Je näher der Wert bei 1 liegt, desto eher kann man davon ausgehen, dass die Daten in \mathbf{x} von einer Verteilung erzeugt wurden, deren Parameter mit der Nullhypothese in Einklang gebracht werden können.

Aus diesem allgemeinen Quotienten lassen sich nahezu alle gängigen Testverfahren herleiten. So kann beispielsweise auch der sicherlich bekannteste und verbreitetste Test, der Gauss-Test, über einen allgemeinen Likelihood-Quotienten bestimmt werden.

In vielen Fällen kann jedoch nicht ohne weiteres eine einfache Verteilung für den Likelihood-Quotienten gefunden werden. Wir werden daher insbesondere an einer Transformation der Testfunktion interessiert sein, deren Verteilung unter bestimmten Annahmen zumindest asymptotisch bekannt ist:

$$\Lambda(\mathbf{x}) = -2 \ln \left(\frac{\sup_{\theta \in \Theta_0} L(\mathbf{x}; \theta)}{\sup_{\theta \in \Theta} L(\mathbf{x}; \theta)} \right)$$

Um die Verteilung für $\Lambda(\mathbf{x})$ zu bestimmen werden wir uns eines Theorems bedienen, welches von Suhov & Kelbert (2005) beschrieben wird:

Wilks Theorem Seien die Realisationen der Stichprobe $\mathbf{x} = (x_1, \dots, x_N)^T$ i.i.d mit Dichtefunktion $f(\cdot; \theta)$, $\theta \in \Theta$. Sei Θ eine offene Menge im euklidischen Raum der Dimension $|\Theta|$ und der Maximum-Likelihood-Schätzer $\hat{\theta}$ des Parametervektors eine asymptotisch normalverteilte Zufallsvariable für $N \rightarrow \infty$. Sei $\Theta_0 \subset \Theta$, sodass Θ_0 ebenfalls eine offene Menge im euklidischen Raum der Dimension $|\Theta_0| < |\Theta|$ ist. Dann gilt unter H_0 , dass

$$\forall \theta \in \Theta_0, h > 0 : \lim_{N \rightarrow \infty} \mathcal{P}_\theta(\Lambda(\mathbf{x}) > h) = \int_h^\infty f_{\chi_d^2}(x) dx$$

mit $d = |\Theta| - |\Theta_0|$

Die Teststatistik Λ ist dann also χ^2 -verteilt mit Freiheitsgrad d . Gemäß Suhov & Kelbert (2005) existiert eine weitere Variante des Theorems, welche für unabhängige, aber nicht notwendigerweise identisch verteilte Zufallsvariablen gilt. Für den recht technischen Beweis des Theorems verweisen wir auf van der Vaart (2000), welcher auch die wesentlichen Restriktionen nochmals zusammenfasst:

1. Die Dichtefunktion muss differenzierbar in θ sein, da der Maximum Likelihood-Schätzer ermittelt werden muss.
2. Θ_0 und Θ müssen lineare Räume sein - dies entspricht einer anderen Form der Bedingung der Offenheit.

Für unsere Zwecke stellen diese beiden Bedingungen keine Einschränkungen dar, da Θ und Θ_0 in unseren Hypothesentests stets lineare Mengen sein werden und, wie sich herausstellen wird, jede Dichtefunktion, welche wir betrachten werden, differenzierbar in θ sein wird.

Als größerer Kritikpunkt ist jedoch vorab zu nennen, dass Wilks Theorem nur asymptotisch gilt. Wir testen oft jedoch nur in vergleichsweise kleinen Stichproben.

Dies ist leider ein übliches Problem in der Statistik, da die meisten Verfahren nur asymptotisch Gültigkeit aufweisen. Wir sollten uns jedoch durch diese Kritik umso mehr bewusst sein, dass Stichproben in der Regel möglichst groß gewählt werden sollten. Zudem müssen wir explizit auf Unabhängigkeit in den Daten achten - dafür werden wir jedoch separate Testverfahren vorstellen.

2.1.4. α - und β -Fehler sowie statistische Macht

Ein Hypothesentest führt immer zu einer von zwei möglichen Entscheidungen: Entweder, wir verwerfen die Nullhypothese, oder wir verwerfen sie nicht. Auch wenn wir es meist nicht explizit wissen können, aber je nachdem, ob die Hypothese in der Realität zutrifft oder nicht, entscheiden wir uns dabei entweder richtig, oder wir begehen einen Fehler.

Dieses Entscheidungsmuster lässt sich gut anhand einer Tabelle zeigen:

	H_0 korrekt	H_1 korrekt
Nicht-Ablehnung von H_0	Korrekte Entscheidung mit Wahrscheinlichkeit $1 - \alpha$	Fehler 2. Art mit Wahrscheinlichkeit β
Ablehnung von H_0	Fehler 1. Art mit Wahrscheinlichkeit α	Korrekte Entscheidung mit Wahrscheinlichkeit $1 - \beta$ ("statistische Macht")

Die Spalten sind dabei getrennt voneinander zu betrachten - es ist zwar immer wahr, dass H_0 oder H_1 korrekt ist, wir werden in der Praxis aber niemals wissen, welche der beiden Möglichkeiten tatsächlich zutrifft. Die Felder geben entsprechende Resultate und Wahrscheinlichkeiten an unter der unbekanntem Tatsache des Zutreffens einer dieser beiden Möglichkeiten.

Fehler 1. Art

Wenn die Nullhypothese in Realität korrekt ist und wir sie basierend auf unserem Hypothesentest tatsächlich nicht ablehnen, dann haben wir die korrekte Entscheidung getroffen. Unter der Bedingung, dass H_0 korrekt ist, passiert dies mit einer Wahrscheinlichkeit von $1 - \alpha$. Sollten wir die Hypothese basierend auf unserem Test verwerfen, so treffen wir die falsche Entscheidung mit einer Wahrscheinlichkeit von α unter der Bedingung dass H_0 korrekt ist.

Bei einem guten Hypothesentest können wir α präzise selbst festlegen. Dies geht schon aus der Definition des Ablehnungsbereichs in Abschnitt 2.1.2 hervor. In der Praxis dieser Arbeit werden wir teils das Problem haben, dass bestimmte Testverfahren nur asymptotisch gelten. Insbesondere gilt diese Einschränkung bei Likelihood-Quotienten-Tests. In vergleichsweise kleinen Stichproben treten hier Ungenauigkeiten auf.

Eine dieser möglichen Ungenauigkeiten ist, dass ein Hypothesentest ein vorgegebenes α -Niveau über- oder unterschreitet. Wenn also in der Theorie korrekte Hypothesen nur in α Prozent aller Fälle abgelehnt werden sollten, so kann es in der Praxis durchaus passieren, dass korrekte Hypothesen wesentlich häufiger oder wesentlich seltener abgelehnt werden.

Dies ist tatsächlich ein erhebliches Problem. Eines der Grundprinzipien von Hypothesentests ist, dass man Nullhypothesen gegebenenfalls signifikant ablehnen kann. Dies ist genau deshalb der Fall, weil man sich in der Regel präzise seiner Fehlerwahrscheinlichkeit bei dieser spezifischen Entscheidung bewusst ist. Ist die Nullhypothese nämlich wahr, obwohl man sie ablehnt, befindet man sich in der Tabelle im Bereich " H_0 korrekt, Ablehnung von H_0 ". Die Wahrscheinlichkeit dafür ist aber genau α . Wenn nun aber das Testverfahren nur asymptotisch gilt und der tatsächliche α -Wert von seinem theoretisch festgelegten Wert abweicht, so sind wir uns letztlich nicht mehr unserer Fehlerwahrscheinlichkeit bewusst.

Leider lässt sich dieses Problem nicht ohne Weiteres beheben, denn auch in kleinen Stichproben ideale Testverfahren lassen sich meist nicht finden. Wir werden bei diversen Simulationen genauer die Auswirkungen des Phänomens betrachten. In der Praxis sollte dieser Aspekt eine Rolle bei der Wahl eines geeigneten Validierungskonzeptes spielen.

Fehler 2. Art und statistische Macht

Angenommen wir befinden uns nun in der Situation, dass H_1 in der Realität korrekt ist. Da wir dies nicht wissen, kann unser Hypothesentest nach wie vor zu einer Nicht-Ablehnung von H_0 oder zu einer Ablehnung von H_0 , also einer Annahme von H_1 , führen.

Auch hierbei können zwei Situationen entstehen: Lehnen wir H_0 nicht ab, so entscheiden wir uns falsch. Dies passiert mit der bedingten Wahrscheinlichkeit β . Lehnen wir H_0 hingegen ab, so entscheiden wir uns mit einer bedingten Wahrscheinlichkeit in Höhe von $1 - \beta$ richtig.

Ein zentraler Begriff in der Statistik ist dabei die “statistische Macht”. Dies ist nichts anderes als die Wahrscheinlichkeit $1 - \beta$. Die Macht sagt also letztlich aus, wie oft wir uns - gegeben die unbekannt reale Tatsache H_1 ist korrekt - auch tatsächlich für eine Ablehnung von H_0 entscheiden. Bei Parametertests lässt sich dies wie folgt ausdrücken:

$$\mathcal{P}(T \in A | \theta \in \Theta_1) = 1 - \beta$$

Perfekt sind dementsprechend natürlich Testverfahren, die eine hohe Macht aufweisen, denn sie minimieren den Fehler 2. Art.

Oft wird zur Analyse der Macht die sogenannte “Gütefunktion”, im Folgenden mit $G(\cdot)$ bezeichnet, herangezogen. Sie gibt in Abhängigkeit des wahren, aber unbekannt Parametervektors θ die Wahrscheinlichkeit an, dass eine Stichprobe X_1, \dots, X_N in den Ablehnungsbereich des Hypothesentests fällt:

$$G(\theta) = \mathcal{P}_\theta((X_1, \dots, X_N) \in A) | \theta \in \Theta$$

Wenn $\theta \in \Theta_0$, so gibt die Gütefunktion für $G(\theta)$ die Wahrscheinlichkeit für eine Fehlentscheidung an, denn in diesem Fall ist die Nullhypothese korrekt und $G(\theta)$ ist die Wahrscheinlichkeit, dass die Stichprobe dennoch in den Ablehnungsbereich fällt.

Für $\theta \in \Theta_1$ gibt die Gütefunktion Wahrscheinlichkeiten für korrekte Entscheidungen an. In diesem Fall ist nämlich die Nullhypothese falsch und $G(\theta)$ entspricht der Wahrscheinlichkeit, die Nullhypothese korrekterweise abzulehnen.

Da θ in Realität unbekannt ist sind die Aussagen der Gütefunktion eher theoretischer Natur. Sie liefert Erkenntnisse darüber, wie hoch die Ablehnungsraten der Nullhypothese unter Eintreffen einer bestimmten Parameterkombination θ wären.

Führen wir das Beispiel der vorherigen Kapitel weiter: Sei $X \sim \mathbb{N}(\vartheta, 1)$ und die zu überprüfende Hypothese

$$H_0 : \theta = (0, 1) \text{ gegen } H_1 : \theta = (\vartheta, 1), \vartheta \neq 0$$

Wenn wir zur Überprüfung der Hypothese einen Gauss-Test mit $\alpha = 5\%$ basierend auf einem Stichprobenumfang $N = 9$ durchführen, so stellt sich die Gütefunktion wie folgt dar:

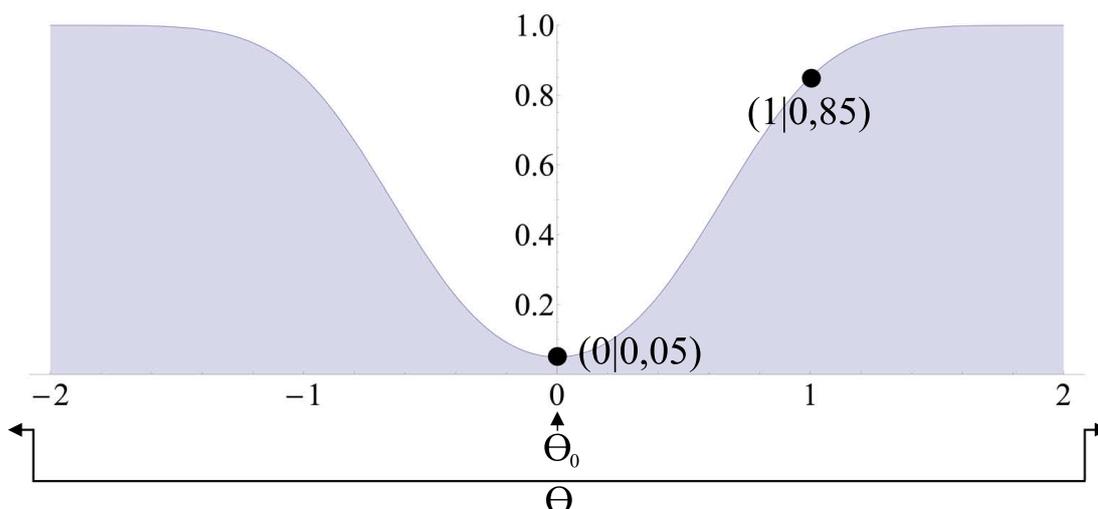


Abbildung 2.1.: Exemplarische Gütefunktion eines Gauss-Tests

Ersichtlich aus dem Schaubild ist beispielsweise, dass eine Stichprobe aus 9 Werten in rund 85% aller Fälle in den Ablehnungsbereich fallen wird, wenn der wahre Parameter $\vartheta = 1$ und somit $\theta = (1, 1)$ ist. Da die Nullhypothese in diesem Fall falsch ist, entspricht die Ablehnung derselbigen einer korrekten Entscheidung. Wir werden im Sprachgebrauch dieser Arbeit sagen, dass die statistische Macht dieser Konstellation 85% beträgt.

Für $\vartheta = 0$ trifft die Nullhypothese zu. Der Wert der Gütefunktion an dieser Stelle ist genau 5% und entspricht der Wahrscheinlichkeit, die Nullhypothese fälschlicherweise abzulehnen. Der Wert von 5% ist genau der für den Hypothesentest spezifizierte α -Fehler.

Für diesen zweiseitigen Gauss-Test können wir die Gütefunktion analytisch mit der Vorschrift

$$G(\vartheta) = \Phi\left(-\Phi^{-1}\left(1 - \frac{\alpha}{2}\right) - \vartheta\sqrt{N}\right) + \Phi\left(-\Phi^{-1}\left(1 - \frac{\alpha}{2}\right) + \vartheta\sqrt{N}\right)$$

angeben, wobei $\Phi(\cdot)$ die Verteilungsfunktion der Standardnormalverteilung und $\Phi^{-1}(\cdot)$ die inverse Verteilungsfunktion der Standardnormalverteilung ist.

Bei nahezu allen Testverfahren, welche im Verlauf dieser Arbeit vorgestellt werden, wird es jedoch nicht möglich sein, eine konkrete Formel für die Gütefunktion zu finden. Um einen konkreten Wert zu berechnen müssten wir nämlich exakt wissen, was passiert, wenn H_0 nicht eintritt. Im obigen Beispiel war dies relativ einfach: H_1 besagt, dass die Werte normalverteilt mit bekannter Varianz und lediglich unbekanntem Erwartungswert sind.

In den meisten Fällen dieser Arbeit umfasst H_1 aber stets eine ganze Bandbreite sehr unterschiedlicher Alternativen. Wenn wir den obigen Hypothesentest etwa umformulieren zu

$$H_0 : X \sim \mathbb{N}(0, 1) \text{ gegen } X \not\sim \mathbb{N}(0, 1)$$

dann umfasst die Alternativhypothese jegliche vorstellbare Verteilungsform angefangen von Gleichverteilungen, t-Verteilungen, Binomialverteilungen bis hin zu gänzlich parameterlosen Verteilungsformen. Wenn wir uns gegen H_0 entscheiden wissen wir also nicht, was stattdessen zutrifft, sondern wir wissen bloß, dass H_0 vermutlich nicht zutrifft.

Dennoch werden wir an einigen Stellen dieser Arbeit die statistische Macht der Testverfahren unter bestimmten Bedingungen simulieren. Dies werden stets hypothetische “was wäre wenn”-Szenarien sein. Wir können beispielsweise simulieren, wie hoch die statistische Macht wäre, wenn die Daten alternativ einer $\mathbb{U}(0, 1)$ -Verteilung oder einer $\mathbb{B}(250, 5\%)$ -Verteilung entstammen würden. Wir führen also gewisse Annahmen und Restriktionen für die Alternativhypothese ein und können mit diesen Annahmen Werte einer Gütefunktion simulieren.

Wenn man sich in der Praxis letztlich für ein konkretes Validierungskonzept entscheiden will, so kann man demzufolge nicht angeben, welches davon *die* höchste statistische Macht hat, denn einen konkreten Wert dafür gibt es nicht. Man kann allerdings sehr wohl betrachten, welche Testverfahren unter diversen hypothetischen Alternativbedingungen jeweils die größte Macht besitzen würden, und kann dies in die Entscheidung mit einfließen lassen.

2.2. Prognose von Risiken

2.2.1. Risikomanagement & Marktrisikomodelle

Der Begriff "Risikomanagement" ist recht allgemein gehalten und umfasst besonders im Finanzsektor eine Vielzahl verschiedener Bereiche.

Die eigentlichen Tätigkeiten, die im Rahmen eines modernen Risikomanagements durchzuführen sind, untergliedern sich im Wesentlichen in Identifikation, Analyse (Quantifizierung), Überwachung und Steuerung von Risiken.

Die Risiken als solche lassen sich wiederum in zahlreiche Unterkategorien aufteilen. Die von Basel II vorgeschriebenen Mindestkapitalanforderungen etwa legen einen Fokus auf Marktpreis-, Kreditausfall- und operationelle Risiken. In der Praxis sind für Unternehmen jedoch noch deutlich mehr Aspekte von Bedeutung; exemplarisch zu nennen sind hier etwa Liquiditätsrisiken, wie nicht zuletzt die jüngste Bankenkrise zeigte, und die momentan stark an unternehmerischem Interesse gewinnenden Reputationsrisiken.

Relevant für die in dieser Arbeit vorgestellten Themen sind in erster Linie Modelle, die der Überwachung von Marktpreisrisiken dienen. Eine für unsere Zwecke angemessene Definition dafür wird von Crnkovic & Drachman (1996) eingeführt. Hier wird das Marktrisikomanagement als Prozess, die Wahrscheinlichkeitsverteilungen der relevanten Risikofaktoren korrekt abzubilden, betrachtet.

Die Grundannahme hierfür ist letztlich, dass wir den Wert unserer Produkte perfekt durch eine Kombination diverser quantifizierbarer Faktoren berechnen können - etwa durch Zinssätze, lineare oder auch nicht lineare Zusammenhänge zu anderen Produkten, Haltekosten usw. Sofern dies zutrifft können wir eine präzise Wertbestimmung zu jedem zukünftigen Zeitpunkt genau dann durchführen, wenn wir die Wahrscheinlichkeiten kennen, mit denen die relevanten Risikofaktoren ihre jeweiligen Wertzustände annehmen können.

Der Schlüssel für ein gutes Management von Marktrisiken ist also eine realitätsnahe Modellierung zukünftiger Risikofaktoren durch Schätzung von geeigneten Wahrscheinlichkeitsverteilungen.

2.2.2. Vorhersage von Wahrscheinlichkeitsverteilungen

Die Auswahl von Wahrscheinlichkeitsverteilungen von Risikofaktoren kann auf vielfältige Weise erfolgen.

In der Praxis werden meist relative oder absolute zukünftige Wertänderungen modelliert. Eine detailliertere Aufteilung in elementarere Risikofaktoren¹ ist zwar möglich, wird jedoch auf Portfolioebene in der Regel nicht vollzogen.

Grundsätzlich kann man die statistischen Probleme in zwei Kategorien eingliedern: Parametrische und nicht-parametrische Verfahren.

Parametrische Verfahren sind solche, bei denen anfangs eine konkrete Klasse von Dichtefunktionen angenommen wird und anschließend die dazugehörigen Parameter geschätzt werden. Die weitverbreitetste Annahme im Finanzbereich ist dabei sicherlich, dass logarithmierte Renditen r approximativ normalverteilt sind. Die Parameter der Verteilung werden dabei oft nach einer der folgenden Methoden bestimmt (Sanddorf-Köhle 2011):

- Gleitende Varianz

$$\hat{\mu}_t = \frac{1}{p} \sum_{k=1}^p r_{t-k} \text{ für } t \in \mathbb{N}, T > p$$
$$\hat{\sigma}_t^2 = \frac{1}{p} \sum_{k=1}^p (r_{t-k} - \hat{\mu}_t)^2$$

- Exponentiell geglättete Varianz

$$\hat{\mu} = \frac{1}{T} \sum_{t=1}^T r_t$$
$$\hat{\sigma}_t^2 = \lambda \hat{\sigma}_{t-1}^2 + (1 - \lambda) (r_{t-1} - \hat{\mu})^2$$

und ein sinnvoll gewählter Startwert $\hat{\sigma}_1^2$, z.B. $\hat{\sigma}_1^2 = (r_1 - \hat{\mu})^2$.

¹beispielsweise Zinssätze, Nutzenfunktionen, Halteprämien usw.

- Ein GARCH-Modell

$$\hat{\mu} = \frac{1}{T} \sum_{t=1}^T r_t$$

$$\sigma_t^2 = \omega + \sum_{i=1}^p \alpha_i \epsilon_{t-1}^2 + \sum_{j=1}^q \beta_j \sigma_{t-1}^2$$

$$\epsilon_t = v_t \sigma_t, v_t \stackrel{i.i.d.}{\sim} \text{mit } E(v_t) = 0, \text{Var}(v_t) = 1$$

und $\omega > 0$,

$$\forall i \in \{1, \dots, p\} : \alpha_i \geq 0,$$

$$\forall j \in \{1, \dots, q\} : \beta_j \geq 0$$

- Ein GJR-GARCH-Modell

$$\hat{\mu} = \frac{1}{T} \sum_{t=1}^T r_t$$

$$\sigma_t^2 = \omega + \sum_{i=1}^p \alpha_i (|\epsilon_{t-1}| - \gamma_i \epsilon_{t-1})^2 + \sum_{j=1}^q \beta_j \sigma_{t-1}^2$$

$$\epsilon_t = v_t \sigma_t, v_t \stackrel{i.i.d.}{\sim} \text{mit } E(v_t) = 0, \text{Var}(v_t) = 1$$

und $\omega > 0$,

$$\forall i \in \{1, \dots, p\} : \alpha_i \geq 0,$$

$$\forall j \in \{1, \dots, q\} : \beta_j \geq 0,$$

$$\gamma_1 \in (-1, 1),$$

$$\forall i \in \{2, \dots, p\} : \gamma_i = 0$$

Nicht-parametrische Verfahren hingegen sind Methoden, die ohne a priori Festlegung auf eine bestimmte Verteilung auskommen. Dann wird meist die Stichprobe als solche als empirische Verteilungsfunktion gedeutet. Der Vorteil eines solchen Verfahrens ist eine recht hohe Flexibilität - in der Tat kennen wir in der Regel nie die unterliegenden Verteilungen, sodass parametrische Verfahren immer auf teils nicht sonderlich realitätsnahe Annahmen gestützt werden müssen. Nicht-parametrische Verfahren kommen hingegen ohne konkrete Annahmen über die Form der tatsächlichen Verteilung und auch ohne eine explizite Modellierung von Abhängigkeiten wie beispielsweise Korrelationswerten aus.

Wie bereits von Kupiec (1998) sowie Dowd (2004) angesprochen, gibt es jedoch auch zahlreiche Nachteile nicht-parametrischer Verfahren, die ihren Einsatz in der

Praxis oft fragwürdig erscheinen lassen. Um die empirische Verteilung möglichst realitätsnah spezifizieren zu können werden enorme Datenmengen gebraucht, insbesondere bei überdurchschnittlicher Leptokurtosis der tatsächlichen Verteilung², denn hierbei ist es wichtig, auch in der Stichprobe schon viele Extremwerte vorliegen zu haben, die selbst unter Leptokurtosis nur mit vergleichsweise geringer Wahrscheinlichkeit und dementsprechend selten auftreten. Genau dieses Phänomen ist jedoch gerade bei Finanzwerten oft anzutreffen. Zudem können nicht-parametrische Verfahren nur verlässliche Ergebnisse liefern, wenn die tatsächliche Verteilung stationär ist, sich also nicht im Zeitablauf verändert, denn es wird implizit davon ausgegangen, dass die Daten der Stichprobe allesamt der selben Verteilung entstammen. Auch dies ist im Finanzbereich in aller Regel höchst unrealistisch. Bei parametrischen Verfahren können sowohl Leptokurtosis als auch Nicht-Stationarität bewusst mit einbezogen werden, wenngleich es teils die entsprechende Modellschätzung erschwert.

Ein typisches Beispiel eines nicht-parametrischen Verfahrens zur Bestimmung möglicher Portfoliowertänderungen ist die sogenannte “historische Simulation”, bei der eine empirische Verteilungsfunktion auf Basis der Stichprobe $\Delta v_t, t \in \{1, \dots, T\}$ berechnet wird:

$$\hat{F}_{\Delta V}(x) = \begin{cases} 0 & \text{wenn } x < \Delta v_{[1]} \\ \frac{1}{T} \max\{t \mid \Delta v_{[t]} \leq x\} & \text{wenn } \Delta v_{[1]} \leq x < \Delta v_{[T]} \\ 1 & \text{wenn } x \geq \Delta v_{[T]} \end{cases}$$

2.2.3. Value at Risk

Nachdem, wie im vorherigen Abschnitt erläutert, die möglichen Wertänderungen eines Portfolios modelliert wurden, stellt sich nun die Frage, wie sich daraus eine Kennzahl für das Risiko quantifizieren lässt. Im Bankbereich wird hierfür gerne der sogenannte “Value at Risk” herangezogen (Kupiec 1998). Mit der formalen Definition folgen wir Cremers (2011):

$$VaR_t(p) = -\sup\{x \in \mathbb{R} \mid \mathcal{P}(\Delta V(t|\Omega_t) \leq x) \leq p\}$$

Er ist also der bezüglich der Verknüpfung + inverse Wert der größten Portfoliowertänderung ΔV , die mit einer Wahrscheinlichkeit von höchstens p unterschritten wird.

ΔV ist dabei abhängig von einem Betrachtungszeitpunkt t sowie der Portfoliozusammensetzung zu diesem Zeitpunkt, Ω_t .

²auch bekannt als “Fat Tails”

Diese recht formale Definition lässt sich jedoch für kontinuierliche Verteilungen anschaulicher interpretieren als negativer Wert der Inversen der Verteilungsfunktion der Wertänderungen:

$$VaR_t(p) = -F_{\Delta V(t|\Omega_t)}^{-1}(p)$$

Der Verlust im Portfolio sollte also nur mit einer Wahrscheinlichkeit von höchstens p niedriger als $-VaR$ sein. Typische Werte für p sind dabei 1% oder 5%.

Bei diskreten Verteilungen müssen Alternativdefinitionen des Value at Risk genutzt werden. Dabei kann es vorkommen, dass die Überschreitungswahrscheinlichkeit eines völlig korrekt berechneten Value at Risk tatsächlich höher als p ist. Sollte mit diskreten Verteilungen gearbeitet werden muss diese Tatsache speziell bei Testverfahren, welche Überschreitungswahrscheinlichkeiten konkret beinhalten, beachtet werden. Diese Konstellation ist jedoch weder im Verlauf dieser Arbeit noch bei gängigen Praxisverfahren von Bedeutung. Wir werden deshalb nicht weiter darauf eingehen.

2.2.4. Überschreitungsfunktion

Im Weiteren wird für uns eine Indikatorvariable relevant sein, die für jeden Zeitpunkt $t \in \tau, \tau = \{1, \dots, T\}$ zwei Zustände annehmen kann:

$$I_t = \begin{cases} 1 & \text{wenn } \Delta V_t < -VaR_t(p^*) \\ 0 & \text{sonst.} \end{cases}$$

ΔV_t ist dabei die Wertänderung eines Portfolios vom Zeitpunkt $t - 1$ auf t . Mit der mathematischen Definition orientieren wir uns an Campbell (2006/2007).³ Die Funktion I_t nimmt also für den Zeitpunkt t den Wert 1 an, wenn die Value at Risk-Prognose für diesen Zeitpunkt tatsächlich überschritten wird, ansonsten 0. Bei einem korrekt spezifizierten VaR -Modell sollte gelten:

$$\forall I_t, t \in \tau : \mathcal{P}(I_t = 1) = E(I_t) \leq p$$

Die Wahrscheinlichkeit, dass die Indikatorfunktion zu einem beliebigen Zeitpunkt t den Wert 1 angibt, sollte genau der Risikowahrscheinlichkeit des zugehörigen

³Wir weichen jedoch bewusst im Bezug auf die zeitabhängige Indizierung von Campbell (2006/2007) ab.

Value at Risk-Modells entsprechen - genau dann gilt im Fall einer kontinuierlichen Werteverteilung $E(I_t) = p$. I_t ist dann bernoulliverteilt:

$$\forall t \in \tau : I_t \sim \mathbb{B}\mathbb{E}(p)$$

$$f_{I_t}(x) = (1-p)^{p-x} p^x \text{ für } x \in \{0, 1\}$$

Wir definieren zur künftigen Vereinfachung I als die Summe aller Überschreitungen. In einem korrekten Modell sollte sie als Summe unabhängig und identisch bernoulliverteilter Zufallsvariablen binomialverteilt sein:

$$I = \sum_{t \in \tau} I_t$$

$$I \sim \mathbb{B}(T, p)$$

2.2.5. Expected Shortfall

Der Expected Shortfall ist eine weitere Risikokennzahl für das Portfolio. Er gibt an, wie hoch der durchschnittliche Portfolioverlust ist unter der Bedingung, dass der Value at Risk überschritten wird. Formal definiert wird er als

$$ES_t(p) = -\frac{1}{p} \left(E(\Delta V(t|\Omega_t) \cdot 1_{\{\Delta V(t|\Omega_t) \leq -VaR_t(p)\}}) \right. \\ \left. + VaR_t(p) \cdot (\mathcal{P}(\Delta V(t|\Omega_t) \leq -VaR_t(p)) - p) \right)$$

oder für kontinuierlich verteilte ΔV auch vereinfacht durch

$$ES_t(p) = -\frac{1}{p} \int_{-\infty}^{-VaR_t(p)} x f_{\Delta V(t|\Omega_t)}(x) dx$$

2.2.6. Regulatorische Auswirkungen der Risikoprognoze

Value at Risk und Expected Shortfall sind im heutigen Bankalltag wichtige Kenngrößen zur Risikosteuerung. Doch nicht nur intern, sondern auch extern sind sie von großer Bedeutung.

Campbell (2006/2007) legt dar, dass der Value at Risk auch von den Aufsichtsbehörden mittlerweile als hauptsächliches Mittel zur Beurteilung von Marktrisiken anerkannt wird. Banken müssen der Aufsichtsbehörde sowohl den VaR mit einem Konfidenzniveau von 99% auf Tagesbasis als auch auf 10-Tages-Basis mitteilen. Das Risikokapital, das es dafür zu hinterlegen gilt, ist schlussendlich entweder der

1% *VaR* auf 10-Tages-Basis oder der durchschnittliche 1% *VaR* auf 10-Tages-Basis der vergangenen 60 Tage.

Tatsächlich können die Banken so das zu hinterlegende Risikokapital durch ihre internen Modelle bestimmen. Natürlich erscheint es hier im ersten Moment ökonomisch, einen möglichst geringen *VaR* auszuweisen, um so wenig Kapital wie möglich hinterlegen zu müssen. Allerdings werden Modelle, die den Value at Risk systematisch unterschätzen, streng von den Aufsichtsbehörden bestraft und die Kapitalanforderungen erhöht. Dies wird im Verlauf dieser Arbeit im Rahmen des Ampelkonzepts noch detaillierter erläutert.

Zusammenfassend lässt sich festhalten, dass der Value at Risk sowohl für die interne Risikosteuerung als auch für externe Mindestkapitalhinterlegungen relevant ist.

2.2.7. Überprüfung der Modelle

Angesichts der enormen Bedeutung des Value at Risk ist es offensichtlich, dass eine Qualitätsprüfung der Modelle notwendig ist. Für die Aufsichtsbehörden ist es unerlässlich beurteilen zu können, ob Modelle das Risiko systematisch unterschätzen.

Für Banken, auf der anderen Seite, kann es durchaus auch von Interesse sein in Erfahrung zu bringen, wenn Modelle das Risiko überschätzen - dadurch würde an die jeweilige Bank nämlich eine höhere Mindestkapitalanforderung als eigentlich nötig gestellt.

In den folgenden Abschnitten werden wir uns also dem Thema widmen, wie man die Prognosegüte von Marktrisikomodellen quantitativ einschätzen kann.

Ein grundsätzliches Problem, das sich dabei stellt, ist jedoch, dass die Investments eines Portfolios selten konstant bleiben. Selbst wenn ein Modell Risikokennzahlen wie den Value at Risk absolut korrekt prognostiziert kann es durchaus sein, dass der zukünftige Value at Risk gänzlich anders ist, da sich bis dahin möglicherweise andere Werte im Portfolio befinden. Beispielsweise dürfte der wahre Value at Risk eines Portfolios, das ausschließlich Kassenhaltung betreibt, für den Zeitpunkt heute in einem Jahr nahezu bei Null liegen. Sollten wir jedoch das Vermögen nach einem Monat von der Kassenhaltung komplett in gehebelte Optionsscheine umschichten, so wird der wahre Value at Risk in 11 Monaten nun weitaus höher sein; die ursprüngliche Prognose wird statistisch gesehen also wesentlich zu häufig überschritten werden. Dies liegt aber nicht an einem ursprünglich falsch spezifizierten Modell, sondern lediglich an einem veränderten Portfolio.

Um diese Problematik zu umgehen gibt es zwei Möglichkeiten:

1. Wir berechnen die Risikokennzahlen stets nur auf sehr kurze Frist, z.B. für 1-Tages-Perioden. Dowd (2004) weist zwar darauf hin, dass es auch innerhalb von nur einem Tag zu deutlichen Änderungen der Bestände kommen kann, die Problematik wird dadurch aber zumindest stark vermindert.
2. Sollten die Bestandsänderungen dennoch zu erheblich sein, so kann man zum Vergleich der Prognosen auch das ursprüngliche Portfolio in Zukunft schlicht neu bewerten und vergleichen, ob diese theoretische Wertänderung mit der ursprünglichen Prognose vereinbar ist. Im Beispiel könnte man also in einem Jahr überprüfen, was der Wert eines ausschließlich aus Kassenhaltung bestehenden Portfolios wäre, und ob dieser Wert mit der Prognose zu t_0 in Einklang zu bringen ist.

Wir werden uns im Weiteren auf die erste Möglichkeit beschränken. Diese Vorgehensweise ist im akademischen und praktischen Bereich üblich und verhindert zudem Probleme, die durch überlappende Perioden bei der Berechnung der Risikokennzahlen entstehen.

2.3. Validierung basierend auf Überschreitungen

Im Folgenden werden wir uns mit Validierungsmodellen beschäftigen, die hauptsächlich auf ein korrektes Verhältnis zwischen Über- und Unterschreitung einer bestimmten Risikogröße testen. Konkreter wird hier meist überprüft, ob der berechnete Value at Risk statistisch signifikant häufiger oder seltener überschritten wird, als die von uns selbst definierte Risikowahrscheinlichkeit p^* dies annehmen lassen würde. Desweiteren werden wir mit dem Markow-Test ein Verfahren beschreiben, welches darauf prüft, ob Überschreitungen zeitlich unabhängig voneinander auftreten.

2.3.1. Time Until First Failure

Dieser Test wurde ursprünglich von Kupiec (1998) eingeführt. Von Interesse ist hier die verstrichene Zeit, bis zum ersten Mal der Value at Risk überschritten wird. Das Grundprinzip hinter dieser Methode ist intuitiv nachvollziehbar: Je geringer die Wahrscheinlichkeit einer Überschreitung, desto länger sollte es statistisch gesehen dauern, bis die erste Überschreitung eintritt.

Zufallsvariable

Wir sind an einer Zufallsvariable X interessiert, die den Zeitpunkt angibt, zu welchem die Überschreitungsfunktion I_t zum ersten mal den Wert 1 annimmt. $X \in \mathbb{N}$ ist daher geometrisch verteilt,⁴

$$\begin{aligned} X &= \min_{t \in \tau} (t \mid I_t = 1) \\ X &\sim \mathbb{G}(p) \\ E(X) &= \frac{1}{p} \end{aligned}$$

Die Wahrscheinlichkeit, dass diese Überschreitung an einem bestimmten Tag x eintritt, lässt sich wie folgt berechnen:

$$\mathcal{P}(X = x) = p(1 - p)^{x-1}$$

⁴Mit $X \in \mathbb{N}$ stellen wir heraus, dass wir von einer geometrischen Verteilung ausgehen, deren Realisationen mindestens den Wert 1 annehmen. X kann nie den Wert 0 annehmen.

Ein Value at Risk mit Konfidenzniveau 95%, d.h. Fehlerwahrscheinlichkeit $p = 5\%$, sollte im Durchschnitt das erste mal in der zwanzigsten Periode ($\frac{1}{5\%}$) überschritten werden. Sollte die tatsächliche Überschreitung im Durchschnitt wesentlich später auftreten, so überschätzt das Modell das Risiko und vice versa.

Hypothesentest

Wir haben ein statistisches parametrisches Problem der Form (X, \mathbb{P}) mit $\mathbb{P} = \{\mathbb{G}(p) \mid p \in (0, 1)\}$.

Wir konstruieren nun einen Test zur Überprüfung, ob unsere angenommene Risikowahrscheinlichkeit p^* der tatsächlichen Risikowahrscheinlichkeit p entsprechen könnte:

$$H_0 : p = p^* \text{ gegen } H_1 : p \neq p^*$$

Testfunktion

Für die Überprüfung der Hypothese bietet sich ein Likelihood-Quotienten-Test der folgenden Form an:

$$\Lambda(x) = -2 \ln \left(\left(\frac{p^*}{\hat{p}} \right) \left(\frac{1-p^*}{1-\hat{p}} \right)^{x-1} \right)$$
$$\hat{p} = \frac{1}{x}$$
$$\Lambda(x) \sim \chi_1^2$$

Der Freiheitsgrad der χ^2 -Verteilung ergibt sich aus $\dim \Theta_0 = 0$ und $\dim \Theta = 1$. Eine Herleitung für \hat{p} wird im Verlauf dieses Abschnitts aufgezeigt.

Ablehnungsbereich

Der Ablehnungsbereich ergibt sich durch das entsprechende Quantil einer χ_1^2 -Verteilung:

$$A = \left((\chi_1^2)^{-1}(1 - \alpha), +\infty \right)$$

Herleitung des Schätzers \hat{p}

Als Schätzer für p verwenden wir den Maximum Likelihood-Schätzer \hat{p} , welcher aus Optimierung der Likelihood-Funktion⁵ resultiert:

$$\begin{aligned}L(p) &= p(1-p)^{x-1} \\ \ln(L(p)) &= \ln(p) + (x-1)\ln(1-p) \\ \frac{\delta \ln(L(p))}{p} &= \frac{1}{p} - \frac{x-1}{1-p} = 0 \\ \frac{1}{p} &= \frac{x-1}{1-p} \\ \frac{1}{p}(1-p) &= x-1 \\ 1-p &= p(x-1) \\ \frac{1-p}{p} &= x-1 \\ \frac{1-p}{p} + 1 &= x \\ \frac{1-p+p}{p} &= x \\ \frac{1}{p} &= x \\ p &= \frac{1}{x}\end{aligned}$$

Berechnung kritischer Werte für x

Anhand dieses vorgestellten Tests lassen sich für ein Konfidenzniveau der Hypothese α und gegebene Wahrscheinlichkeiten p^* kritische Werte für X errechnen, bei deren Überschreitung die Nullhypothese abgelehnt werden muss.

Hierfür wird die LR-Funktion mit dem entsprechenden Quantil der χ_1^2 -Verteilung gleichgesetzt:

$$-2 \ln \left(\left(\frac{p^*}{\frac{1}{x}} \right) \left(\frac{1-p^*}{1-\frac{1}{x}} \right)^{x-1} \right) = (\chi_1^2)^{-1}(1-\alpha)$$

⁵Die Likelihood-Funktion entspricht hier der Dichtefunktion einer geometrischen Verteilung.

Für gegebene Werte für p^* und α lässt sich diese Gleichung numerisch nach x auflösen. Wir führten dies für verschiedene p^* und α durch. Eine Tabelle mit den Resultaten für kritische Werte befindet sich in Anhang A.1 und A.2.

Grafisch lässt sich die obige Gleichung visualisieren, indem man die Schnittpunkte der LR-Funktion mit einer horizontalen Geraden auf Höhe des $(1 - \alpha)$ -Quantils der χ_1^2 -Verteilung darstellt. Wir stellen dies für $p^* = 0,5\%$ und $\alpha = 5\%$ im Folgenden grafisch dar.

Der Term links des Gleichheitszeichens resultiert dabei in der konvexen Kurve des Schaubilds. Der Teil rechts des Gleichheitszeichens resultiert in der horizontalen Gerade auf Höhe des 95%-Quantils der χ_1^2 -Verteilung von ungefähr 3,84. Die x-Werte der Schnittpunkte der beiden Funktionen stellen die Lösungen der Gleichung dar:

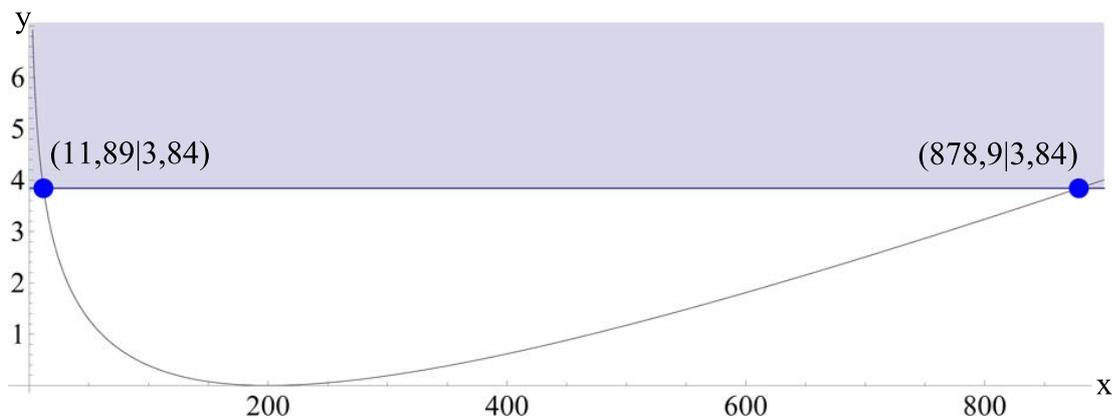


Abbildung 2.2.: TUFF LR-Test, kritische Werte von X als Schnittpunkt zwischen LR-Funktion und 95%-Quantil der χ_1^2 -Verteilung

Dies kann nun wie folgt interpretiert werden: Der Hypothesentest $H_0 : p = 0,5\%$ gegen $H_1 : p \neq 0,5\%$ sollte dann abgelehnt werden, wenn der erste Fehler vor der Periode 11,89 (in der Praxis also in der 11. Periode oder früher) oder nach der Periode 878,9 (in der Praxis also in der 879. Periode oder später) auftritt, denn genau für diese Werte nimmt die Testfunktion einen Wert an, der oberhalb des 95%-Quantils der χ_1^2 -Verteilung liegt.

Die eben beschriebene Methode dient zum Berechnen kritischer Werte für feste α und p^* . Wir führen diese Berechnung nun mehrmals für weiterhin feste α , aber variierende p^* durch. Im Folgenden eine Darstellung der kritischen Werte für X auf der y-Achse für ein festes $\alpha = 5\%$ in Abhängigkeit von verschiedenen p^* -Werten von 1% bis 10% auf der x-Achse:

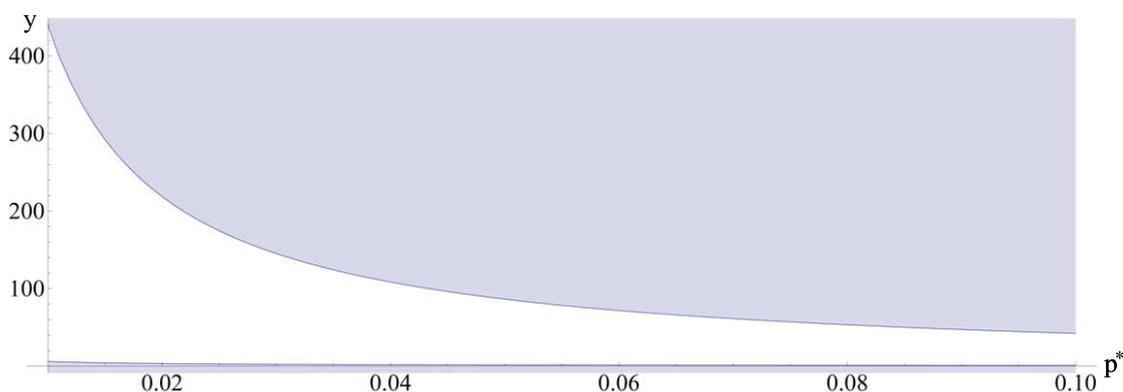


Abbildung 2.3.: Kritische x -Werte für $\alpha = 5\%$ und variable p^*

Die hellblau markierten Regionen stellen den Ablehnungsbereich, die weiß markierten Bereiche den Annahmehereich der Nullhypothese dar für das jeweils zugehörige p^* mit einer Konfidenz von 95%. Erwartungsgemäß sinken die kritischen Werte mit wachsendem p^* , der Annahmehereich wird dabei stets schmaler.

Kritik - Mangelnde Ablehnungsmöglichkeiten für große p^*

Kupiec (1998) weist darauf hin, dass für eine Konfidenz in Höhe von $1 - \alpha = 95\%$ keine Modelle mit einem Fehler von 5% oder mehr erkannt werden können. Tatsächlich weist das Modell ein Problem bei der Ablehnung von Hypothesen mit großem p^* auf. Hierbei ist nämlich der erste Fehler bereits zu einem sehr frühen Zeitpunkt zu erwarten. Wenn nun die tatsächliche Fehlerwahrscheinlichkeit nochmals größer als die angenommene Wahrscheinlichkeit ist, $p > p^*$, dann kann die Nullhypothese zugunsten einer höheren Risikowahrscheinlichkeit nur dann verworfen werden, wenn der erste Fehler extrem früh eintritt. Konkret heißt dies, dass der erste Fehler teils bei $X < 1$ eintreten müsste, um die Nullhypothese ablehnen zu können.

Da jedoch $X \in \mathbb{N}$ ist kann dies in der Praxis nicht passieren: Der erste Fehler kann schlicht nicht vor der ersten Periode auftreten. Dadurch wird es für große p^* unmöglich, eine Risikounterschätzung zu identifizieren.

Unsere Ergebnisse zeigen aber, dass dieses Problem erst später als von Kupiec berechnet auftritt; für eine Konfidenz in Höhe von 95% ist damit erst ab $p^* > 14\%$ zu rechnen. Dies kann mittels der Tabelle in Anhang A.2 nachvollzogen werden: Für $\alpha = 5\%$ unterschreitet der untere kritische Wert für X die Zahl 1 zum ersten mal bei einem p^* zwischen 14% und 15%.

Gütefunktion und statistische Macht

Wir simulieren nun Gütefunktionen des TUFF-Verfahrens, indem wir die Ablehnungsraten für diverse Kombinationen von p und p^* berechnen. Wir gehen dazu wie folgt vor:

1. Wähle $p^*, p \in \{0,01, 0,01005, 0,01001, \dots, 0,2\}$, sodass $p^* \neq p$. Setze $\alpha = 5\%$.
2. Erzeuge eine Realisation x der Zufallsvariable $X \sim \mathbb{G}(p)$
3. Berechne $\Lambda(x)$ mit dem im ersten Schritt gewählten $p^* \neq p$
4. Lehne die Nullhypothese des TUFF-Tests ab, wenn $\Lambda(x) > (\chi_1^2)^{-1}(1 - \alpha)$

Diese Prozedur wird für jede einzelne Kombination von p und p^* 1.000.000 mal durchgeführt, sodass insgesamt ca. 145 Milliarden Tests berechnet werden. Für jede Kombination von p und p^* wird nun die gesamte Anzahl an Ablehnungen der Nullhypothese durch 1.000.000 geteilt, um die relative Ablehnungshäufigkeit zu berechnen.

Für ein festes p^* entsprechen die Simulationsergebnisse einer Gütefunktion des Testverfahrens, hier am Beispiel für $p^* = 5\%$:

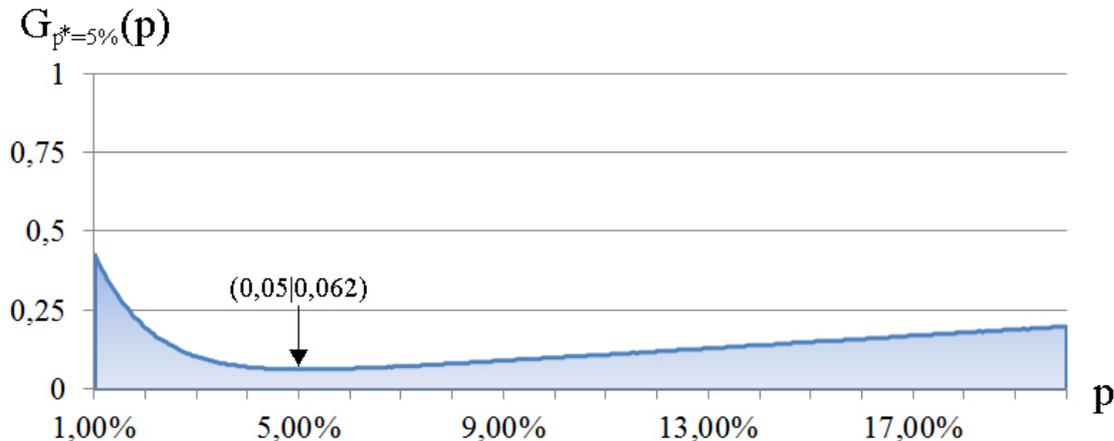


Abbildung 2.4.: Gütefunktion, $\alpha = 5\%$, $p^* = 5\%$

Erwartungsgemäß hat die Funktion ihr Minimum ungefähr⁶ im Bereich von $p = p^* = 5\%$ und umso höhere Werte, je weiter sich p von $p^* = 5\%$ entfernt. Eigentlich

⁶Eigentlich sollte die Funktion ihr Minimum exakt bei $p = p^*$ haben. Wie später noch herausgestellt wird ist der TUFF-Test hier in der Tat unpräzise.

sollte $G_{p^*=5\%}(p^*) = \alpha = 5\%$ gelten, denn für $p = p^*$ ist die Nullhypothese wahr und die Gütefunktion sollte demzufolge Wert des α -Fehlers annehmen. Tatsächlich ist aber $G_{p^*=5\%}(p^*) \approx 6,2\% \neq \alpha$. Diesem Phänomen werden wir uns später genauer widmen.

Wir erweitern nun die Darstellung der Gütefunktion, indem wir auch p^* variieren lassen. Jeder Querschnitt des folgenden Schaubilds entspricht einer Gütefunktion⁷ für das jeweilige p^* :

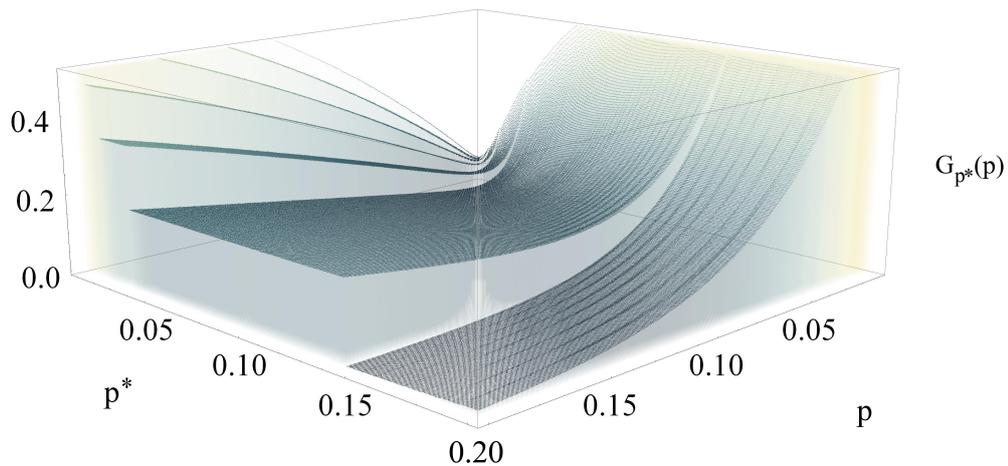


Abbildung 2.5.: Gütefunktionen für verschiedene p^* mit $\alpha = 5\%$

⁷Zur besseren Darstellung wurde die Höhenachse des Schaubilds 2.5 auf 0,5 begrenzt

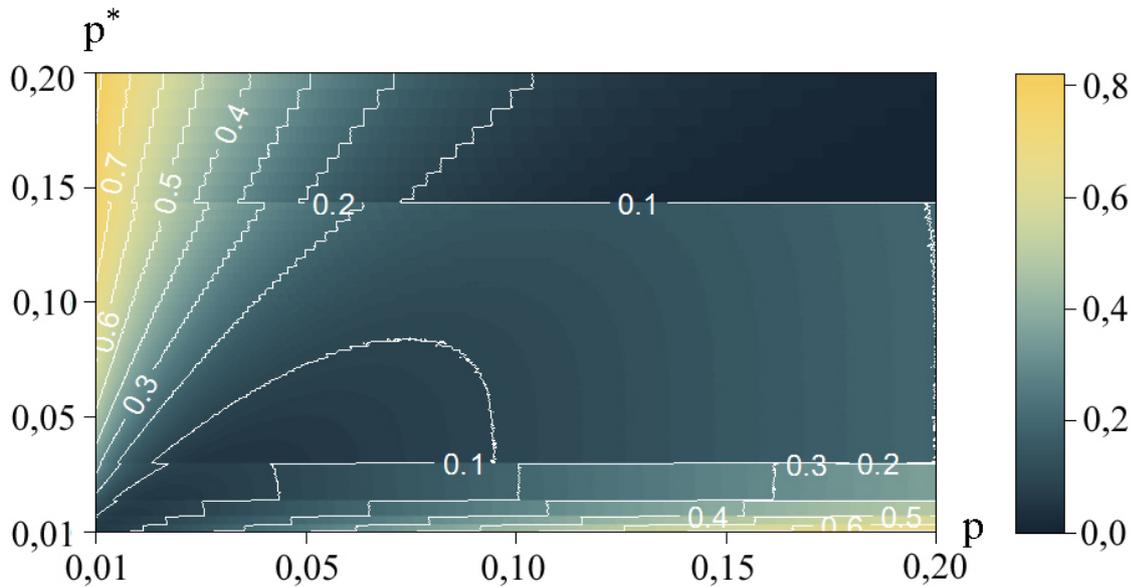


Abbildung 2.6.: Konturplot der Gütefunktionen für verschiedene p^* , $\alpha = 5\%$

Gut zu sehen ist der sprunghafte Verlauf der Funktion im Bezug auf die angenommene Wahrscheinlichkeit p^* . Dies liegt daran, dass der LR-Test Ablehnungs- und Annahmehereich zwar auf eine Teilmenge aus \mathbb{R} abbildet, geometrische Zufallsvariablen aber stets aus \mathbb{N} stammen. An bestimmten Punkten unter- oder überschreitet der Ablehnungsbereich eine Ganzzahl - und genau an diesen Stellen verändert sich die Funktion sprunghaft. Deutlich erkennbar ist auch eine auffällige Verringerung der Funktionswerte ab einer angenommenen Wahrscheinlichkeit von ca. 14,7% oder höher - hier tritt genau das oben genannte Phänomen auf, dass es effektiv keinen unteren Ablehnungsbereich mehr gibt und somit keine Risikounterschätzung mehr erkannt werden kann. Der Eintritt dieses Phänomens kann durch eine Erhöhung des Fehlers 1. Art hinausgezögert werden. Eine genaue Tabellierung kritischer Werte befindet sich in Anhang A.1.

Insgesamt nehmen die Gütefunktionen nur in sehr wenigen Konstellationen Werte größer als 0,4 an. Dies bedeutet, dass selbst falsche Modelle recht selten selbst bei großen Unterschieden zwischen p und p^* abgelehnt werden, die statistische Macht ist also meist sehr niedrig. Selbst in der extremen Kombination $p = 1\%$, $p^* = 20\%$ wird die Nullhypothese nur in knapp 82% aller Fälle abgelehnt.

Kritik an der Verteilung der Testfunktion

Durch die Simulation der Gütefunktionen zeigte sich ein weiteres kritisches Phänomen: Der spezifizierte α -Fehler des Hypothesentests wurde nicht eingehalten. Dies wird deutlich, wenn wir die Simulationsdaten im Falle eines korrekt geschätzten Modells betrachten - also in den Fällen, in denen die geschätzte Wahrscheinlichkeit genau der tatsächlichen Wahrscheinlichkeit des datengenerierenden Prozesses entspricht, $p = p^*$.

Die konkreten Schritte einer solchen Simulation sind dabei:

1. Wähle einen bestimmten Wert für $p \in \{0,01, 0,02, \dots, 0,20\}$. Setze $\alpha = 5\%$.
2. Erzeuge eine Realisation x der Zufallsvariable $X \sim \mathbb{G}(p)$
3. Berechne $\Lambda(x)$ für $p^* = p$
4. Lehne die Nullhypothese des TUFF-Tests ab, wenn $\Lambda(x) > (\chi_1^2)^{-1}(1 - \alpha)$

Die Schritte 2-4 werden 1.000.000 mal wiederholt. Die relative Ablehnungshäufigkeit ergibt sich dann aus der Anzahl aller Ablehnungen der Nullhypothese geteilt durch 1.000.000. Anschließend wird ein anderer Wert für p gewählt die Prozedur erneut durchgeführt.

Die Variable X simuliert also den Zeitpunkt der ersten Value at Risk-Überschreitung in einem Portfolio, das in jeder einzelnen Periode eine unabhängige Value at Risk-Überschreitungswahrscheinlichkeit von p hat. Der entscheidende Schritt in der Simulation ist der Dritte, denn dieser besagt, dass $p^* = p$ gesetzt wird. Das heißt, dass das Modell von einer Überschreitungswahrscheinlichkeit ausgeht, die gleich der tatsächlichen Überschreitungswahrscheinlichkeit ist. Das Modell ist also korrekt spezifiziert und sollte eigentlich nicht abgelehnt werden.

Dies zeigt sich auch, wenn wir die Hypothese des TUFF-Tests betrachten: $H_0 : p = p^*$ gegen $H_1 : p \neq p^*$. Mit $p^* = p$ gilt $H_0 : p = p$ gegen $H_1 : p \neq p$. Die Nullhypothese ist eine Tautologie und somit immer wahr. Wenn der TUFF-Test die Nullhypothese dennoch ablehnt, so begeht er einen Fehler 1. Art. Dies sollte in genau α Prozent aller Fälle passieren.

Völlig unabhängig von p sollte der TUFF-Test dann im Durchschnitt eigentlich stets in 5% aller Fälle der simulierten Konstellationen zur Ablehnung der Nullhypothese führen. Die Ergebnisse der Simulation sind im folgenden Schaubild dargestellt:

Prozentuale Ablehnungen der Nullhypothese

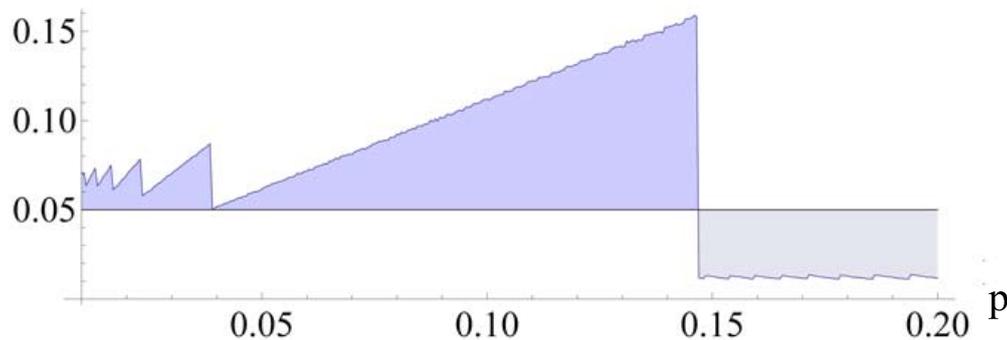


Abbildung 2.7.: Ablehnungsraten eines korrekten Modells, $\alpha = 5\%$

Es ist sehr deutlich zu erkennen, dass die Ablehnungsrate der Nullhypothese deutlich in Abhängigkeit von p variiert und fast nie genau bei ihrem eigentlich zu erwartenden Wert in Höhe von 5% liegt.

Die gleiche Simulation erweitern wir nun, indem wir nicht mehr von einem festen $\alpha = 5\%$ ausgehen, sondern auch α zwischen 1% und 20% variieren lassen. Zur besseren Interpretierbarkeit des Schaubildes stellen wir nun auf der Höhenachse nicht mehr die relative Ablehnungshäufigkeit der Nullhypothese dar, sondern die Differenz zwischen α und der relativen Ablehnungshäufigkeit. Diese sollte im Durchschnitt eigentlich konstant 0 sein.

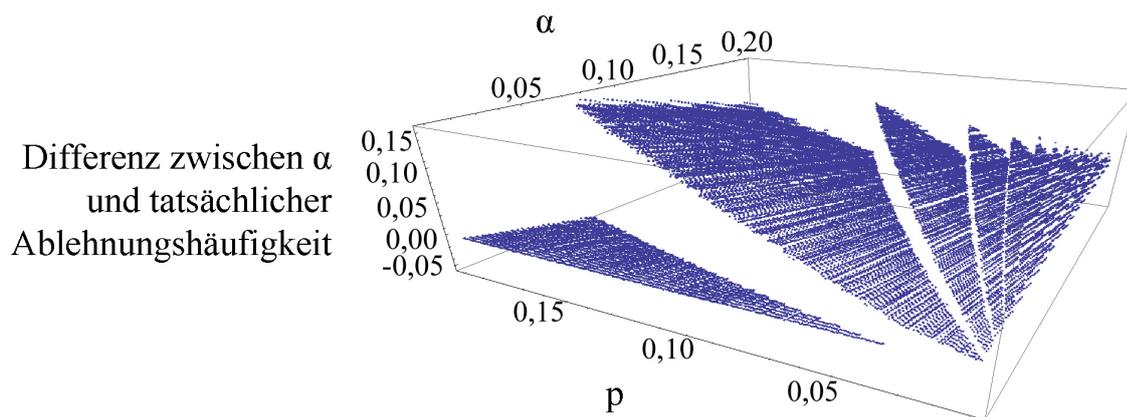


Abbildung 2.8.: Ablehnungsraten eines korrekten Modells für verschiedene Signifikanzniveaus

Auch hier bestätigt sich, dass der TUFF-Test nicht in der Lage ist, einen vorgegebenen α -Fehler einzuhalten. Es zeigt sich, dass die Abweichungen charakteristisch für unterschiedlichste Parameterkombinationen sind. Das Problem ist am geringsten, wenn das Signifikanzniveau der Hypothese ungefähr der Fehlerwahrscheinlichkeit des Modells entspricht. Im Zweifelsfall sollte das Signifikanzniveau höher angesetzt werden als die Fehlerwahrscheinlichkeit des Modells, denn speziell im umgekehrten Fall weicht der Fehler erster Art stark von seinem zu erwartenden Wert ab.

Zusammenfassung des Testverfahrens

Hypothesentest	
1. Testproblem	$H_0 : p = p^*$ gegen $H_1 : p \neq p^*$
2. Testfunktion	$\Lambda(x) = -2 \ln \left(\left(\frac{p^*}{\hat{p}} \right) \left(\frac{1-p^*}{1-\hat{p}} \right)^{x-1} \right)$ $\Lambda(x) \sim \chi_1^2$
	Der Stichprobenwert x ist dabei die Realisation der Zufallsvariablen $X = \min_{t \in \tau} (t I_t = 1)$.
3. Ablehnungsbereich	$A = \left((\chi_1^2)^{-1}(1 - \alpha), +\infty \right)$
4. Entscheidungsregel	H_0 ablehnen, wenn die Testfunktion im Ablehnungsbereich liegt. Ansonsten H_0 nicht ablehnen.

Insgesamt ist der Time Until First Failure-Test wenig praxistauglich, da er kaum zwischen verschiedenen Alternativhypothesen zu unterscheiden vermag. Unter bestimmten Konstellationen ist der Likelihood-Ratio-Test sehr unzuverlässig. Seine statistische Macht ist selbst unter Idealbedingungen⁸ niedrig. Dies ist nicht verwunderlich - die einzige Information, die bei diesem Verfahren beachtet wird, ist

⁸Idealbedingungen bezieht sich vor allem darauf, dass Value at Risk-Überschreitungen unabhängig sind und mit gleicher Wahrscheinlichkeit zu jedem Zeitpunkt auftreten

die Zeit bis zur ersten Überschreitung des Value at Risk. Zwar liegen in aller Regel noch wesentlich mehr Informationen vor, beispielsweise die Höhe der Überschreitung sowie möglicherweise noch weitere Überschreitungen, diese werden jedoch vollständig im Test vernachlässigt. Wir werden im Folgenden daher Verfahren untersuchen, die weitere Informationen einbeziehen.

2.3.2. Proportion of Failure

Der Proportion of Failure-Test ist, wie von Kupiec (1998) und Campbell (2006/2007) beschrieben, eine logische Weiterentwicklung des Time Until First Failure-Verfahrens. Statt reiner Fokussierung auf den erste Fehler werden nun zusätzliche Informationen einbezogen, indem eine Teststatistik konstruiert wird, die alle in der Stichprobe vorkommenden Value at Risk-Überschreitungen einbezieht.

Zufallsvariable

Wir betrachten nun nicht mehr nur die erste Überschreitung, sondern die Summe aller Überschreitungen, I :⁹

$$I \sim \mathbb{B}(T, p)$$
$$f_I(k) = \binom{T}{k} p^k (1-p)^{T-k} \quad \text{für } k = 0, 1, \dots, T$$

mit den ersten beiden Momenten

$$E(I) = Tp$$
$$Var(I) = Tp(1-p)$$

Hypothesentest

Für unser statistisches parametrisches Problem (I, \mathbb{P}) mit $\mathbb{P} = \{\mathbb{B}(T, p) \mid p \in (0, 1)\}$ können wir einen Test konstruieren und basierend auf allen vorliegenden Überschreitungen testen, ob unsere angenommene Fehlerwahrscheinlichkeit der empirischen Schätzung der realen Fehlerwahrscheinlichkeit entsprechen könnte:

$$H_0 : p = p^* \quad \text{gegen} \quad H_1 : p \neq p^*$$

⁹siehe dazu auch Abschnitt 2.2.4

Testfunktion

Ähnlich dem Time Until First Failure-Verfahren bietet sich ein Likelihood-Quotienten-Test zum Test der Hypothese an. Mit i als Realisation der Zufallsvariable I nutzen wir folgende Testfunktion:

$$\Lambda(T, i) = -2 \ln \left(\left(\frac{(1-p^*)}{(1-\hat{p})} \right)^{T-i} \left(\frac{p^*}{\hat{p}} \right)^i \right)$$

$$\hat{p} = \frac{i}{T}$$

$$\Lambda(T, i) \sim \chi_1^2$$

Mit $d = 1$ da $\dim \Theta_0 = 0$ und $\dim \Theta = 1$.

Ablehnungsbereich

Erneut ergibt sich der Ablehnungsbereich aus dem jeweiligen Quantil der χ_1^2 -Verteilung:

$$A = \left((\chi_1^2)^{-1}(1 - \alpha), +\infty \right)$$

Herleitung des Schätzers \hat{p}

Als Schätzer für p kommt erneut ein Maximum Likelihood Schätzer \hat{p} zum Einsatz, den wir durch Ableiten, Nullsetzen und Auflösen nach p der Likelihood-Funktion erhalten. i ist dabei die Realisation der Zufallsvariablen I .

$$\begin{aligned}L(T, i, p) &= \binom{T}{i} p^i (1-p)^{T-i} \\ \ln(L(T, i, p)) &= \ln\left(\binom{T}{i}\right) + i \ln(p) + (T-i) \ln(1-p) \\ \frac{\delta \ln(L(T, i, p))}{\delta p} &= \frac{i}{p} - \frac{T-i}{1-p} = 0 \\ \frac{i}{p} &= \frac{T-i}{1-p} \\ i(1-p) &= (T-i)p \\ \frac{1-p}{p} &= \frac{T-i}{i} \\ &= \frac{T}{i} - \frac{i}{i} \\ &= \frac{T}{i} - 1 \\ \frac{1-p}{p} + 1 &= \frac{T}{i} \\ \frac{1-p+p}{p} &= \frac{T}{i} \\ \frac{1}{p} &= \frac{T}{i} \\ p &= \frac{i}{T}\end{aligned}$$

Gütefunktion und statistische Macht

Wir simulieren nun die Ablehnungsraten verschieden spezifizierter Modelle gemäß folgender Vorschrift:¹⁰

1. Setze $\alpha = 5\%$, $T = 250$
2. Wähle p^* , $p \in \{0,01, 0,01005, 0,01001, \dots, 0,2\}$, sodass $p^* \neq p$
3. Erzeuge eine Realisation i der Zufallsvariable $I \sim \mathbb{B}(T, p)$
4. Berechne $\Lambda(i)$ mit dem im ersten Schritt gewählten $p^* \neq p$
5. Lehne die Nullhypothese des POF-Tests ab, wenn $\Lambda(i) > (\chi_1^2)^{-1}(1 - \alpha)$

Erneut wird die Ablehnungsquote der Nullhypothese durch 1.000.000 Simulationen für jede Kombination von p und p^* berechnet.

Analog zum vorherigen Abschnitt lässt sich auch hier eine Gütefunktion aus den Daten für ein festes p^* , im beispiel mit $p^* = 5\%$, konstruieren:

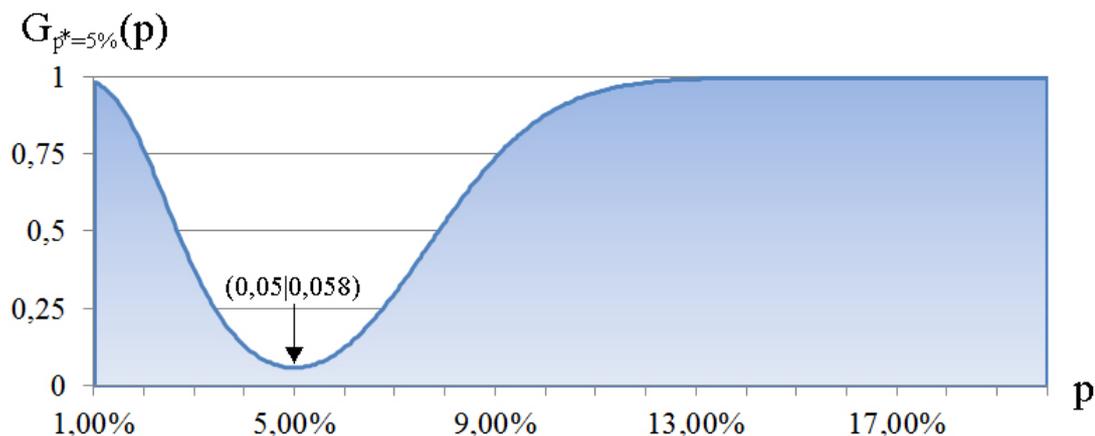


Abbildung 2.9.: Gütefunktion für $T = 250$, $\alpha = 5\%$, $p^* = 5\%$

Wir sehen, dass $G_{p^*}(p^*) = 5,8\% \neq \alpha$ gilt, auch hier wird also der α -Fehler des Hypothesentests nicht eingehalten. Wir werden im kommenden Abschnitt genauer darauf eingehen.

¹⁰Die tatsächlich durchgeführte Simulation wurde wesentlich stärker hinsichtlich ihrer Berechnungseffizienz optimiert und weicht deshalb teilweise von den aufgeführten Schritten ab, funktioniert aber nach dem gleichen Prinzip und liefert die selben Resultate.

Wir werden nun wieder die Darstellung der Gütefunktion erweitern, indem wir p^* variieren und somit verschiedene Gütefunktionen in einem Schaubild darstellen:

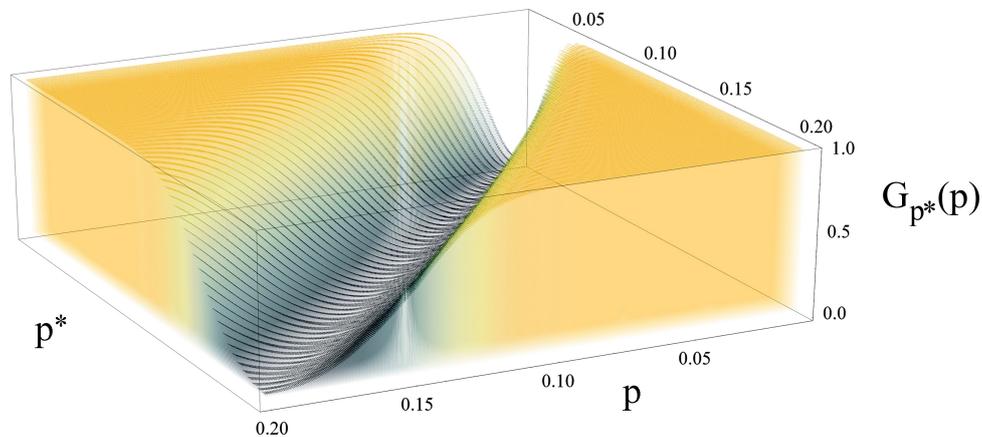


Abbildung 2.10.: Gütefunktionen für variable p^* mit $T = 250$, $\alpha = 5\%$

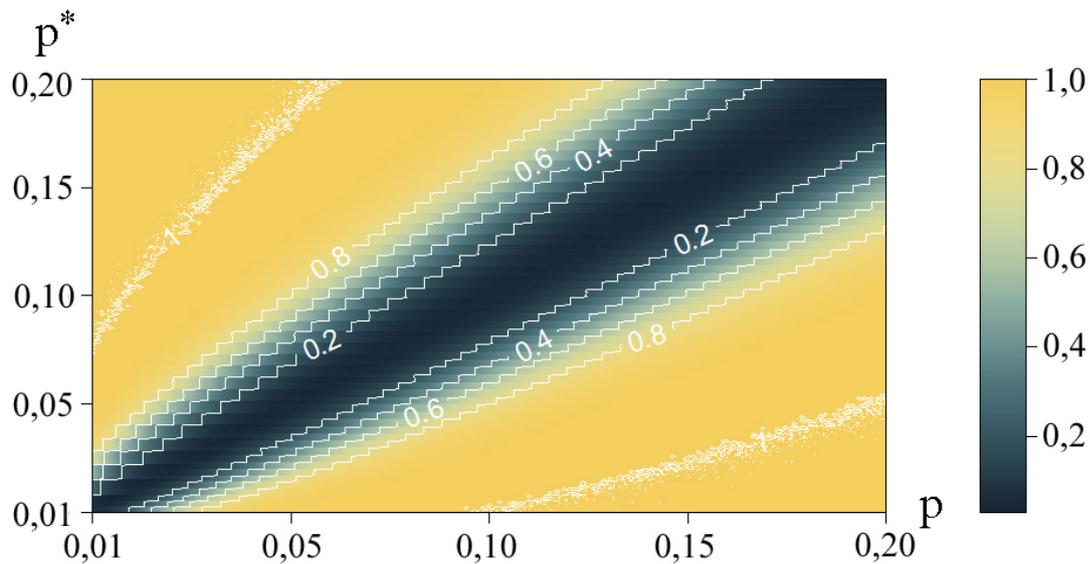


Abbildung 2.11.: Konturplot der Gütefunktionen für variable p^* mit $T = 250$,
 $\alpha = 5\%$.

Auffällig ist hier zum einen, dass der Verlauf der Funktion relativ konstant ist. Es gibt keine erkennbaren Sprünge mehr, wie dies noch beim TUFF-Test der Fall war. Zudem erreicht die Funktion recht schnell hohe Werte bis hin zu Niveaus von

100%. Im Bezug auf die statistische Macht ist der POF-Test also erwartungsgemäß wesentlich besser als der TUFF-Test.

Abweichungen vom erwarteten Fehler 1. Art

Wir betrachten abschließend noch die Ablehnungsraten eines korrekt geschätzten Modells. Diese sollte bei den Simulationsdaten eigentlich konstant auf dem α -Niveau von 5% liegen. Wir gehen dabei nach diesem Prinzip vor:

1. Setze $\alpha = 5\%$, $T = 250$
2. Wähle $p \in \{0,01, 0,01005, 0,01001, \dots, 0,2\}$ und setze $p^* = p$
3. Erzeuge eine Realisation i der Zufallsvariable $I \sim \mathbb{B}(T, p)$
4. Berechne $\Lambda(i)$ mit dem im ersten Schritt gewählten $p^* = p$
5. Lehne die Nullhypothese des POF-Tests ab, wenn $\Lambda(i) > (\chi_1^2)^{-1}(1 - \alpha)$

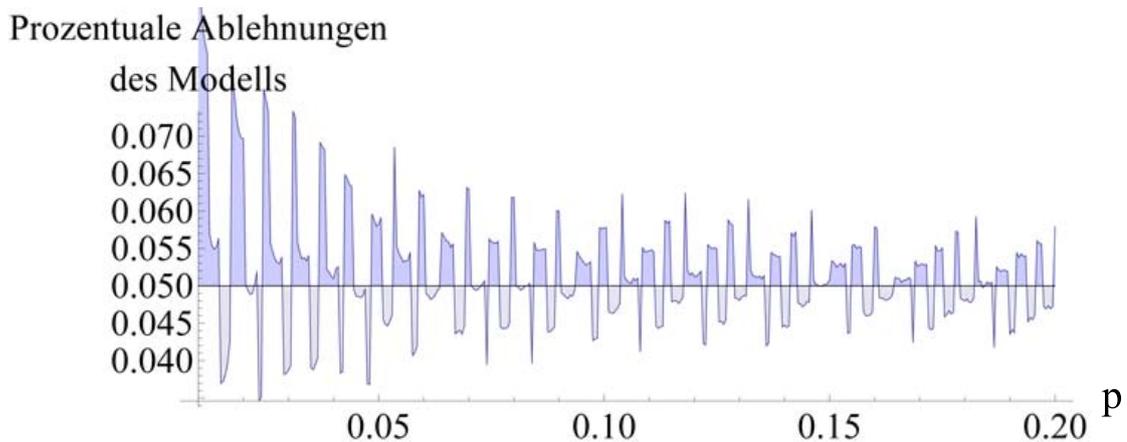


Abbildung 2.12.: Ablehnungsraten eines korrekten Modells, $T = 250$

Es zeigt sich, dass es auch hier Abweichungen vom theoretisch zu erwartenden Niveau der Ablehnungsraten gibt. Erneut lieben selbige nicht auf dem eigentlich zu erwartenden, konstanten 5%-Niveau. Die Abweichungen werden jedoch umso kleiner, je größer die reale und korrekt geschätzte Wahrscheinlichkeit für eine Value at Risk-Überschreitung wird. Die Fehler erweisen sich zudem immerhin als kleiner im Vergleich zum TUFF-Test.

Zusammenfassung des Testverfahrens

Hypothesentest	
1. Testproblem	$H_0 : p = p^*$ gegen $H_1 : p \neq p^*$
2. Testfunktion	$\Lambda(T, i) = -2 \ln \left(\left(\frac{(1-p^*)}{(1-\hat{p})} \right)^{T-i} \left(\frac{p^*}{\hat{p}} \right)^i \right)$ $\hat{p} = \frac{i}{T}$ $\Lambda(T, i) \sim \chi_1^2$
3. Ablehnungsbereich	$A = \left((\chi_1^2)^{-1}(1-\alpha), +\infty \right)$
4. Entscheidungsregel	H_0 ablehnen, wenn die Testfunktion im Ablehnungsbereich liegt. Ansonsten H_0 nicht ablehnen.

Das POF-Verfahren ist insgesamt unbedingt dem TUFF-Verfahren vorzuziehen, denn seine statistischen Eigenschaften erweisen sich wie erwartet als erheblich besser. Dies ist kaum verwunderlich, da es auf eine größere Anzahl an Informationen als das TUFF-Verfahren zurückgreift.

Nichtsdestotrotz wird jedoch auch beim POF-Test eine hohe Menge an Daten benötigt, um verlässliche Ergebnisse berechnen zu können.

2.3.3. Baseler Ampelansatz / Traffic Light Approach

Wir kommen mit dem sogenannten "Traffic Light Approach" nun zu einem von der Bundesanstalt für Finanzdienstleistungsaufsicht verwendeten Verfahren, welches auf Überschreitungen basiert. Es ist als maßgebliches Verfahren zur Beurteilung von Marktrisikomodellen aus Sicht der Aufsichtsbehörde von besonderer Bedeutung.

Schröder & Schulte-Mattler (1997) liefern eine sehr gute Einführung in dieses Konzept. Auf Vorschlag des Baseler Ausschusses für Bankenaufsicht dient dieses Verfahren zur Beurteilung der Qualität eines Risikomodells aus Sicht der Aufsichtsbehörden. Konkret heißt dies, dass die Ergebnisse des Baseler Ampelansatzes Auswirkungen auf die Kapitalanforderungen an das jeweilige Institut haben.

Ähnlich dem Proportion of Failure-Test wird auch hier versucht, ein Risikomodell basierend auf den beobachteten Value at Risk-Überschreitungen zu beurteilen. Von Interesse ist also auch hier die Zufallsvariable der Summe aller Überschreitungen I , wie sie bereits im vorherigen Kapitel definiert wurde.

Vorgeschrieben wird nun eine Datenbasis von 250 Handelstagen. Von Interesse ist der Value at Risk auf einem Konfidenzniveau in Höhe von 99%. Dieser sollte in einem korrekten Modell, wie im vorherigen Kapitel beschrieben, im Durchschnitt $250 \cdot 1\% = 2,5$ mal überschritten werden.

Jedoch ist es selbst bei einwandfreien Modellen sehr wahrscheinlich, eine höhere Anzahl an Überschreitungen zu erhalten. Die einzelnen Wahrscheinlichkeiten, in einem richtig spezifizierten Modell eine bestimmte Anzahl an Überschreitungen zu erhalten, lassen sich an der Dichtefunktion der entsprechenden Binomialverteilung ablesen:

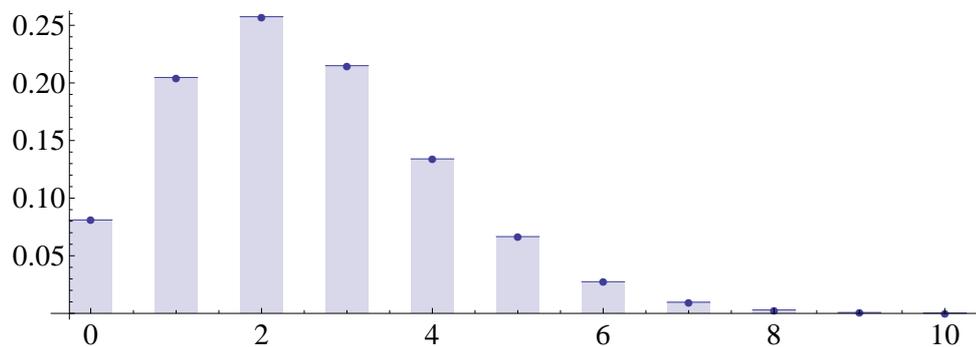


Abbildung 2.13.: Diskrete Dichtefunktion einer Binomialverteilung, $T = 250, p = 0,01$.

Selbst korrekte Modelle können also mit durchaus realistischer Wahrscheinlichkeit eine größere Anzahl an Überschreitungen generieren. Dies führt zu der Frage, ab wie vielen Überschreitungen man ein Modell ablehnen sollte.

Wenn beispielsweise alle Modelle abgelehnt werden würden, die 4 oder mehr Überschreitungen generieren, dann würde man damit auch zahlreiche korrekte Modelle ablehnen. Die Wahrscheinlichkeit, dass ein richtiges Modell 4 oder mehr Fehler in den 250 Tagen generiert liegt dabei bei $1 - F_I(3) \approx 24,19\%$.

Diese Fehlerwahrscheinlichkeit lässt sich in Abhängigkeit der Ablehnungsgrenze visualisieren:

Fehlerwahrscheinlichkeit

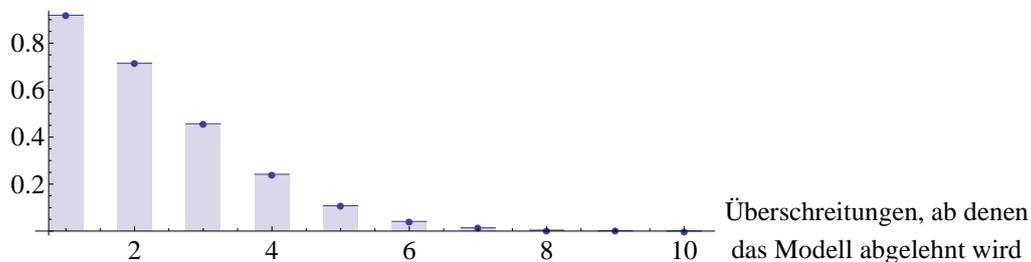


Abbildung 2.14.: Hypothetische Fehlerwahrscheinlichkeiten beim TLA-Verfahren

Setzt man nun also den Ablehnungsbereich zu niedrig an, so wird man auch sehr viele Modelle ablehnen, die eigentlich richtig wären. Setzt man ihn auf der anderen Seite zu hoch an, so akzeptiert man zu viele Modelle, die in Wirklichkeit eigentlich falsch sind.

Bedingt durch diese Problematik entschied sich die Kommission dazu, nicht eine einzelne statische Grenze im Bezug auf die maximale Anzahl an Überschreitungen zu setzen, sondern verschiedene Zonen in Abhängigkeit von den realisierten Überschreitungen zu definieren. Je nachdem, in welche Zone ein Modell fällt, entstehen dabei unterschiedliche Konsequenzen. Diese können von höheren Eigenkapitalanforderungen bis hin zu einem Verbot, das bestehende Modell weiterhin zu nutzen, reichen.

Zone	Überschreitungen	Erhöhung des Multiplikators	Kumulierte Wahrscheinlichkeit
Grüne Zone	0	0,00	8,11%
	1	0,00	28,58%
	2	0,00	54,32%
	3	0,00	75,81%
	4	0,00	89,22%
Gelbe Zone	5	0,40	95,88%
	6	0,50	98,63%
	7	0,65	99,60%
	8	0,75	99,89%
	9	0,85	99,97%
Rote Zone	≥ 10	1,00	99,99% - 100%

Tabelle 2.1.: Quelle: *Supervisory framework for the use of "Backtesting" in conjunction with the internal models approach to market risk capital requirements* (1998)

Die grüne, gelbe und rote Zone wurden dabei in Abhängigkeit von den kumulativen Wahrscheinlichkeiten eingeteilt.

Ein korrektes Modell wird mit 89,2% in die grüne Zone gelangen und mit einer Wahrscheinlichkeit von 99,97% in die grüne oder in die gelbe Zone. Die Wahrscheinlichkeit, dass ein korrektes Modell in die rote Zone gelangt, beträgt also nur 0,03%. In solchen Fällen besteht jedoch die Möglichkeit, den Aufsichtsbehörden glaubhaft darzulegen, dass das betroffene Modell eigentlich dennoch korrekt ist, um etwaige Konsequenzen zu vermeiden.

Die kumulativen Wahrscheinlichkeiten, in einem korrekten Modell höchstens eine bestimmte Anzahl an Überschreitungen zu erhalten, lassen sich auch an folgendem Schaubild ablesen:

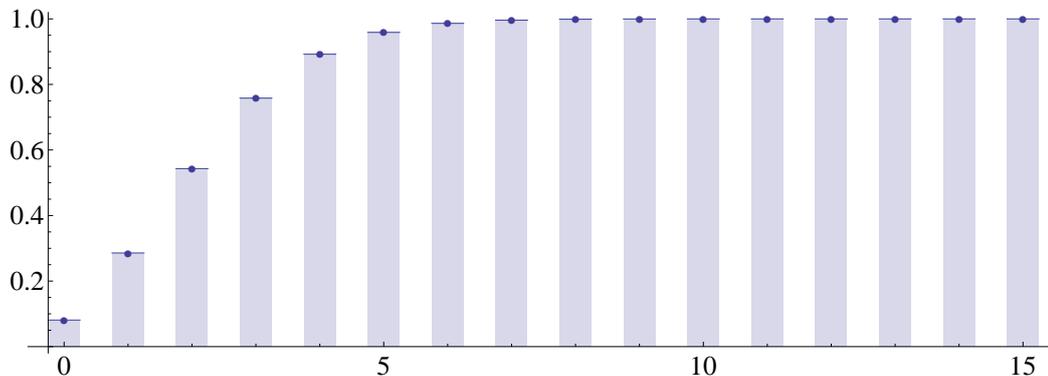


Abbildung 2.15.: Verteilungsfunktion der Binomialverteilung mit $T = 250, p = 0,01$

Das Baseler Ampelkonzept ist letztlich eine Art Weiterentwicklung des Proportion of Failure-Verfahrens. Mathematisch basiert es auf den selben Techniken und Gegebenheiten und führt lediglich eine Interpretation der Ergebnisse aus Risikogesichtspunkten mitsamt Auswirkung auf Eigenkapitalanforderungen ein. Die Güte der Ergebnisse sind somit aber eng Verbunden mit den Ergebnissen des POF-Verfahrens. Es leidet deshalb unter den bereits im vorherigen Abschnitten präsentierten Kritikpunkten. Wie wir im Verlauf dieser Arbeit herausstellen werden gibt es zahlreiche Verfahren, deren Eigenschaften sich besser für die Validierung von Risikomodellen eignen. Es ist durchaus vorstellbar, dass die Aufsichtsbehörden in künftigen Novellierungen vermehrt andere Modelle fokussieren werden.

2.3.4. Magnitude of Loss Function

Das nun folgende Verfahren orientiert sich an Goethe Universität (2004). Wir erweitern es jedoch um eine Korrektur der Standardabweichung.

Das in der Literatur eher selten diskutierte MLF-Verfahren erweitert die Überschreitungsfunktion I_t und lässt auch das Ausmaß einer Überschreitung mit einfließen. In welcher Art man das Ausmaß in die Funktion einfließen lässt ist vielfältig wählbar; wir entscheiden uns für folgende Funktion:

$$\iota_t(\Delta V_t) = \begin{cases} 1 + \left[\frac{VaR_t(p^*) + \Delta V_t}{\sigma_t} \right]^2 & \text{wenn } \Delta V_t \leq -VaR_t(p^*) \\ 0 & \text{sonst.} \end{cases}$$
$$\iota(\Delta V_t) = \sum_{t \in \tau} \iota_t(\Delta V_t)$$

Wie Lopez (1998) beschreibt ist die genaue Verteilung von ι unbekannt. Indem wir im Gegensatz zur gängigen Praxis den Abweichungsterm durch die Standardabweichung des Portfolios zum jeweiligen Zeitpunkt dividieren erreichen wir aber zumindest eine gewisse parametrische Unabhängigkeit bei der Simulation kritischer Werte. σ_t ist dabei die Standardabweichung der Verteilung von ΔV_t .

Wir müssen uns nunmehr lediglich auf eine Verteilungsform festlegen und können so kritische Werte für ι bestimmen.

Im Folgenden berechnen wir kritische Werte für ι mit der Annahme, dass ΔV normalverteilt ist. In diesem Fall ist σ_t die Standardabweichung der Normalverteilung von ΔV zum Zeitpunkt t . Wir gehen bei der Simulation wie folgt vor:

1. Setze $p \in \{0,001, 0,01, 0,05, 0,1\}$, $T = 250$
2. Setze $\iota = 0$
3. Erzeuge T standardnormalverteilte Werte $\Delta v_t, t \in \{1, \dots, T\}$.
4. Für jedes $\Delta v_t < -VaR_t(p)$, addiere $1 + (VaR_t(p) + \Delta v_t)^2$ zu ι hinzu.¹¹

Schritte 2-4 werden 1.000.000 mal wiederholt. Der kritische Wert ergibt sich dann aus dem $(1 - \alpha)$ -Quantil der berechneten ι -Werte. Anschließend kann die gesamte Prozedur für ein anderes p wiederholt werden.

¹¹Man beachte, dass die Daten hier aus einer Standardnormalverteilung stammen und die Verteilung somit eine Standardabweichung von 1 hat. In der Simulation muss der Abweichungsterm somit nicht mehr durch die Standardabweichung dividiert werden.

Es ergeben sich folgende kritische Werte für ι in Abhängigkeit von p und α :

	$\alpha = 0,1\%$	$\alpha = 1\%$	$\alpha = 5\%$	$\alpha = 10\%$
$p = 0,1\%$	3,7697	2,3815	1,3313	1,0747
$p = 1\%$	11,4310	8,7251	6,6593	5,6623
$p = 5\%$	34,5357	29,1690	24,9571	22,8517
$p = 10\%$	60,6839	53,5564	47,5656	44,5336

Eine grafische Darstellung der kritischen Werte in Abhängigkeit von p und α liefert interessante Erkenntnisse:

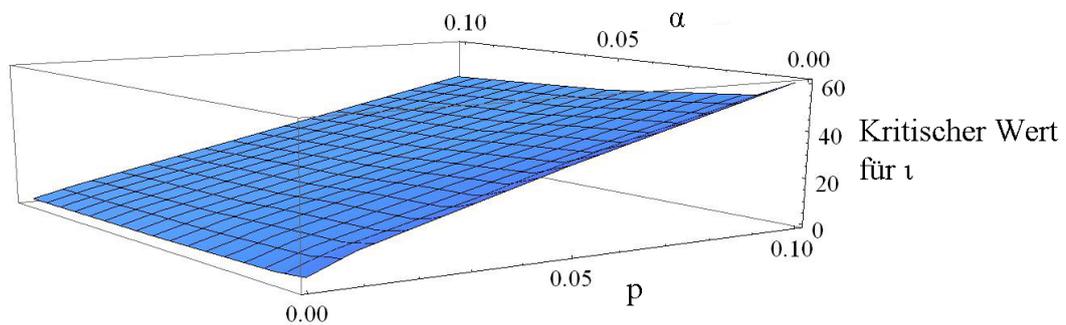


Abbildung 2.16.: Kritische Werte für $T = 250$

Der Zusammenhang scheint hier annähernd linear zu sein. Es bietet sich daher an, kritische Werte K_ι für ι mit $T = 250$ in Abhängigkeit von p und α per Regression zu approximieren. Wir nutzen dazu folgendes Regressionsmodell:

$$K_\iota = a + b \cdot p + c \cdot \alpha + \epsilon$$

K_ι , p und α sind dabei aus der Simulation bekannt, ϵ stellt das Regressionsresiduum dar und die Parameter a, b, c sollen durch die Regression gefunden werden und Aufschluss über den Zusammenhang der Werte geben. Die Methode kleinster Quadrate liefert für das Regressionsmodell:

$$K_\iota = 6.098 + 463,213 \cdot p - 75,415 \cdot \alpha + \epsilon$$

Selbstverständlich ist eine Approximation der kritischen Werte mittels dieser Regression weniger präzise als die eigentlichen simulierten Werte. Die Berechnung kritischer Werte mit der angegebenen Gleichung sollte daher nur genutzt werden, wenn Kombinationen von p und α benötigt werden, die nicht an der von uns angegebenen Tabelle ablesbar sind.

Zusammenfassung des Testverfahrens

Hypothesentest	
1. Testproblem	H_0 : Der Value at Risk unterschätzt das Risiko nicht signifikant gegen H_1 : Der VaR unterschätzt das Risiko signifikant
2. Testfunktion	$\iota(\Delta v_t) = \sum_{t \in \tau} \iota_t(\Delta v_t)$
3. Ablehnungsbereich	$A = \left(\widehat{F}_{\text{MLF}}^{-1}(1 - \alpha), +\infty \right)$ $\widehat{F}_{\text{MLF}}^{-1}$ ist inverse Verteilungsfunktion von $\iota(\Delta v_t)$ unter H_0
4. Entscheidungsregel	H_0 ablehnen, wenn die Testfunktion im Ablehnungsbereich liegt. Ansonsten H_0 nicht ablehnen.

2.3.5. Markow-Test auf Unabhängigkeit

Eine Value at Risk-Überschreitung wird oft als ein Worst Case-Szenario bei Finanzinstituten angesehen. Damit keine Gefahr für die Solvenz eines Instituts besteht sollten eventuelle Überschreitungen daher nicht nur in ihrer durchschnittlichen Häufigkeit den Prognosen entsprechen, sondern sie sollten zudem nur zeitlich unabhängig voneinander auftreten. Finanzielle Aufwendungen für zeitlich stark konzentrierte Überschreitungen könnten ansonsten das Kapital eines Instituts schnell übersteigen.

Die Auswirkungen von serieller Abhängigkeit lassen sich gut veranschaulichen. Betrachten wir zuerst eine Folge von Realisierungen unabhängig bernoulliverteilter Zufallsvariablen mit einer Erfolgswahrscheinlichkeit in Höhe von 50%:

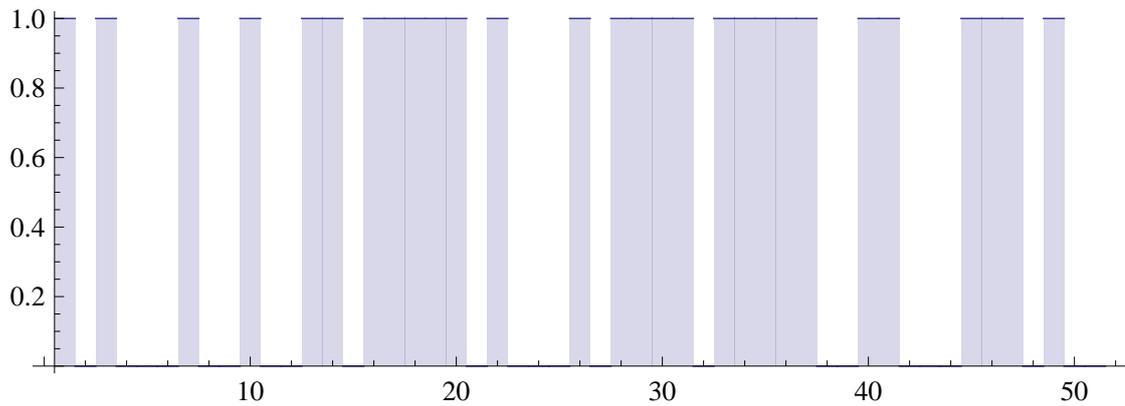


Abbildung 2.17.: Unabhängiger Bernoulli-Prozess

Die Wahrscheinlichkeit, zu einem bestimmten Zeitpunkt eine 1 (respektive 0) zu erhalten erscheint hier weitestgehend unabhängig von dem jeweils vorherigen Wert. Es ist keine wirkliche Struktur im Schaubild zu erkennen.

Schauen wir uns nun einen stochastischen Prozess mit abhängig bernoulliverteilten Zufallsvariablen an:¹²

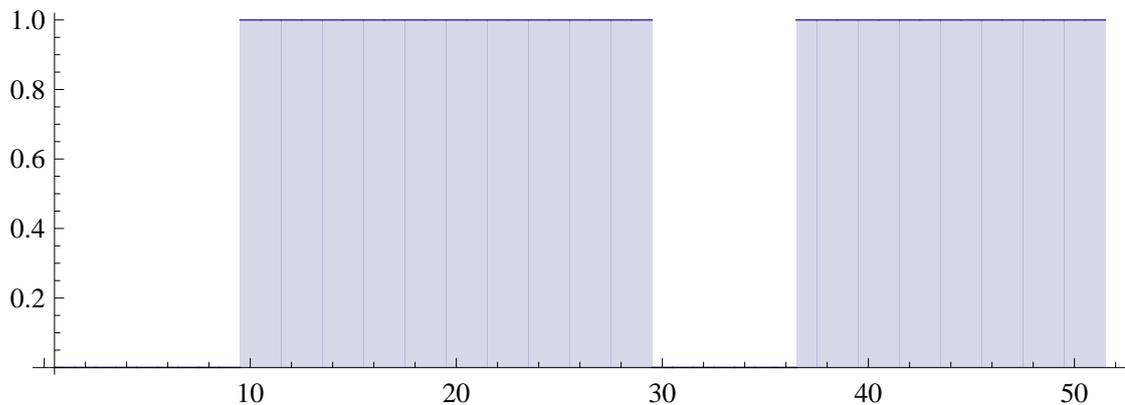


Abbildung 2.18.: Abhängiger Bernoulli-Prozess

Der Unterschied ist hier sehr klar ersichtlich. Ein einmal angenommener Wert 0 oder 1 bleibt für eine lange Zeit erhalten. Die Wahrscheinlichkeit, zum Zeitpunkt t also einen bestimmten Wert zu realisieren, hängt hier stark von der vorherigen Realisation in $t - 1$ ab.

¹²Zu jedem Zeitpunkt wird eine Realisation einer bernoulliverteilten Zufallsvariable mit Erfolgswahrscheinlichkeit p erzeugt. War der vorherige realisierte Wert 1, so wurde dafür $p = 0,9$ gesetzt, andernfalls $p = 0,1$.

Die den beiden Schaubildern unterliegenden Bernoulliprozesse haben jedoch beide die selbe durchschnittliche Erfolgswahrscheinlichkeit in Höhe von 0,5, obwohl die konkreten Prozesse anschließend grundlegend unterschiedlich aussehen. Übertragen auf unsere Risikomodelle würde dies bedeuten, dass es so zu sehr langen Zeiten kommen könnte, in denen der Value at Risk jeden Tag überschritten wird, obwohl wir die durchschnittliche Wahrscheinlichkeit einer Überschreitung möglicherweise durchaus korrekt abgebildet haben.

Wir sehen also sehr deutlich, dass es nicht ausreicht, nur auf die durchschnittliche Überschreitungswahrscheinlichkeit zu prüfen - es ist unerlässlich, auch die Unabhängigkeitseigenschaft zu testen. Im Folgenden werden wir zu diesem Zweck den Markow-Test auf Unabhängigkeit, zu welchem bereits Campbell (2006/2007) eine sehr gute Einführung liefert, präsentieren.

Wir werden uns im Folgenden darauf fokussieren, auf Autokorrelation erster Ordnung zu prüfen - das heißt, nur auf mögliche Abhängigkeiten von dem direkt vorhergegangenen Wert $t - 1$. Nicht prüfen werden wir auf Abhängigkeit von den Werten zu den Zeitpunkten $t - 2$, $t - 3$ etc.

Zufallsvariable

Von Interesse ist hier eine Folge an Realisationen i_1, \dots, i_T der Zufallsvariablen I_1, \dots, I_T .

Hypothesentest

Wir testen auf Autokorrelation erster Ordnung:

$$\begin{aligned} H_0 : I_t \text{ und } I_{t+1} \text{ sind unabhängig} \\ \text{gegen } H_1 : I_t \text{ und } I_{t+1} \text{ sind nicht unabhängig} \end{aligned}$$

Die Nullhypothese ist also, dass die Folge serielle Unabhängigkeit erster Ordnung aufweist. Die Alternativhypothese besagt, dass diese Annahme nicht zutrifft.

Testfunktion

Wir testen dies mittels Likelihood-Quotienten-Test,

$$\begin{aligned} \Lambda(\hat{\Pi}_{ML}, \hat{\Pi}_U) &= -2 \ln \left(\frac{L(\hat{\Pi}_U)}{L(\hat{\Pi}_{ML})} \right) \\ \Lambda(\hat{\Pi}_{ML}, \hat{\Pi}_U) &\sim \chi_1^2 \end{aligned}$$

Eine Herleitung der Transitionsmatrizen erfolgt im Verlauf dieses Abschnitts. Dieser Test erlaubt uns, recht elegant auf Unabhängigkeit zu testen. Bei Bedarf kann er problemlos auch abgeändert werden, um Tests auf serielle Autokorrelation größer der ersten Ordnung zu erlauben.

Ablehnungsbereich

Durch die Verteilung der Teststatistik können wir den Ablehnungsbereich wieder als Quantil der χ_1^2 -Verteilung ausdrücken:

$$A = \left((\chi_1^2)^{-1} (1 - \alpha), +\infty \right)$$

Transitionsmatrizen

Transitionsmatrizen geben systematisch Wahrscheinlichkeiten an, dass ein bestimmter heutiger Zustand (Überschreitung des Value at Risk / Keine Überschreitung) in einen bestimmten Zustand zum nächsten Zeitpunkt (Überschreitung / keine Überschreitung) übergeht. Transitionsmatrizen haben dabei die Form

$$\Pi = \begin{pmatrix} \pi_{00} & \pi_{01} \\ \pi_{10} & \pi_{11} \end{pmatrix} = \begin{pmatrix} 1 - \pi_{01} & \pi_{01} \\ 1 - \pi_{11} & \pi_{11} \end{pmatrix}$$

Die Bedeutungen der Symbole π_{ab} sind dabei:

Symbol	VaR heute	VaR morgen
π_{00}	Keine Überschreitung	Keine Überschreitung
π_{10}	Überschreitung	Keine Überschreitung
π_{01}	Keine Überschreitung	Überschreitung
π_{11}	Überschreitung	Überschreitung

Likelihood-Funktion eines Markow-Prozesses

Wir betrachten nun unsere gesamte Stichprobe vom Umfang T und definieren \mathfrak{Z}_{ab} , $a, b \in \{0, 1\}$ als Anzahl aller Beobachtungen, bei denen die Überschreitungsfunktion zu einem Zeitpunkt den Wert a annimmt und im nächsten Zeitpunkt den Wert b .

Das heißt also, dass \mathfrak{Z}_{01} beispielsweise die Anzahl aller Beobachtungen ist, bei denen auf eine Nicht-Überschreitung des Value at Risk eine Überschreitung folgt. Formal definieren wir dies wie folgt:

$$\mathfrak{Z}_{ab} = |\{(I_t, I_{t+1}) | I_t = a, I_{t+1} = b, t = 1, \dots, T - 1\}| \text{ für } a, b \in \{0, 1\}$$

Die zugehörige Likelihood-Funktion des Markow-Prozesses ist

$$\begin{aligned} L(\Pi) &= \pi_{00}^{\mathfrak{T}_{00}} \cdot \pi_{01}^{\mathfrak{T}_{01}} \cdot \pi_{10}^{\mathfrak{T}_{10}} \cdot \pi_{11}^{\mathfrak{T}_{11}} \\ &= (1 - \pi_{01})^{\mathfrak{T}_{00}} \cdot \pi_{01}^{\mathfrak{T}_{01}} \cdot (1 - \pi_{11})^{\mathfrak{T}_{10}} \cdot \pi_{11}^{\mathfrak{T}_{11}} \end{aligned}$$

ML-Schätzung der Transitionsmatrix

Wir können die Maximum Likelihood-Schätzer für $\hat{\pi}_{01}$ und $\hat{\pi}_{11}$ ermitteln, indem wir die Likelihood-Funktion ableiten, gleich 0 setzen und entsprechend auflösen. Dazu bestimmen wir zur Vereinfachung zuerst die logarithmierte Likelihood-Funktion:

$$\begin{aligned} \ln L(\Pi) &= \mathfrak{T}_{00} \cdot \ln(1 - \pi_{01}) + \mathfrak{T}_{01} \cdot \ln(\pi_{01}) \\ &\quad + \mathfrak{T}_{10} \cdot \ln(1 - \pi_{11}) + \mathfrak{T}_{11} \cdot \ln(\pi_{11}) \end{aligned}$$

Ableiten und Auflösen nach π_{01} liefert

$$\begin{aligned} \frac{\delta \ln L(\Pi)}{\delta \pi_{01}} &= -\frac{\mathfrak{T}_{00}}{1 - \pi_{01}} + \frac{\mathfrak{T}_{01}}{\pi_{01}} \stackrel{!}{=} 0 \\ \frac{\mathfrak{T}_{01}}{\pi_{01}} &= \frac{\mathfrak{T}_{00}}{1 - \pi_{01}} \\ \frac{\pi_{01}}{\mathfrak{T}_{01}} &= \frac{1 - \pi_{01}}{\mathfrak{T}_{00}} = \frac{1}{\mathfrak{T}_{00}} - \frac{\pi_{01}}{\mathfrak{T}_{00}} \\ \frac{\pi_{01}}{\mathfrak{T}_{01}} + \frac{\pi_{01}}{\mathfrak{T}_{00}} &= \frac{1 - \pi_{01}}{\mathfrak{T}_{00}} = \frac{1}{\mathfrak{T}_{00}} \\ \pi_{01} \cdot \left(\frac{1}{\mathfrak{T}_{01}} + \frac{1}{\mathfrak{T}_{00}} \right) &= \pi_{01} \cdot \left(\frac{\mathfrak{T}_{00} + \mathfrak{T}_{01}}{\mathfrak{T}_{00} \cdot \mathfrak{T}_{01}} \right) = \frac{1}{\mathfrak{T}_{00}} \\ \pi_{01} &= \frac{\mathfrak{T}_{00} \cdot \mathfrak{T}_{01}}{\mathfrak{T}_{00} \cdot (\mathfrak{T}_{00} + \mathfrak{T}_{01})} \\ \hat{\pi}_{01} &= \frac{\mathfrak{T}_{01}}{\mathfrak{T}_{00} + \mathfrak{T}_{01}} \end{aligned}$$

Analog können wir π_{11} bestimmen:

$$\hat{\pi}_{11} = \frac{\mathfrak{T}_{11}}{\mathfrak{T}_{10} + \mathfrak{T}_{11}}$$

Die Schätzer $\hat{\pi}_{00}$ und $\hat{\pi}_{10}$ können wir nun sehr einfach bestimmen. Da zum morgigen Zeitpunkt zwangsläufig irgendein Zustand - d.h. eine Überschreitung oder keine Überschreitung - eintreten muss, müssen folgende Gleichungen gelten:

$$\begin{aligned}\hat{\pi}_{01} + \hat{\pi}_{00} &= 1 \\ \hat{\pi}_{10} + \hat{\pi}_{11} &= 1\end{aligned}$$

Entsprechendes Umstellen liefert die ML-Schätzer

$$\begin{aligned}\hat{\pi}_{00} &= 1 - \hat{\pi}_{01} \\ \hat{\pi}_{10} &= 1 - \hat{\pi}_{11}\end{aligned}$$

Eine ML-Schätzung der Transitionsmatrix ergibt sich damit durch

$$\hat{\Pi}_{ML} = \begin{pmatrix} \frac{\mathfrak{I}_{00}}{\mathfrak{I}_{00} + \mathfrak{I}_{01}} & \frac{\mathfrak{I}_{01}}{\mathfrak{I}_{00} + \mathfrak{I}_{01}} \\ \frac{\mathfrak{I}_{10}}{\mathfrak{I}_{10} + \mathfrak{I}_{11}} & \frac{\mathfrak{I}_{11}}{\mathfrak{I}_{10} + \mathfrak{I}_{11}} \end{pmatrix}$$

Schätzung einer Transitionsmatrix unter Unabhängigkeit

Unter Unabhängigkeit hingegen müsste das Eintreten des morgigen Zustands vom heutigen unabhängig sein, das heißt es müsste gelten $\pi_{01} = \pi_{11} = \pi_U$. Da, wie bereits ausgeführt, zum zukünftigen Zeitpunkt einer der beiden möglichen Zustände eintreten muss ergeben sich zudem direkt die Werte für $\pi_{00} = \pi_{10} = 1 - \pi_U$. Unter Unabhängigkeit ist also die folgende Transitionsmatrix anzunehmen:

$$\hat{\Pi}_U = \begin{pmatrix} 1 - \hat{\pi}_U & \hat{\pi}_U \\ 1 - \hat{\pi}_U & \hat{\pi}_U \end{pmatrix}$$

Der ML-Schätzer für $\hat{\pi}_U$ kann dabei durch entsprechendes Ableiten, Nullsetzen und Auflösen der Likelihood-Funktion

$$L(\Pi_U) = (1 - \pi_U)^{\mathfrak{I}_{00} + \mathfrak{I}_{10}} \cdot \pi_U^{\mathfrak{I}_{01} + \mathfrak{I}_{11}}$$

ermittelt werden:

$$\begin{aligned}\ln L(\Pi_U) &= (\mathfrak{I}_{00} + \mathfrak{I}_{10}) \cdot \ln(1 - \pi_U) + (\mathfrak{I}_{01} + \mathfrak{I}_{11}) \cdot \ln(\pi_U) \\ \frac{\delta L(\Pi_U)}{\delta \pi_U} &= -\frac{\mathfrak{I}_{00} + \mathfrak{I}_{10}}{1 - \pi_U} + \frac{\mathfrak{I}_{01} + \mathfrak{I}_{11}}{\pi_U} \stackrel{!}{=} 0 \\ \hat{\pi}_U &= \frac{\mathfrak{I}_{01} + \mathfrak{I}_{11}}{\mathfrak{I}_{00} + \mathfrak{I}_{10} + \mathfrak{I}_{01} + \mathfrak{I}_{11}}\end{aligned}$$

Zusammenfassung des Testverfahrens

Hypothesentest

1. Testproblem

$$H_0 : \mathcal{P}(I_t(p) = 1 | I_{t-1}(p) = 1) = \mathcal{P}(I_t(p) = 1 | I_{t-1}(p) = 0)$$

gegen $H_1 : \mathcal{P}(I_t(p) = 1 | I_{t-1}(p) = 1) \neq \mathcal{P}(I_t(p) = 1 | I_{t-1}(p) = 0)$

2. Testfunktion

$$\Lambda(\hat{\Pi}_{\text{ML}}, \hat{\Pi}_{\text{U}}) = -2 \ln \left(\frac{L(\hat{\Pi}_{\text{U}})}{L(\hat{\Pi}_{\text{ML}})} \right)$$

$$L(\Pi_{\text{ML}}) = (1 - \pi_{01})^{\mathfrak{I}_{00}} \cdot \pi_{01}^{\mathfrak{I}_{01}} \cdot (1 - \pi_{11})^{\mathfrak{I}_{10}} \cdot \pi_{11}^{\mathfrak{I}_{11}}$$

$$\hat{\pi}_{01} = \frac{\mathfrak{I}_{01}}{\mathfrak{I}_{00} + \mathfrak{I}_{01}}$$

$$\hat{\pi}_{11} = \frac{\mathfrak{I}_{11}}{\mathfrak{I}_{10} + \mathfrak{I}_{11}}$$

$$\hat{\pi}_{00} = 1 - \hat{\pi}_{01}$$

$$\hat{\pi}_{10} = 1 - \hat{\pi}_{11}$$

$$L(\Pi_{\text{U}}) = (1 - \pi_{\text{U}})^{\mathfrak{I}_{00} + \mathfrak{I}_{10}} \cdot \pi_{\text{U}}^{\mathfrak{I}_{01} + \mathfrak{I}_{11}}$$

$$\hat{\pi}_{\text{U}} = \frac{\mathfrak{I}_{01} + \mathfrak{I}_{11}}{\mathfrak{I}_{00} + \mathfrak{I}_{10} + \mathfrak{I}_{01} + \mathfrak{I}_{11}}$$

$$\Lambda(\hat{\Pi}_{\text{ML}}, \hat{\Pi}_{\text{U}}) \sim \chi_1^2$$

3. Ablehnungsbereich

$$A = \left((\chi_1^2)^{-1}(1 - \alpha), +\infty \right)$$

4. Entscheidungsregel

H_0 ablehnen, wenn die Testfunktion im Ablehnungsbereich liegt. Ansonsten H_0 nicht ablehnen.

2.4. Validierung basierend auf der gesamten Verteilung

Wie bereits zu Anfang dieser Arbeit erläutert ist eine der wichtigsten Aufgaben des Marktrisikomanagements die korrekte Prognose zukünftiger Wahrscheinlichkeitsverteilungen. Wir werden nun Tests fokussieren, welche die korrekte Spezifikation der gesamten Risikoverteilung sicherstellen sollen. Im Folgenden ist also nicht nur von Interesse, ob ein ganz bestimmtes Quantil der Verteilung richtig geschätzt wird, sondern, ob die Verteilung in ihrer Gesamtheit richtig geschätzt wird. Sollte dies der Fall sein, so impliziert dies immer auch eine korrekte Schätzung von Value at Risk sowie Expected Shortfall.

Als Kritikpunkt solcher Tests wird oft angeführt, dass es nicht nur wesentlich schwerer ist, eine komplette Verteilung korrekt zu bestimmen, sondern auch, dass dies zu einer Verwässerung des Risikobegriffs führt, da ein Risikomanagement in aller Regel eigentlich nur an den Randbereichen einer Verteilung interessiert ist. Campbell (2006/2007) und Berkowitz (2001) diskutieren diese Kritikpunkte exemplarisch.

Dennoch erfreuen sich diese Tests in der Praxis großer Beliebtheit. Durch die Einbeziehung aller verfügbaren Informationen kann die statistische Macht teils deutlich erhöht werden. Da gerade dies bei überschreitungsbasierten Tests ein erhebliches Problem darstellt, sind die Tests der gesamten Verteilung eine praktikable Alternative.

2.4.1. Rosenblatt-Transformation zur Gleichverteilung

Wir wollen im Folgenden überprüfen, ob unsere prognostizierten Verteilungen mit den real eingetretenen Werten vereinbar sind. Problematisch dabei ist, dass die prognostizierten Verteilungen abhängig von den jeweiligen Anforderungen und Annahmen unterschiedlichste Formen und Parameter annehmen können.

Idealerweise sollten die Testverfahren dabei auf jede beliebige Verteilungsform anwendbar sein. Bei diesem Problem hilft uns die sogenannte "Rosenblatt-Transformation", welche in diesem Kontext auch von Campbell (2006/2007) angewandt wird.

Nehmen wir an, die Portfoliowertänderung $\Delta V(t|\Omega_t)$ hat eine bestimmte, uns bekannte Verteilung mit Verteilungsfunktion $F_{\Delta V(t|\Omega_t)}(\cdot)$.

Wir definieren nun eine Zufallsvariable Z_t als Verknüpfung der Zufallsvariablen ΔV_t und ihrer angenommenen Verteilungsfunktion. Die Rosenblatt-Transformation besagt nun, dass die Zufallsvariable Z_t gleichverteilt auf dem Intervall $(0, 1)$ ist:

$$Z_t = F_{\Delta V(t|\Omega_t)}(\Delta V_t) \text{ für } t = 1, \dots, T$$
$$Z_t \stackrel{i.i.d.}{\sim} \mathbb{U}(0, 1)$$

Crnkovic & Drachman (1996) liefern eine noch intuitiver verständliche Zusammenfassung dieser Transformation: Die Quantile dieser Dichtefunktionen sind gleichverteilt.

Dies ist natürlich nur wahr, wenn die beobachteten Werte Δv_t auch tatsächlich den angenommenen Verteilungsfunktionen $F_{\Delta V(t|\Omega_t)}$ entstammen.

Mit anderen Worten: Wenn wir die Wahrscheinlichkeitsverteilungen korrekt prognostizieren, dann werden die Rosenblatt-Transformationen der realisierten Werte gleichverteilt sein. Wenn wir hingegen falsch prognostizieren, dann werden unsere transformierten Werte nicht gleichverteilt sein. Je schlechter unsere Prognose die Realität abbildet, desto stärker entfernt sich die Verteilung der Z -Werte von einer Gleichverteilung.

Wie Berkowitz (2001) herausstellt gilt dies auch, wenn sich die Verteilungen im Laufe der Zeit verändern. Wir können die Transformation also auch problemlos für Portfolios mit sich ändernden Beständen und unterschiedlichen Verteilungsparametern, d.h. beispielsweise selbst bei Heteroskedastizität, verwenden.

Damit haben wir die Aufgabe, die Korrektheit unserer prognostizierten Verteilungen zu bestimmen, wesentlich vereinfacht. Statt für jede erdenkliche Verteilungsform separate Testverfahren zu verwenden können wir nun eine einfache Transformation durchführen und überprüfen, ob die resultierenden Werte gleichverteilt sind. Können wir dies bejahen, so sind unsere Prognosen korrekt, andernfalls ist unser Modell fehlerhaft.

Es müssen zwei wesentliche Eigenschaften geprüft werden:

1. Gilt $Z \sim \mathbb{U}(0, 1)$, d.h. sind die Rosenblatt-Transformationen der realisierten Werte tatsächlich gleichverteilt auf dem Intervall $(0, 1)$?
2. Sind die Elemente der Folge Z_1, \dots, Z_T unabhängig voneinander?

Wir werden versuchen, die erste Frage über das CD-Verfahren sowie über einen Kolmogorov-Smirnov-Test zu überprüfen.

Auf die Unabhängigkeitseigenschaft werden wir erst im nächsten Abschnitt im Rahmen einer Normalverteilung weiter eingehen. Sollte explizit die Unabhängigkeitsprüfung der Rosenblatt-transformierten Daten gewünscht sein, so könnte dies,

wie von Crnkovic & Drachman (1996) vorgeschlagen, beispielsweise durch einen hier nicht näher erläuterten Brock-Dechert-Scheinkman-Test durchgeführt werden.

2.4.2. CD-Verfahren: Kuiper-Statistik mit Worry Function

Wir wollen die Hypothese

$$H_0 : Z \sim \mathbb{U}(0, 1) \text{ gegen } H_1 : Z \not\sim \mathbb{U}(0, 1)$$

überprüfen. Wohlgermerkt fordern wir hier nicht explizit die *i.i.d.*-Eigenschaft, Z muss also nicht unabhängig verteilt sein.

Crnkovic & Drachman (1996) schlagen für die Überprüfung dieser Hypothese ein Verfahren vor, welches auf der Kuiper-Statistik zur Überprüfung auf eine bestimmte Verteilungsform basiert.

Grundsätzlich wird hier die empirische Verteilungsfunktion von Z ermittelt und mit der Verteilungsfunktion einer perfekten $\mathbb{U}(0, 1)$ -Verteilung verglichen. Als Kennzahl für den Unterschied zwischen diesen beiden Funktionen dient dann die Summe aus größter Abweichung nach unten und größter Abweichung nach oben:

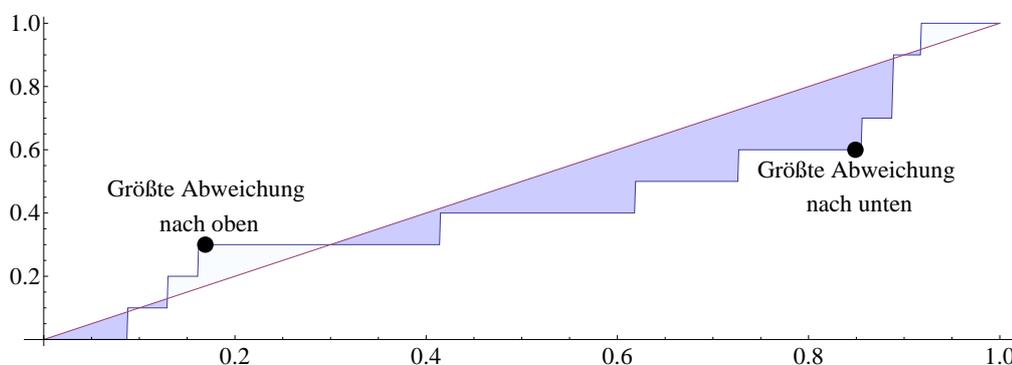


Abbildung 2.19.: Zusammensetzung der Kuiper-Statistik

Für dieses Verfahren müssen wir zuerst aus unseren Rosenblatt-transformierten Werten Z eine empirische Verteilungsfunktion berechnen:

$$\hat{F}(x) = \frac{n(x)}{T}, \quad x \in (0, 1)$$

$$n(x) = \max_{t \in \tau} \{t | z_{[t]} \leq x\}$$

$n(x)$ berechnet also die Anzahl aller Werte z_t , die kleiner als x sind. $F(x)$ ist demzufolge die Funktion, welche den Anteil an Werten $z_t \leq x$ relativ zur gesamten Stichprobe angibt, also die Verteilungsfunktion von Z . Wir definieren nun die Kuiper-Statistik für eine Gleichverteilung:

$$Q = \max_{0 \leq x \leq 1} \{\widehat{F}(x) - x\} + \max_{0 \leq x \leq 1} \{x - \widehat{F}(x)\}$$

Die Statistik Q ist die Summe der größten Differenzen nach unten und nach oben zwischen tatsächlicher Verteilungsfunktion und empirischer Verteilungsfunktion. Der Wert sollte möglichst klein sein.

In Abhängigkeit der Stichprobengröße T lassen sich nun kritische Werte für Q per Monte Carlo Simulation approximieren, für die die Nullhypothese abgelehnt werden muss.

Dazu wird ein Vektor \mathbf{x} mit T Realisationen gleichverteilter Zufallsvariablen erzeugt und eine Teststatistik $Q(\mathbf{x})$ berechnet,

$$Q(\mathbf{x}) = \max_{t=1, \dots, T} \left(\widehat{F}(x_t) - F(x_t), \widehat{F}(x_{t+1}) - F(x_t) \right) \\ + \max_{t=1, \dots, T} \left(F(x_t) - \widehat{F}(x_t), F(x_t) - \widehat{F}(x_{t+1}) \right)$$

Dies wird möglichst oft wiederholt. Alle dadurch berechneten Werte der Teststatistik werden schließlich im Vektor \mathbf{s} zusammengefasst. Der Ablehnungsbereich dieses Testverfahrens zum Konfidenzniveau $1 - \alpha$ ergibt sich nun aus den entsprechenden Quantilen der Verteilung der Teststatistik,

$$A = (s_{[gT(T(1-\alpha))+1]}, +\infty)$$

Für gängige Stichprobengrößen sind kritische Werte bereits berechnet und verfügbar.

Doch erscheint es sinnvoll, die Abweichungen von der Gleichverteilung zu gewichten. Zwar haben wir in diesem Abschnitt durchaus Interesse daran, die korrekte Spezifikation der gesamten Verteilung zu beurteilen, dennoch interessieren uns im Bereich des Risikomanagements natürlich meist besonders die Randbereiche. Eine Abweichung an den Randbereichen sollte also mit größerem Gewicht in die Statistik eingehen als eine Abweichung in den Zentralbereichen.

Genau dies lässt sich durch eine Gewichtung der tatsächlichen Verteilungsfunktion sowie der empirischen Verteilungsfunktion mit einer sogenannten "Worry Function"

bewerkstelligen. In Anlehnung an Crnkovic & Drachman (1996) entscheiden wir uns für die Funktion

$$W(x) = -\frac{\ln(x \cdot (1-x))}{2}$$

Diese Gewichtungsfunktion verläuft wie folgt auf dem Intervall (0, 1):

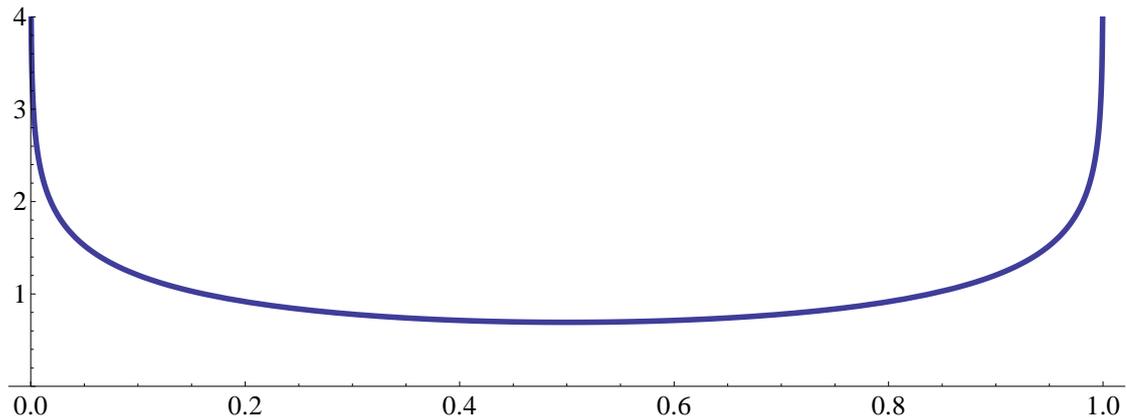


Abbildung 2.20.: Worry Function

Wir wenden die Worry Function nun auf die theoretische sowie auf die empirische Verteilungsfunktion an und erhalten dadurch eine gewichtete Statistik Q_w , mittels welcher den Randbereichen eine größere Bedeutung zugemessen wird:

$$Q_w = \max_{0 \leq x \leq 1} \{W(x) \hat{F}(x) - W(x) x\} + \max_{0 \leq x \leq 1} \{W(x) x - W(x) \hat{F}(x)\}$$

Auch hier müssen in Abhängigkeit von Stichprobengröße und Ausgestaltung der Worry Function kritische Werte wieder per Monte Carlo Simulation gefunden werden, indem bestimmte Quantile des Vektors \mathbf{s}_w , welcher die Teststatistiken Q_w der simulierten Stichproben beinhaltet, betrachtet werden. Die folgende Tabelle stellt kritische Werte in Abhängigkeit von Stichprobengröße T und α -Wert basierend auf 10.000.000 unabhängigen Simulationen pro Zeitraum T dar:

$T \backslash \alpha$	0.1%	1%	5%	10%	15%
1	3.808	2.652	1.857	1.523	1.334
10	0.849	0.681	0.549	0.488	0.451
21	0.541	0.443	0.367	0.332	0.311
125	0.2	0.17	0.147	0.136	0.129
250	0.139	0.119	0.104	0.096	0.092
500	0.097	0.084	0.073	0.068	0.065

Tabelle 2.4.: Kritische Werte der Teststatistik Q_w

Das Problem des CD-Verfahrens ist die relativ große Datenmenge, die es benötigt. In der Regel wird von zwischen 500 bis 1000 Datenwerten ausgegangen, die für signifikante Ergebnisse mit in die Berechnung einfließen sollten. Dies entspricht in der Praxis einem Zeitraum von zwei bis vier Jahren, um genügend Daten akkumulieren zu können. Kupiec (1998) geht zudem davon aus, dass weitere Modifikationen des Tests bei einer asymmetrischen Verteilung nötig wären.

Das CD-Verfahren ist also eine interessante Möglichkeit, die Verteilung mit einem Fokus auf bestimmte Teilbereiche zu überprüfen, scheitert in der Praxis oft aber an der großen Datenmenge, die es - wie die meisten nicht-parametrischen Testverfahren bedingen - benötigt. Zudem sind die Ergebnisse nur aussagekräftig, wenn keine Autokorrelation in den Daten vorliegt - dies muss gegebenenfalls separat überprüft werden.

Zusammenfassung des Testverfahrens

Hypothesentest
<p>1. Testproblem</p> $H_0 : Z \sim \mathbb{U}(0, 1) \text{ gegen } H_1 : Z \not\sim \mathbb{U}(0, 1)$
<p>2. Testfunktion</p> $Q_w(\mathbf{x}) = \max_{t=1, \dots, T} \left(W(x_t) \widehat{F}(x_t) - W(x_t) F(x_t), W(x_t) \widehat{F}(x_{t+1}) - W(x_t) F(x_t) \right) + \max_{t=1, \dots, T} \left(W(x_t) F(x_t) - W(x_t) \widehat{F}(x_t), W(x_t) F(x_t) - W(x_t) \widehat{F}(x_{t+1}) \right)$
<p>3. Ablehnungsbereich</p> $A = (s_{w gT(T(1-\alpha))+1}, +\infty)$
<p>4. Entscheidungsregel H_0 ablehnen, wenn die Testfunktion im Ablehnungsbereich liegt. Ansonsten H_0 nicht ablehnen.</p>

2.4.3. Kolmogorov-Smirnov-Test mit Worry Function

Wir kommen nun zu einer weiteren, in der Literatur bislang wenig diskutierten Methode, welche dem CD-Verfahren ähnelt. Ziel wird es erneut sein, zu überprüfen, ob unsere empirisch ermittelte Verteilungsfunktion mit der Verteilungsfunktion einer $\mathbb{U}(0, 1)$ -Verteilungsfunktion übereinstimmen könnte. Wir werden dazu allerdings eine andere Teststatistik als beim CD-Verfahren verwenden.

Wir definieren also erneut die Hypothesen sowie die empirische Verteilungsfunktion durch

$$H_0 : Z \sim \mathbb{U}(0, 1) \text{ vs. } H_1 : Z \not\sim \mathbb{U}(0, 1)$$

$$\widehat{F}(x) = \frac{n(x)}{T}, \quad x \in (0, 1)$$

$$n(x) = \max_{t \in \tau} \{t | z_{[t]} \leq x\}$$

Der Unterschied zum CD-Verfahren liegt nun in der Art und Weise, auf die wir die Nullhypothese prüfen. Dazu führen wir hier die Kolmogorov-Smirnov-Teststatistik des Anpassungstests ein:

$$D = \sup_{0 \leq x \leq 1} \{|\hat{F}(x) - x|\}$$

Im Gegensatz zur Kuiper-Teststatistik, die eine Summe der größten Abweichung nach unten und nach oben hin war, ist die Kolmogorov-Smirnov-Statistik also lediglich die insgesamt größte absolut gesehene Abweichung:

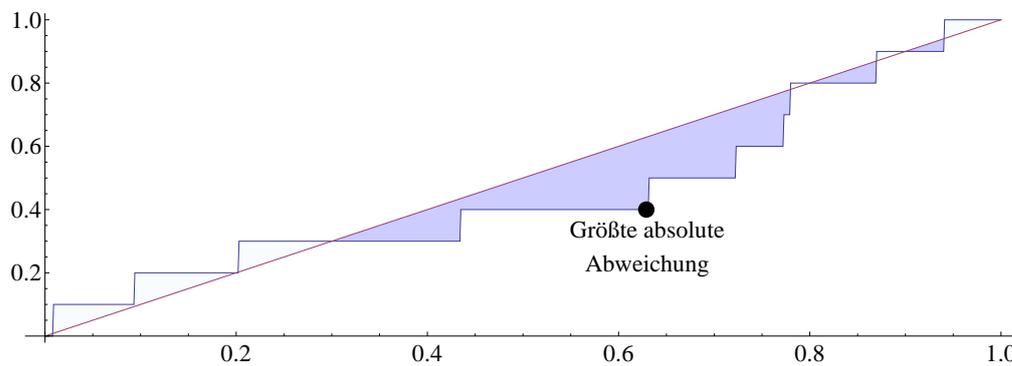


Abbildung 2.21.: Veranschaulichung einer Simulation einer exemplarischen Teststatistik für $T = 10$

Nun gibt es wohlbekannte Testverfahren zum Überprüfen der Nullhypothese anhand der Teststatistik D . Doch analog zum vorherigen Abschnitt wollen wir auch hier explizit Abweichungen an den Rändern der Verteilung stärker gewichten. Wir nutzen dafür erneut die bereits erläuterte Worry Function:

$$W(x) = -\frac{\ln(x \cdot (1-x))}{2}$$

Durch Anwendung auf empirische und theoretische Verteilungsfunktion erhalten wir eine neue Teststatistik,

$$D_w = \sup_{0 \leq x \leq 1} \{|W(x) \hat{F}(x) - W(x)x|\}$$

Durch Anwendung der Worry Function können wir nicht mehr etwa auf die Kolmogorovverteilung zur Ermittlung kritischer Werte der Teststatistik zurückgreifen, sondern müssen selbige per Monte Carlo Simulation finden. Wie im vorherigen Abschnitt ausführlicher dargestellt ergeben sich diese aus den Quantilen der

simulierten Verteilung der Teststatistik s_{dw} . In einer Simulation mit 10.000.000 Wiederholungen pro Zeitraum ergeben sich folgende kritische Werte:

$T \backslash \alpha$	0.1%	1%	5%	10%	15%
1	3.798	2.637	1.81	1.447	1.233
10	0.656	0.514	0.406	0.357	0.326
21	0.406	0.325	0.263	0.234	0.216
125	0.145	0.12	0.101	0.092	0.087
250	0.1	0.084	0.071	0.065	0.061
500	0.07	0.059	0.05	0.046	0.043

Tabelle 2.5.: Kritische Werte der Teststatistik D_w

Zusammenfassung des Testverfahrens

Hypothesentest	
1. Testproblem	$H_0 : Z \sim \mathbb{U}(0, 1)$ gegen $H_1 : Z \not\sim \mathbb{U}(0, 1)$
2. Testfunktion	$D_w(\mathbf{x}) = \max_{t=1, \dots, T} \left(W(x_t) \hat{F}(x_t) - W(x_t) F(x_t) , W(x_t) \hat{F}(x_{t+1}) - W(x_t) F(x_t) \right)$
3. Ablehnungsbereich	$A = (s_{dw[gT(T(1-\alpha))+1]}, +\infty)$
4. Entscheidungsregel	H_0 ablehnen, wenn die Testfunktion im Ablehnungsbereich liegt. Ansonsten H_0 nicht ablehnen.

2.4.4. Wahrscheinlichkeitsintegral zur Transformation in eine Normalverteilung

In den vorherigen Abschnitten haben wir die Rosenblatt-Transformation auf unsere Daten angewandt und schließlich versucht zu berechnen, ob die daraus resultierende empirische Verteilungsfunktion der einer Gleichverteilung entsprechen könnte. Wir haben dabei aber festgestellt, dass diese Tests Schwächen in kleinen Stichproben haben.

Wir werden nun einen Schritt weiter gehen und nach der Rosenblatt-Transformation noch eine weitere Transformation per Wahrscheinlichkeitsintegral anwenden, um unsere Daten in eine idealerweise standardnormalverteilte Form zu bringen. Dies bietet für uns den Vorteil, dass wir parametrische Tests zum Prüfen der Verteilungsannahme anwenden können, welche mitunter bessere Eigenschaften in kleineren Stichproben haben.

Wir wenden auf eine Serie von Zufallsvariablen, welche unabhängig gleichverteilt auf $(0, 1)$ sind, die inverse Verteilungsfunktion der Standardnormalverteilung, $\Phi^{-1}(\cdot)$, an. Das daraus resultierende Ergebnis ist eine Serie unabhängig standardnormalverteilter Zufallsvariablen (Christoffersen 2003):

$$Q_t = \Phi^{-1}(Z_t)$$
$$Q_t \stackrel{i.i.d.}{\sim} \mathbb{N}(0, 1)$$

Wenden wir auf unsere Wertänderungen ΔV_t nun also erst die Rosenblatt-Transformation und anschließend die hier erläuterte Transformation an, so sollten wir eine Serie unabhängig standardnormalverteilter Werte bekommen:

$$Q_t = \Phi^{-1}(F_{\Delta V(t|\Omega_t)}(\Delta V_t))$$
$$Q \stackrel{i.i.d.}{\sim} \mathbb{N}(0, 1)$$

Dies gilt natürlich nur, wenn die Rosenblatt-transformierten Daten tatsächlich unabhängig gleichverteilt auf $(0, 1)$ sind. Es gilt also nur, wenn sich unsere prognostizierten Wahrscheinlichkeitsverteilungen als korrekt herausstellen. Je weniger korrekt unsere Prognosen waren, desto stärker entfernt sich die Serie Q_t von unabhängig standardnormalverteilten Werten.

Dies lässt sich visualisieren, indem man zuerst ein korrekt spezifiziertes Modell betrachtet. In diesem Fall lassen wir tatsächliche Daten nach einer Standardnor-

malverteilung generieren und nutzen zur Prognose ebenfalls eine Standardnormalverteilung:¹³

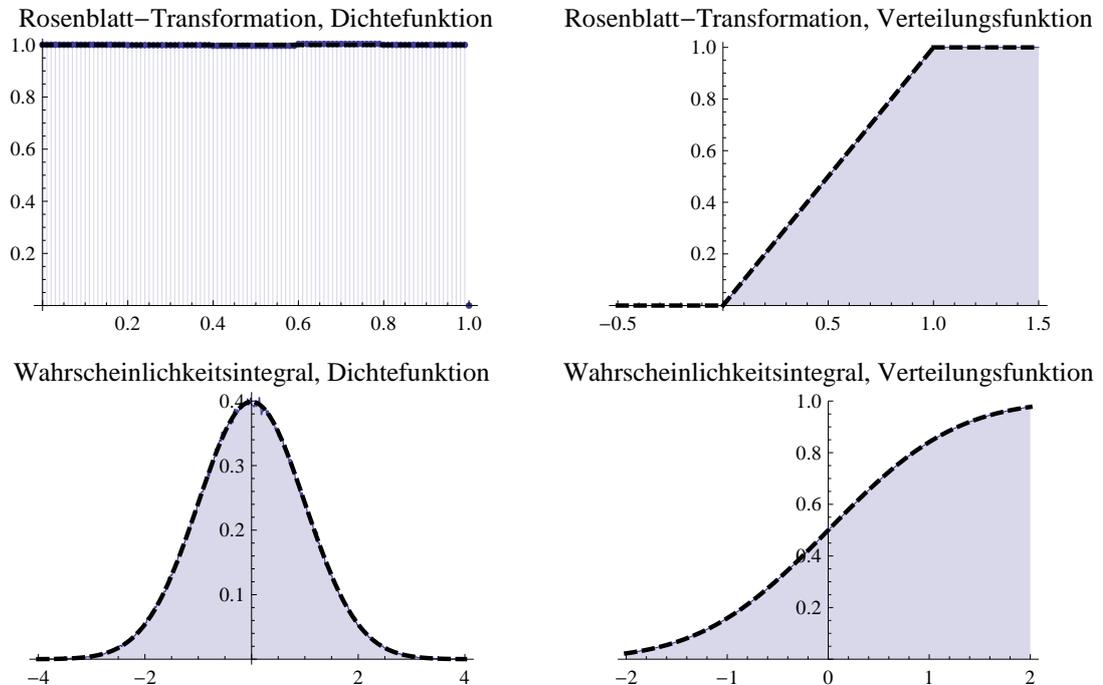


Abbildung 2.22.: Transformation von korrekt geschätzten Daten

In den Schaubildern stellen die schwarz gepunkteten Linien die Idealwerte dar, also z.B. Dichte- und Verteilungsfunktion einer $\mathbb{U}(0,1)$ -Verteilung im Fall der Rosenblatt-Transformation und eine $\mathbb{N}(0,1)$ -Verteilung für die Transformation per Wahrscheinlichkeitsintegral. Die blau gezeichneten Werte stellen die empirischen Dichte- und Verteilungsfunktionen dar. Man sieht, dass die empirischen Funktionen praktisch identisch mit den theoretischen Idealfunktionen sind.

Doch wie sehen die transformierten Daten bei einem falsch spezifizierten Modell aus? Wir illustrieren dies anhand eines aus Risikosicht sehr relevanten Beispiels: Wir unterschätzen das Risiko nun bewusst, indem wir von standardnormalverteilten Daten ausgehen, die tatsächlichen Daten aber von einer Verteilung mit höherer Varianz, $\mathbb{N}(0,2)$, generieren lassen:

¹³Basiert auf jeweils 1,000,000 Simulationen

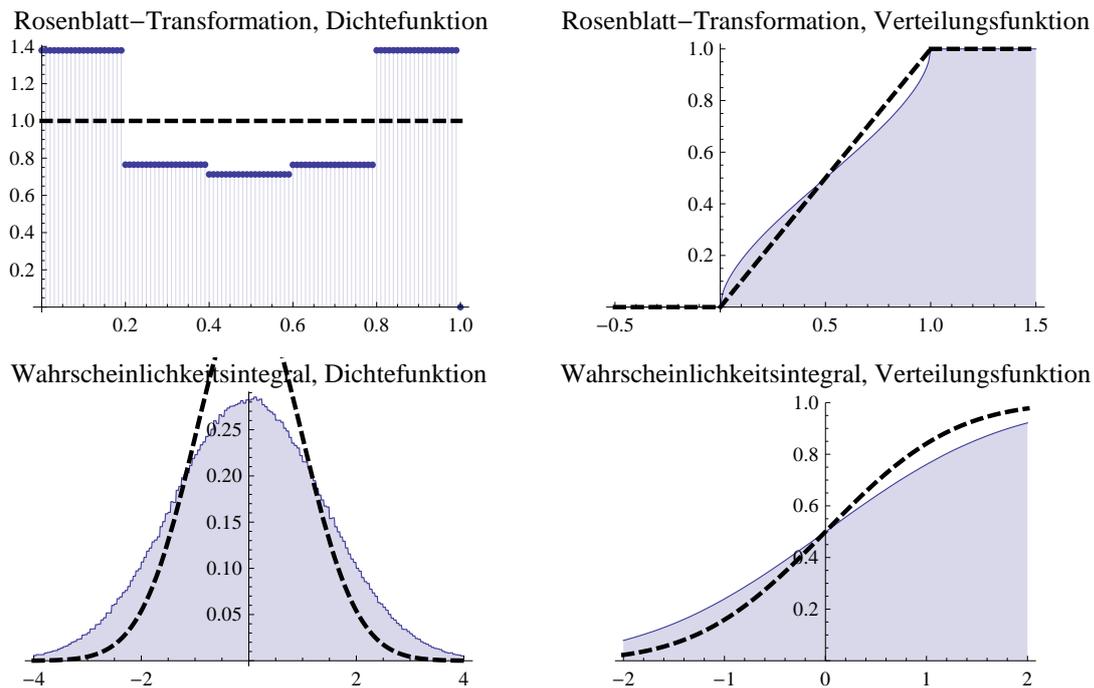


Abbildung 2.23.: Transformation von Daten, welche das Risiko unterschätzen

Bei einer Risikounterschätzung weist sowohl die Transformation zur Gleichverteilung als auch die nachfolgende Transformation zur Standardnormalverteilung zu große Wahrscheinlichkeit in den Randbereichen auf. Dies ist intuitiv nachvollziehbar: Der datengenerierende Prozess hat eine höhere Varianz als im Modell angenommen. In unserem Modell, $N(0, 1)$ würden wir beispielsweise erwarten, in rund 2,28% aller Fälle einen Wert ≤ -2 zu erhalten. Der datengenerierende Prozess, $N(0, 2)$, wird jedoch in ca. 7,86% aller Fälle einen Wert ≤ -2 erzeugen. Wir erhalten also mit einer unerwartet hohen Wahrscheinlichkeit Werte, die in den Randbereichen unserer Modellverteilung liegen.

Eine Risikoüberschätzung würde sich ebenfalls in empirischer Dichte- und Verteilungsfunktion widerspiegeln. Hierbei würden die empirischen Funktionen eine zu große Fläche im Zentralbereich aufweisen.

Es wird also deutlich, dass die empirische Verteilung der per Wahrscheinlichkeitsintegral transformierten Daten eine Standardnormalverteilung approximiert, sofern das Modell korrekt ist. Ungenauigkeiten im Modell spiegeln sich sowohl nach der Rosenblatt-Transformation als auch in gleicher Weise nach der Wahrscheinlichkeitsintegral-Transformation in Abweichungen von der jeweils erwarteten Dichte- und Verteilungsfunktion wider.

Diese Erkenntnisse helfen uns nun dabei, die Angemessenheit unseres Modells zu überprüfen. Nach Transformation per Wahrscheinlichkeitsintegral sollten unsere Daten unabhängig und identisch standardnormalverteilt sein. Dies werden wir anhand von drei Kriterien überprüfen:

1. Ist Q_t normalverteilt?
2. Hat die Verteilung von Q_t einen Mittelwert von 0 und eine Varianz von 1?
3. Sind die einzelnen Werte aus Q_t unabhängig voneinander?

Wenn all diese Fragen mit Ja beantwortet werden können, dann ist das Prognosemodell richtig spezifiziert. In den kommenden Abschnitten werden daher Tests erläutert, um genau diese drei Punkte zu beantworten.

2.4.5. Jarque-Bera-Test auf Normalverteilung

Der Jarque-Bera-Test prüft, wie von Jost, Spree, Hapke, Mehmke & Huck (2011) beschrieben, ob vorliegende Daten einer Normalverteilung entstammen.

Dazu vergleicht er Schiefe und Kurtosis, d.h. Wölbung, der empirischen Verteilung mit den entsprechenden Werten einer Normalverteilung.

Selbige hat eine Schiefe von 0, ist also symmetrisch, und weist eine Kurtosis von 3 auf. Wenn die empirische Verteilung zu stark von diesen Werten abweicht, so ist davon auszugehen, dass der datengenerierende Prozess keiner Normalverteilung folgt.

Schiefe S und Kurtosis K entsprechen einer Funktion $\mu_n(X)/\sigma^n(X)$, also das n -te zentrale Moment normiert auf die n -te Potenz der Standardabweichung, für $n = 3$ (Schiefe) bzw. $n = 4$ (Kurtosis):

$$S = \frac{\frac{1}{T} \sum_{t=1}^T (q_t - \hat{q})^3}{\left(\frac{1}{T} \sum_{t=1}^T (q_t - \hat{q})^2 \right)^{\frac{3}{2}}}, \quad K = \frac{\frac{1}{T} \sum_{t=1}^T (q_t - \hat{q})^4}{\left(\frac{1}{T} \sum_{t=1}^T (q_t - \hat{q})^2 \right)^2}$$

mit $\hat{q} = \frac{1}{T} \sum_{t=1}^T q_t$

Wir testen nun die Hypothese

$$H_0 : Q \sim \mathbb{N}(\cdot, \cdot) \text{ gegen } H_1 : Q \not\sim \mathbb{N}(\cdot, \cdot)$$

mittels der Jarque-Bera-Teststatistik, welche unter der Nullhypothese χ_2^2 -verteilt ist:

$$JB = \frac{T}{6} \left(S^2 + \frac{(K-3)^2}{4} \right)$$
$$JB \sim \chi_2^2$$

Ein Problem ist freilich, dass wir die Nullhypothese durch diesen Test zwar gegebenenfalls signifikant ablehnen können, wir können sie aber nicht gesichert akzeptieren. Wir können uns also niemals statistisch wirklich gesichert dafür entscheiden, dass die Daten tatsächlich normalverteilt sind. Eine Analyse der statistischen Macht zeigt sich zudem als kompliziert, da die Verteilung unter der Alternativhypothese gänzlich unbekannt ist.

Es ist jedoch anzunehmen, dass der Jarque-Bera-Test für eine ausreichend große Stichprobe in Verbindung mit den anderen beiden Tests zur Überprüfung auf Vorliegen einer Standardnormalverteilung genügend Aussagekraft besitzt, um eine Nicht-Zurückweisung der Nullhypothese als Annahme interpretieren zu können.

Zusammenfassung des Testverfahrens

Hypothesentest	
1. Testproblem	$H_0 : Q \sim \mathbb{N}(\cdot, \cdot)$ gegen $H_1 : Q \not\sim \mathbb{N}(\cdot, \cdot)$
2. Testfunktion	$JB = \frac{T}{6} \left(S^2 + \frac{(K-3)^2}{4} \right)$ $JB \sim \chi_2^2$
3. Ablehnungsbereich	$A = \left((\chi_2^2)^{-1}(1 - \alpha), +\infty \right)$
4. Entscheidungsregel	H_0 ablehnen, wenn die Testfunktion im Ablehnungsbereich liegt. Ansonsten H_0 nicht ablehnen.

2.4.6. Regression auf Mittelwert, Varianz und Autokorrelation

Berkowitz (2001) schlägt vor, Mittelwert, Varianz und eine mögliche Autokorrelation erster Ordnung in den Daten durch eine Regression zu überprüfen. Wir orientieren uns hier an seinem Konzept und führen einen AR(1)-Prozess ein:

$$Q_t = \mu + \rho Q_{t-1} + \epsilon_t$$
$$\epsilon_t \sim \mathbb{N}(0, \sigma^2)$$

Dieses Modell geht also davon aus, dass unsere Daten normalverteilt mit einer bestimmten Varianz sind und einem Erwartungswert, der sich aus einem konstanten Ausdruck und einem Anteil des vorherigen realisierten Wertes zusammensetzt. Wir wollen am Ende auf Vorliegen einer Standardnormalverteilung testen. Dafür sollte die Varianz σ^2 genau 1 sein, der konstante Erwartungswert μ sollte 0 sein. Zudem sollten die Daten unabhängig voneinander sein, das heißt ein Wert sollte

unbeeinflusst von der jeweils vorherigen Realisation sein. Dies wäre für $\rho = 0$ der Fall.

Die empirischen Werte dieser Parameter können wir durch einfache Regression erhalten. Doch müssen wir schließlich testen, ob sie gegebenenfalls auch signifikant von den unter Standardnormalverteilung zu erwartenden Werten sind - wir müssen also einen Hypothesentest durchführen. Dazu werden wir nun die der obigen Modellgleichung zugehörigen Likelihood-Funktion herleiten, über die wir letztlich einen Likelihood-Quotienten-Test der Hypothese durchführen können.

Mit Parametervektor

$$\theta = (\mu, \sigma^2, \rho)$$

gilt für Q_1 :

$$E(Q_1) = \frac{\mu}{1 - \rho}$$

$$Var(Q_1) = \frac{\sigma^2}{1 - \rho^2}$$

Annahmegemäß ist ϵ_t normalverteilt. Da dies die einzige Zufallsgröße ist, die in Y_1 eingeht, ist auch Y_1 normalverteilt mit Dichtefunktion

$$f_{Q_1}(q_1; \theta) = \frac{1}{\sqrt{\frac{2\pi\sigma^2}{1 - \rho^2}}} \exp\left(-\frac{\left(q_1 - \frac{\mu}{1 - \rho}\right)^2}{\frac{2\sigma^2}{1 - \rho^2}}\right)$$

Auch die bedingte Wahrscheinlichkeit aller nachfolgenden Werte (Q_2, Q_3, \dots, Q_T) sind, gegeben einer Realisation des jeweiligen vorherigen Wertes, normalverteilt mit

$$E(Q_t | Q_{t-1}) = \mu + \rho q_{t-1}$$

$$Var(Q_t | Q_{t-1}) = \sigma^2$$

$$(Q_t | Q_{t-1}) \sim \mathbb{N}(\mu + \rho q_{t-1}, \sigma^2)$$

für $2 \leq t \leq T$. Die bedingte Dichtefunktion für Q_t ist also

$$f_{Q_t}(q_t | q_{t-1}; \theta) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(q_t - \mu - \rho q_{t-1})^2}{2\sigma^2}\right)$$

Die Dichtefunktion eines beliebigen Q_t ergibt sich nun also durch:

$$f_{Q_t}(q_t, q_{t-1}, q_{t-2}, \dots, q_1; \theta) = f_{Q_1}(q_1; \theta) \cdot \prod_{i=2}^t f_{Q_i|Q_{i-1}}(q_i|q_{i-1}; \theta)$$

Die Likelihood der gesamten Daten ergibt sich für $t = T$:

$$\begin{aligned} L(q_1, \dots, q_T; \theta) &= f_{Q_1}(q_1; \theta) \cdot \prod_{i=2}^T f_{Q_i|Q_{i-1}}(q_i|q_{i-1}; \theta) \\ &= \frac{1}{\sqrt{\frac{2\pi\sigma^2}{1-\rho^2}}} \exp\left(-\frac{\left(q_1 - \frac{\mu}{1-\rho}\right)^2}{\frac{2\sigma^2}{1-\rho^2}}\right) \cdot \prod_{i=2}^T \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(q_i - \mu - \rho q_{i-1})^2}{2\sigma^2}\right) \end{aligned}$$

Durch logarithmieren ergibt sich die Log-Likelihood-Funktion:

$$\begin{aligned} \ln L(q_1, \dots, q_T; \theta) &= \ln \left(\frac{1}{\sqrt{\frac{2\pi\sigma^2}{1-\rho^2}}} \exp\left(-\frac{\left(q_1 - \frac{\mu}{1-\rho}\right)^2}{\frac{2\sigma^2}{1-\rho^2}}\right) \right) + \sum_{i=2}^T \ln \left(\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(q_i - \mu - \rho q_{i-1})^2}{2\sigma^2}\right) \right) \\ &= \ln \left(\frac{1}{\sqrt{\frac{2\pi\sigma^2}{1-\rho^2}}} \right) + \left(-\frac{\left(q_1 - \frac{\mu}{1-\rho}\right)^2}{\frac{2\sigma^2}{1-\rho^2}} \right) + \sum_{i=2}^T \ln \left(\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(q_i - \mu - \rho q_{i-1})^2}{2\sigma^2}\right) \right) \\ &= \ln(1) - \ln \left((2\pi)^{\frac{1}{2}} \left(\frac{\sigma^2}{1-\rho^2} \right)^{\frac{1}{2}} \right) + \left(-\frac{\left(q_1 - \frac{\mu}{1-\rho}\right)^2}{\frac{2\sigma^2}{1-\rho^2}} \right) \\ &\quad + \sum_{i=2}^T \ln \left((2\pi\sigma^2)^{-\frac{1}{2}} - \frac{(q_i - \mu - \rho q_{i-1})^2}{2\sigma^2} \right) \end{aligned}$$

also

$$\ln L(q_1, \dots, q_T; \theta) = -\frac{1}{2} \ln(2\pi) - \frac{1}{2} \ln\left(\frac{\sigma^2}{1-\rho^2}\right) + \left(-\frac{\left(q_1 - \frac{\mu}{1-\rho}\right)^2}{\frac{2\sigma^2}{1-\rho^2}}\right) \\ - \frac{T-1}{2} \ln(2\pi) - \frac{T-1}{2} \ln(\sigma^2) - \sum_{i=2}^T \frac{(q_i - \mu - \rho q_{i-1})^2}{2\sigma^2}$$

Uns bietet sich nun die Möglichkeit, verschiedene Hypothesen mittels Likelihood-Quotienten-Test zu überprüfen. Wir können dabei bestimmte Aussagen über den Erwartungswert, die Varianz und die Autokorrelation erster Ordnung und jede beliebige Kombination dieser drei Charakteristika treffen.

Da wir mit diesem Test Aussagen über eine mögliche unabhängige Standardnormalverteilung unserer Daten treffen wollen bietet sich ein gemeinsamer Test an, welcher auf einen Mittelwert von 0, eine Varianz von 1 und eine Autokorrelation erster Ordnung von 0 prüft:

$$H_0 : \theta = (\mu, \sigma^2, \rho) = (0, 1, 0) \text{ vs. } H_1 : \theta \neq (0, 1, 0)$$

Als Teststatistik nutzen wir einen Likelihood-Quotienten-Test,

$$\Lambda(q, \hat{\mu}, \hat{\sigma}^2, \hat{\rho}) = -2(\ln L(\mathbf{q}; (0, 1, 0)) - \ln L(\mathbf{q}; (\hat{\mu}, \hat{\sigma}^2, \hat{\rho}))) \\ \Lambda(\mathbf{q}, \hat{\mu}, \hat{\sigma}^2, \hat{\rho}) \sim \chi_3^2$$

Zum bestimmen der empirischen Parameter $\hat{\theta} = (\hat{\mu}, \hat{\sigma}^2, \hat{\rho})$ könnte nun die Log-Likelihood-Funktion optimiert werden - in unserem Fall bietet es sich aber an, eine einfache Regression nach der Methode kleinster Quadrate durchzuführen, welche hier auf die selben Ergebnisse wie die exakte Maximum Likelihood-Parameterschätzung kommt.

Zusammenfassung des Testverfahrens

Hypothesentest	
1. Testproblem	$H_0 : (\mu, \sigma^2, \rho) = (0, 1, 0) \text{ gegen } H_1 : (\mu, \sigma^2, \rho) \neq (0, 1, 0)$
2. Testfunktion	$\Lambda(\mathbf{q}, \hat{\mu}, \hat{\sigma}^2, \hat{\rho}) = -2 (\ln L(\mathbf{q}; (0, 1, 0)) - \ln L(\mathbf{q}; (\hat{\mu}, \hat{\sigma}^2, \hat{\rho})))$ $\ln L(\mathbf{q}; \theta) = -\frac{1}{2} \ln(2\pi) - \frac{1}{2} \ln\left(\frac{\sigma^2}{1-\rho^2}\right) + \left(-\frac{\left(q_1 - \frac{\mu}{1-\rho}\right)^2}{\frac{2\sigma^2}{1-\rho^2}} \right)$ $- \frac{T-1}{2} \ln(2\pi) - \frac{T-1}{2} \ln(\sigma^2) - \sum_{i=2}^T \frac{(q_i - \mu - \rho q_{i-1})^2}{2\sigma^2}$ $\Lambda(\mathbf{q}, \hat{\mu}, \hat{\sigma}^2, \hat{\rho}) \sim \chi_3^2$ <p>$\hat{\mu}, \hat{\sigma}^2, \hat{\rho}$ können per linearer Regression gefunden werden.</p>
3. Ablehnungsbereich	$A = \left((\chi_3^2)^{-1}(1-\alpha), +\infty \right)$
4. Entscheidungsregel	<p>H_0 ablehnen, wenn die Testfunktion im Ablehnungsbereich liegt. Ansonsten H_0 nicht ablehnen.</p>

2.4.7. Fazit zu ganzheitlichen Validierungsmodellen

In diesem Abschnitt sind wir zuerst auf die Rosenblatt-Transformation eingegangen, welche, bei korrekter Spezifikation des Risikomodells, die Realisationen unserer Risikoverteilung in eine Gleichverteilung transformieren kann. Wir testeten anschließend auf die Form der Gleichverteilung.

Um die Gleichverteilungsannahme zu überprüfen mussten wir auf nicht-parametrische Testverfahren zurückgreifen. Ein Nachteil dieser Testverfahren ist jedoch ihr schlechtes Abschneiden in Stichproben kleinen Umfangs.

Um auf parametrische Tests zurückgreifen zu können haben wir die Daten anschließend nochmals transformiert, um - zumindest ebenfalls bei korrekt spezifiziertem Modell - standardnormalverteilte Daten zu erhalten. Wir testeten erneut auf die Form der eigentlichen Verteilung sowie auf Unabhängigkeit.

Sowohl bei nicht-parametrischen als auch bei parametrischen Tests auf eine Verteilung stellte sich uns jedoch das Problem, dass das Zutreffen der Verteilungsannahme in der Nullhypothese angegeben werden musste. Damit kann die Verteilungsannahme, also die Annahme auf Gleich- oder Standardnormalverteilung, zwar signifikant abgelehnt werden, jedoch nicht statistisch gesichert angenommen werden. Eine genauere Analyse des β -Fehlers bzw. der statistischen Macht zeigt der Verfahren sich problematisch, da dazu konkrete Verteilungen für die Alternativhypothese angenommen werden müssten - dies wäre jedoch reine Spekulation und damit nicht aussagekräftig.

Technisch gesehen können wir also auch durch die vorgestellten Tests keine absolut gesicherte Aussage über die Verteilungsform treffen. Vor allem in Zusammenhang mit ergänzenden Tests erscheint es jedoch gleichermaßen legitim wie unumgänglich, ein nicht-Ablehnen der Nullhypothese hier mit einer Annahme derselbigen gleichzusetzen. Mögliche praktische Auswirkungen werden sich gegebenenfalls in der Simulation in Abschnitt 3 zeigen.

2.5. Validierung basierend auf einem Teilbereich der Verteilung

Im vorherigen Abschnitt versuchten wir abzuschätzen, ob die Prognose der kompletten Risikoverteilung mit der Realität vereinbar ist. Dies hatten den Vorteil, dass wir alle uns zur Verfügung stehenden Informationen über unsere Risikovariablen nutzen konnten. Problematisch ist jedoch, dass für das Risikomanagement in aller Regel nicht die korrekte Spezifikation der gesamten Verteilung, sondern lediglich die korrekte Spezifikation der Randbereiche von Interesse ist.

Wir werden uns daher nun der Frage widmen, wie man überprüfen kann, ob der Randbereich beginnend ab $-\infty$ bis $-VaR$ korrekt spezifiziert ist. Bei Zutreffen impliziert dies direkt eine korrekte Spezifikation des Expected Shortfall. Die nun vorgestellten Verfahren eignen sich also hervorragend zur Validierung desselben, da Abweichungen im relevanten Randbereich erkannt werden können, ohne dabei von Abweichungen in den für das Risikomanagement weniger interessanten Verteilungsbereichen beeinflusst zu werden.

2.5.1. Gestutzter Likelihood-Quotienten-Test

Berkowitz (2001) schlägt vor, eine neue Zufallsvariable Q^* einzuführen, welche den Wert unserer Risikovariablen ΔV unter der Bedingung annimmt, dass das p -Quantil unterschritten wurde.

Man kann sich dies in etwa so vorstellen, dass nur noch der Randbereich einer Verteilung von Bedeutung ist, exemplarisch durch den blau unterlegten Bereich folgender Abbildung dargestellt:

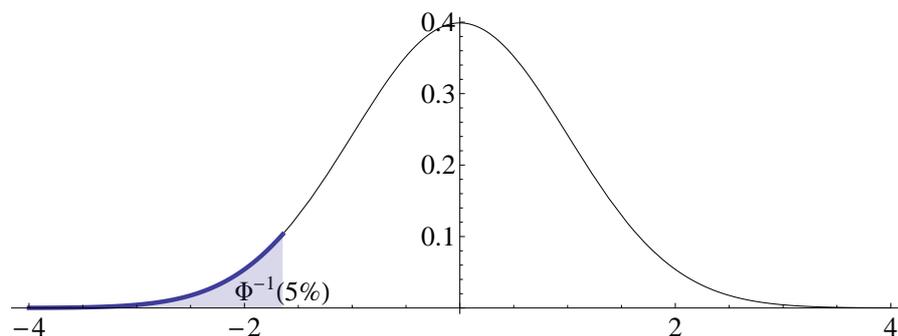


Abbildung 2.24.: Stutzung einer Standardnormalverteilung bei $\Phi^{-1}(5\%)$

Die Abbildung ist insofern nicht ganz korrekt, als dass die Dichtefunktion der gestutzten Verteilung bezogen auf ihre Fläche noch auf 1 normiert werden muss. Dies ändert jedoch nur die Größenverhältnisse, nicht die eigentliche Form.

Mathematisch definieren wir eine neue Zufallsvariable Q^* wie folgt:

$$\mathcal{P}(Q^* \leq q^*) = \mathcal{P}(Q \leq q^* | Q \leq F_Q^{-1}(p^*))$$

Wir weichen von Berkowitz (2001) ab, indem wir die neue Zufallsvariable nur für tatsächliche Quantilsüberschreitungen definieren.¹⁴ Wir gehen also von einer gestutzten statt von einer zensierten Verteilung aus. Die Dichtefunktion von Q^* entspricht unter der Nullhypothese von der Form her einer Standardnormalverteilung, durch die Einschränkung jedoch nur noch von $-\infty$ bis F_Q^{-1} . Die Wertemenge von Q^* ist also $[-\infty, F_Q^{-1}(p)]$.

Die Dichtefunktion von Q^* ist

$$f_{Q^*}(x) = \frac{\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)}{\Phi\left(\frac{F_Q^{-1}(p) - \mu}{\sigma}\right)} = \frac{\frac{1}{\sigma} \phi\left(\frac{x-\mu}{\sigma}\right)}{\Phi\left(\frac{F_Q^{-1}(p) - \mu}{\sigma}\right)}$$

Zur Parameterschätzung und für Hypothesentests können wir nun erneut eine Likelihood-Funktion konstruieren. Sei T^* die die Anzahl der Q , die kleiner oder gleich dem p^* -Quantil der Standardnormalverteilung sind. Dies entspricht im Übrigen genau unserer Zufallsvariable der summierten Anzahl aller Value at Risk-Überschreitungen I , also $T^* = I$. Mit $\theta = (\mu, \sigma)$ gilt:

$$\begin{aligned} L(q_1^*, \dots, q_{T^*}^*; \theta) &= \prod_{i=1}^{T^*} f_{Q_i^*}(q_i^*; \theta) \\ \ln L(q_1^*, \dots, q_{T^*}^*; \theta) &= \sum_{i=1}^{T^*} \left[\frac{1}{\sigma} \phi\left(\frac{q_i^* - \mu}{\sigma}\right) \right] - \sum_{t=1}^{T^*} \ln \left(\Phi\left(\frac{F_Q^{-1}(p) - \mu}{\sigma}\right) \right) \\ &= \sum_{i=1}^{T^*} \ln \left[(2\pi\sigma^2)^{-\frac{1}{2}} \exp\left(-\frac{(q_i^* - \mu)^2}{2\sigma^2}\right) \right] - \sum_{t=1}^{T^*} \ln \left(\Phi\left(\frac{F_Q^{-1}(p) - \mu}{\sigma}\right) \right) \\ &= \sum_{i=1}^{T^*} \left[-\frac{1}{2} \ln(2\pi\sigma^2) - \frac{(q_i^* - \mu)^2}{2\sigma^2} \right] - T^* \ln \left(\Phi\left(\frac{F_Q^{-1}(p) - \mu}{\sigma}\right) \right) \end{aligned}$$

¹⁴Berkowitz definiert Q_t^* für jedes Q_t und berechnet dies als $Q_t^* = \min(Q_t, \Phi^{-1}(p))$

Wir testen nun die Hypothesen

$$H_0 : Q \sim \mathbb{N}_{\mathcal{T}}(0, 1, F_Q^{-1}(p)) \text{ gegen } H_1 : Q \not\sim \mathbb{N}_{\mathcal{T}}(0, 1, F_Q^{-1}(p))$$

$\mathbb{N}(\mu, \sigma^2, \mathcal{T})$ ist dabei die Normalverteilung $\mathbb{N}(\mu, \sigma^2)$ mit Stutzung ab der Stelle \mathcal{T} , d.h. Wertebereich ihrer Dichtefunktion von $[-\infty, \mathcal{T}]$. Wir testen hier also letztlich, ob der linke Rand Q dem einer Standardnormalverteilung entspricht. Werte, die im Zentrum oder am rechten Rand der Verteilung liegen werden bewusst ignoriert.

Den eigentlichen Test führen wir erneut durch einen Likelihood-Quotienten-Test durch:

$$\begin{aligned} \Lambda(\hat{\mu}, \hat{\sigma}^2) &= -2(\ln L(0, 1) - \ln L(\hat{\mu}, \hat{\sigma}^2)) \\ \Lambda(\hat{\mu}, \hat{\sigma}^2) &\sim \chi_2^2 \end{aligned}$$

Die Schätzer $\hat{\mu}$ und $\hat{\sigma}^2$ müssen numerisch durch Optimierung der logarithmierten Likelihood-Funktion gefunden werden.

Zusammenfassend prüfen wir hier also darauf, ob der linke Randbereich der Verteilung korrekt abgebildet wird. Wir haben dabei jedoch nicht auf eine mögliche Autokorrelation der Daten getestet. Dies muss bei Bedarf separat, beispielsweise anhand des ebenfalls im Rahmen dieser Arbeit vorgestellten Markow-Tests, durchgeführt werden.

Zusammenfassung des Testverfahrens

Hypothesentest

1. Testproblem

$$H_0 : Q \sim \mathbb{N}_T(0, 1, F_Q^{-1}(p)) \text{ gegen } H_1 : Q \not\sim \mathbb{N}_T(0, 1, F_Q^{-1}(p))$$

2. Testfunktion

$$\Lambda(\hat{\mu}, \hat{\sigma}^2) = -2(\ln L(0, 1) - \ln L(\hat{\mu}, \hat{\sigma}^2))$$

$$L(\mathbf{q}^*; \theta) = \sum_{i=1}^{T^*} \left[-\frac{1}{2} \ln(2\pi\sigma^2) - \frac{(q_i^* - \mu)^2}{2\sigma^2} \right] - T^* \ln \left(\Phi \left(\frac{F_Q^{-1}(p) - \mu}{\sigma} \right) \right)$$

$$\Lambda(\hat{\mu}, \hat{\sigma}^2) \sim \chi_2^2$$

3. Ablehnungsbereich

$$A = \left((\chi_2^2)^{-1}(1 - \alpha), +\infty \right)$$

4. Entscheidungsregel

H_0 ablehnen, wenn die Testfunktion im Ablehnungsbereich liegt. Ansonsten H_0 nicht ablehnen.

2.5.2. Gauss-Test des Expected Shortfall bei gestutzter Verteilung

Unter der Nullhypothese einer korrekten Modellspezifikation sind die per Wahrscheinlichkeitsintegral transformierten Daten approximativ standardnormalverteilt.

Im vorherigen Abschnitt definierten wir eine Zufallsvariable Q^* , welche lediglich die transformierten Werte kleiner gleich des p^* -Quantils der Standardnormalverteilt erfasst. Die Dichtefunktion dieser Variable wurde bereits angegeben.

Unter Zutreffen der Nullhypothese können wir diese jedoch noch weiter vereinfachen, da die sich in diesem Fall auf die Standardnormalverteilung bezieht:

$$f_{Q^*}^{H_0}(x) = \frac{\phi(x)}{p^*}$$

$$E(Q^*|H_0) = -\frac{\phi(\Phi^{-1}(p^*))}{p^*}$$

$$Var(Q^*|H_0) = 1 - \frac{\Phi^{-1}(p^*)\phi(\Phi^{-1}(p^*))}{p^*} - \left(\frac{-\phi(\Phi^{-1}(p^*))}{p^*}\right)^2$$

Da Q^* unter der Nullhypothese i.i.d. ist konvergiert der Erwartungswert der Folge $Q_1^*, \dots, Q_{T^*}^*$ gegen eine Normalverteilung mit Erwartungswert und Varianz wie oben angegeben.

Aufgrund des zentralen Grenzwertsatzes gilt nun:

$$\Upsilon_{T^*}(Q^*) = \sqrt{T^*} \frac{\overline{Q^*} - E(Q^*|H_0)}{\sqrt{Var(Q^*|H_0)}}$$

mit $\overline{Q^*} = \frac{1}{T^*} \sum_{i=1}^{T^*} Q_i^*$

$$\lim_{T^* \rightarrow \infty} \mathcal{P}(\Upsilon_{T^*} \leq z) = \Phi(z)$$

Das standardisierte arithmetische Mittel der Folge $Q_1^*, \dots, Q_{T^*}^*$ konvergiert also unter der Nullhypothese gegen eine Standardnormalverteilung.

Wir prüfen nun, ob das arithmetische Mittel unserer realisierten Daten, $\widehat{\overline{Q^*}}$, in Einklang zu bringen ist mit dem eigentlich anzunehmenden Erwartungswert der gestutzten Standardnormalverteilung, $E(Q^*|H_0)$:

$$H_0 : E(Q^*|H_0) = \widehat{\overline{Q^*}} \text{ gegen } H_1 : E(Q^*|H_0) \neq \widehat{\overline{Q^*}}$$

Der Ablehnungsbereich A mit Signifikanzniveau α wird berechnet durch:

$$A = \left(-\infty, \Phi^{-1}\left(\frac{\alpha}{2}\right)\right) \cup \left(\Phi^{-1}\left(1 - \frac{\alpha}{2}\right), +\infty\right)$$

Wir lehnen die Nullhypothese ab, wenn die Teststatistik in den Ablehnungsbereich fällt, also bei $\Upsilon_{T^*}(Q^*) \in A$.

Dieses eigentlich sehr elegante Verfahren wird in der Praxis leider oft von der geringen Mengen an verfügbaren Daten negativ beeinflusst. Die hier aus dem zentralen Grenzwertsatz gezogenen Schlussfolgerungen gelten nur für eine große Anzahl

verschiedener Realisation von Q^* . Bei etwa 250 Handelstagen würden, in einem korrekt spezifizierten Modell, jedoch im Durchschnitt weniger als 3 Realisationen von Q^* vorliegen. Im nächsten Abschnitt werden wir versuchen, dieses Problem zu verbessern.

Zusammenfassung des Testverfahrens

Hypothesentest
<p>1. Testproblem</p> $H_0 : E(Q^* H_0) = \widehat{Q}^* \text{ gegen } H_1 : E(Q^* H_0) \neq \widehat{Q}^*$
<p>2. Testfunktion</p> $\Upsilon(Q^*) = \sqrt{T^*} \frac{\overline{Q^*} - E(Q^* H_0)}{\sqrt{\text{Var}(Q^* H_0)}}$ $\overline{Q^*} = \frac{1}{T^*} \sum_{i=1}^{T^*} Q_i^*$ $E(Q^* H_0) = -\frac{\phi(\Phi^{-1}(p))}{p}$ $\text{Var}(Q^* H_0) = 1 - \frac{\Phi^{-1}(p)\phi(\Phi^{-1}(p))}{p} - \left(\frac{-\phi(\Phi^{-1}(p))}{p}\right)^2$ $\Upsilon(Q^*) \sim \mathbb{N}(0, 1)$
<p>3. Ablehnungsbereich</p> $A = \left(-\infty, \Phi^{-1}\left(\frac{\alpha}{2}\right)\right) \cup \left(\Phi^{-1}\left(1 - \frac{\alpha}{2}\right), +\infty\right)$
<p>4. Entscheidungsregel H_0 ablehnen, wenn die Testfunktion im Ablehnungsbereich liegt. Ansonsten H_0 nicht ablehnen.</p>

2.5.3. Sattelpunktapproximation mittels Lugannani Rice-Formel zum Überprüfen des Expected Shortfall

Die bislang vorgestellten Verfahren zur Überprüfung einer Expected Shortfall-Prognose litten vor allem an schlechten Eigenschaften in kleinen Stichproben - ein großes Problem, da die Anzahl an Überschreitungen in der Praxis tatsächlich sehr klein ist.

Die von Wong (2008) präsentierte, in diesem Abschnitt diskutierte Methode zielt darauf ab, auch in kleinen Stichproben bessere Aussagen erzielen zu können. Detailliertere Ausführungen zur Sattelpunktapproximation werden von DasGupta (2008) vorgestellt. Prinzipiell gesehen wird über die Inversion einer Fouriertransformation versucht, die Dichtefunktion zu rekonstruieren und diese zu integrieren, um die Verteilungsfunktion zu erhalten. Die Lugannani Rice-Formel, die auf diesem Prinzip basiert, gilt dabei als besonders präzise allgemein, aber auch im Speziellen in den Randbereichen.

Ausgangspunkt ist unsere bereits definierte Zufallsvariable Q^* , welche alle Werte, die unterhalb des $F_Q^{-1}(p^*)$ -Quantils liegen, beinhaltet. Als Expected Shortfall der Verteilung von Q definieren wir daher

$$ES = -\bar{Q}^* = -\frac{1}{T^*} \sum_{i=1}^{T^*} Q_i^*$$

Sei $q = \Phi^{-1}(p^*)$. Die momenterzeugende Funktion und ihre ersten drei Ableitungen bezogen auf die Zufallsvariable Q^* sind definiert durch

$$\begin{aligned} M(t) &= \frac{1}{p^*} \exp\left(\frac{t^2}{2}\right) \cdot \Phi(q-t) \\ M'(t) &= t \cdot M(t) - \exp(t \cdot q) \frac{\phi(q)}{p^*} \\ M''(t) &= t \cdot M'(t) + M(t) - \exp(q \cdot t) \cdot \frac{q \cdot \phi(q)}{p^*} \\ M'''(t) &= t \cdot M''(t) + 2M'(t) - \exp(q \cdot t) \cdot \frac{q^2 \phi(q)}{p^*} \end{aligned}$$

und kumulantengenerierender Funktion

$$\begin{aligned} K(t) &= \ln(M(t)) \\ K'(t) &= \frac{M'(t)}{M(t)} \\ K''(t) &= -\left(\frac{M'(t)}{M(t)}\right)^2 + \frac{M''(t)}{M(t)} \\ K'''(t) &= \frac{2M'(t)^3}{M(t)^3} - \frac{3M'(t)M''(t)}{M(t)^2} + \frac{M'''(t)}{M(t)} \end{aligned}$$

Wir bestimmen nun den Sattelpunkt ϖ so, dass er dem arithmetischen Mittel der realisierten Werte unserer Zufallsvariable \bar{q}^* entspricht, indem wir die folgende Gleichung numerisch nach ϖ lösen:

$$K'(\varpi) = \bar{q}^*$$

Anschließend berechnen wir zwei Werte η und ζ wie folgt:

$$\begin{aligned} \eta &= \varpi \sqrt{T^* K''(\varpi)} \\ \zeta &= \operatorname{sgn}(\varpi) \sqrt{2T^* (\varpi \bar{q}^* - K(\varpi))} \end{aligned}$$

$\operatorname{sgn}(x)$ ist dabei die Signumfunktion.

Mit diesen berechneten Werten können wir nun die Lugannani-Rice-Formel anwenden, welche die Wahrscheinlichkeit angibt, dass \bar{Q}^* unter der Nullhypothese kleiner gleich unserem empirischen Wert \bar{q}^* ist:

$$P(\bar{Q}^* \leq \bar{q}^*) = \begin{cases} \Phi(\zeta) - \phi(\zeta) \cdot \left(\frac{1}{\eta} - \frac{1}{\zeta} + \mathcal{O}(T^{*-3/2})\right) & \text{für } \bar{Q}^* < q \\ -\frac{1}{2} + \frac{K'''(0)}{6\sqrt{2\pi T^* (K''(0))^3}} + \mathcal{O}(T^{*-3/2}) & \text{für } \bar{Q}^* = \mu_{Q^*} \\ 1 & \text{für } \bar{Q}^* \geq q \end{cases}$$

\mathcal{O} ist hierbei ein Landau-Symbol, welches das Konvergenzverhalten des Ausdrucks beschreibt.

Wir testen die Hypothese, ob der theoretisch zu erwartende Wert unserer Zufallsvariable mit ihrem realisierten arithmetischen Mittel vereinbar ist:

$$H_0 : E(Q^*) \geq \bar{q}^* \text{ gegen } H_1 : E(Q^*) < \bar{q}^*$$

Bei Zutreffen der Nullhypothese liegt der Expected Shortfall der Stichprobe in der Nähe des theoretisch zu erwartenden Expected Shortfall eines korrekt spezifizierten Modells.

Zusammenfassung des Testverfahrens

Hypothesentest	
1. Testproblem	
$H_0 : E(Q^*) \geq \bar{q}^*$ gegen $H_1 : E(Q^*) < \bar{q}^*$	
2. Testfunktion	
$P(\bar{Q}^* \leq \bar{q}^*) =$	$\begin{cases} \Phi(\zeta) - \phi(\zeta) \cdot \left(\frac{1}{\eta} - \frac{1}{\zeta} + \mathcal{O}(T^{*-3/2}) \right) & \text{für } \bar{Q}^* < q \\ -\frac{1}{2} + \frac{K'''(0)}{6\sqrt{2\pi T^* (K''(0))^3}} + \mathcal{O}(T^{*-3/2}) & \text{für } \bar{Q}^* = \mu_{Q^*} \\ 1 & \text{für } \bar{Q}^* \geq q \end{cases}$
$P(\bar{Q}^* \leq \bar{q}^*) \sim \mathbb{U}(0, 1)$	
3. Ablehnungsbereich	
$A = (0, \alpha)$	
4. Entscheidungsregel	
H_0 ablehnen, wenn die Testfunktion im Ablehnungsbereich liegt. Ansonsten H_0 nicht ablehnen.	

3. Simulation

Im Folgenden werden wir die vorgestellte Validierungsmodelle in einer Simulation überprüfen. Von Interesse wird dabei zum einen die Ablehnungsrate korrekter Risikomodelle (α -Fehler) als auch die Ablehnungsrate falscher Risikomodelle (statistische Macht) sein. Die Hypothese H_0 wird dabei stets für die Angemessenheit eines Risikomodells stehen. Die Hypothese H_1 steht für einen Fehler im Risikomodell, der beispielsweise in einer Über- oder Unterschätzung des Risikos resultieren könnte.

Es werden für jede Iteration der Simulation von einem datengenerierenden Prozess 350 normalverteilte Werte erzeugt. Erwartungswert und Varianz folgen dabei bestimmten, später im Abschnitt genauer aufgeführten Vorschriften. Diese jeweils 350 Daten werden als "reale Renditen" betrachtet, die von einem Risikomodell abgebildet und eingeschätzt werden sollen.

Ein Risikomodell trifft schließlich Prognosen für diese realen Werte. Die ersten 100 Realisationen werden nur zur Parameteradjustierung genutzt, die darauf folgenden 250 Werte werden vom Risikomodell schließlich auf täglicher Basis prognostiziert. Für jeden Zeitpunkt $t = 1, \dots, 250$ existiert also eine Schätzung bezüglich Erwartungswert und Varianz der Portfoliorenditen, die vom Risikomodell am Vortag $t - 1$ berechnet wurde, und ein realer Wert, der im ersten Schritt vom datengenerierenden Prozess erzeugt wurde.

Basierend auf diesen Daten können nun alle Werte berechnet werden, die von den Validierungsverfahren benötigt werden - etwa realisierte Value at Risk-Überschreitungen, die Rosenblatt-Transformation der realen Werte etc. Das Konfidenzniveau des Value at Risk wird in der Simulation stets 95% sein.

Die Algorithmen sind dabei so konzipiert, dass zu jeder Teststatistik ein konkreter Signifikanzwert, auch bekannt als p-Wert, berechnet wird. Die Hypothese H_0 wird verworfen, wenn der Signifikanzwert geringer als das gewünschte α -Niveau des Hypothesentests ist.

Wir werden im Folgenden sechs verschiedene Ausgangssituationen betrachten:

1. Das Risikomodell ist korrekt spezifiziert. Wie hoch ist der α -Fehler der Validierungsmodelle?

2. Das Risikomodell ist falsch spezifiziert. Die Daten werden von einem GJR-GARCH-Prozess erzeugt, das Risikomodell berechnet Parameter auf einfacher, rollierender Basis.
3. Das Risikomodell ist falsch spezifiziert. Die Daten werden von einem GJR-GARCH-Prozess erzeugt, das Risikomodell berechnet Parameter mit exponentieller Glättung.
4. Das Risikomodell ist falsch spezifiziert. Die Daten werden von einem GJR-GARCH-Prozess erzeugt, das Risikomodell berechnet Parameter mittels GARCH-Prognose.
5. Das Risikomodell ist falsch spezifiziert und unterschätzt das reale Risiko konstant.
6. Das Risikomodell ist falsch spezifiziert. Zwar wird der Value at Risk korrekt abgebildet, jedoch wird der Expected Shortfall unterschätzt.

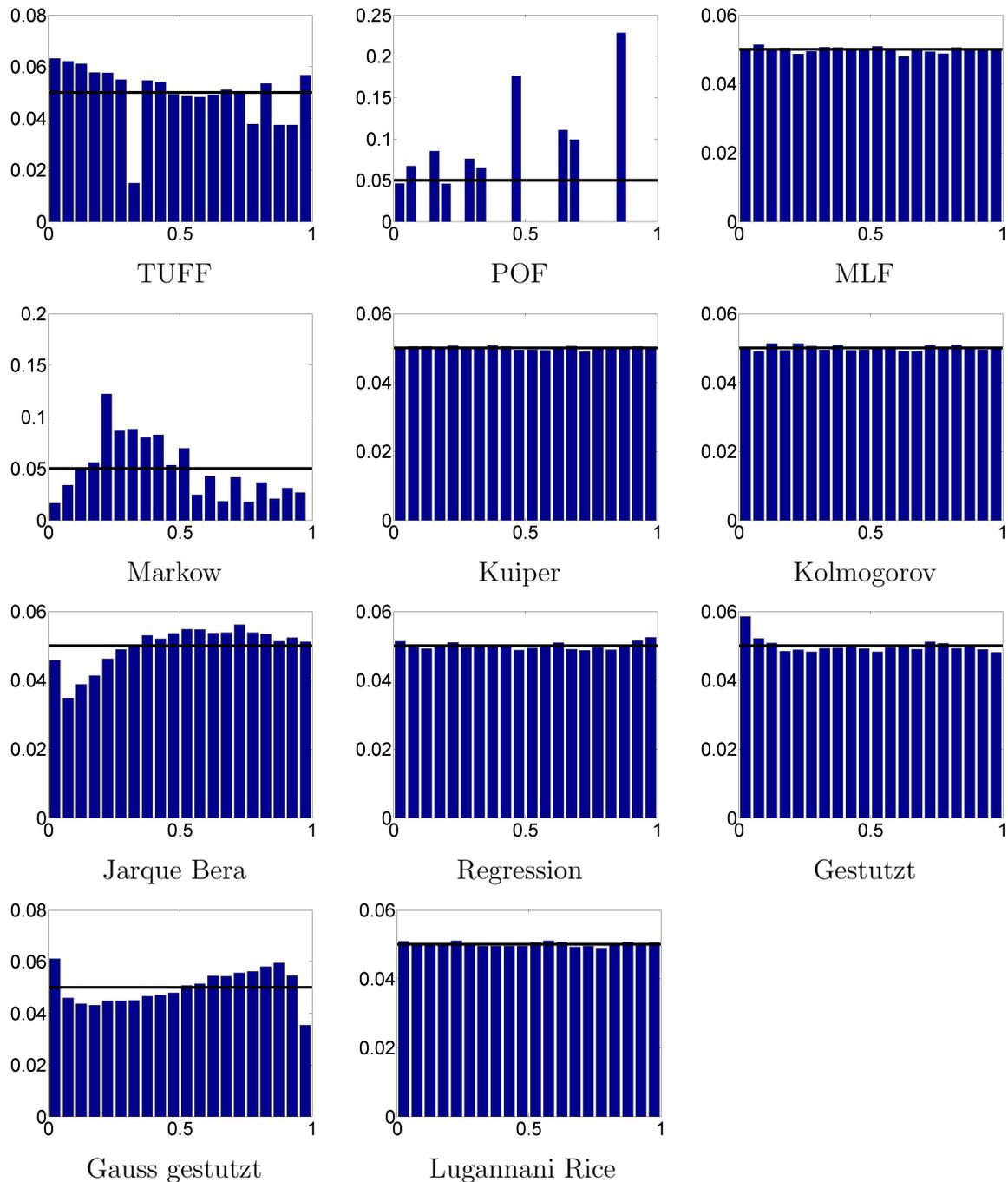
Bei den letzten drei der genannten Punkte steht dabei die Analyse der statistischen Macht im Vordergrund. Die genutzten Verfahren zur Erzeugung der Daten und Prognose der Parameter entsprechen den in Abschnitt 2.2.2 erläuterten Methoden. Für jede Ausgangssituation werden wir 100.000 Simulationsdurchgänge durchführen.

3.1. Korrekt spezifiziertes Risikomodell

Mit einem korrekt spezifizierten Risikomodell sollten die Signifikanzwerte einer Gleichverteilung folgen. Dies ergibt sich daraus, dass die Ablehnungsrate eines korrekten Modells genau dem α -Fehler des Hypothesentests entsprechen sollte - der Signifikanzwert sollte also beispielsweise genau in 5% aller Fälle kleiner gleich 5% sein.

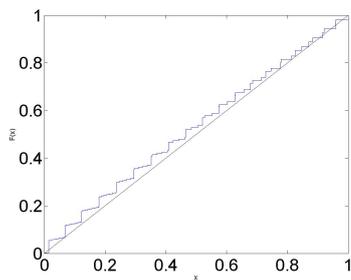
Diese Eigenschaft werden wir nun für alle Testverfahren prüfen. Der datengenerierende Prozess der "realen Renditen" folgt zu jedem Zeitpunkt einer Normalverteilung mit Erwartungswert 0 und Varianz 1. Das Risikoprognosemodell geht ebenfalls in jedem Zeitpunkt von einem Erwartungswert in Höhe von 0 und einer Varianz in Höhe von 1 aus und ist somit korrekt spezifiziert. Wir betrachten nun die Histogramme der Signifikanzwerte der verschiedenen Testverfahren. Auf der x-Achse ist jeweils der Wert der Signifikanzwerte in Abständen von 5% pro Balken abgetragen, auf der y-Achse werden deren relative Häufigkeiten dargestellt. Idealerweise sollten alle relativen Häufigkeiten genau 5% betragen, d.h. genau auf Höhe der jeweils eingezeichneten schwarzen Linie sein.

Validierung von Konzepten zur Messung des Marktrisikos -
insbesondere des Value at Risk und des Expected Shortfall

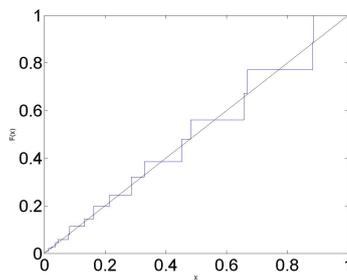


Die meisten Testverfahren erzeugen in der Tat weitestgehend gleichverteilte Signifikanzwerte. Lediglich POF- und Markow-Verfahren weichen deutlich von einer Gleichverteilung ab. Dies wird nochmals verdeutlicht, wenn man die empirischen Verteilungsfunktionen der Signifikanzwerte betrachtet:

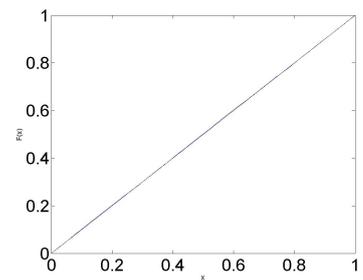
Validierung von Konzepten zur Messung des Marktrisikos -
insbesondere des Value at Risk und des Expected Shortfall



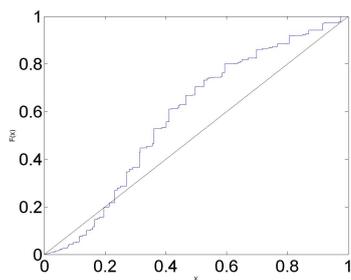
TUFF



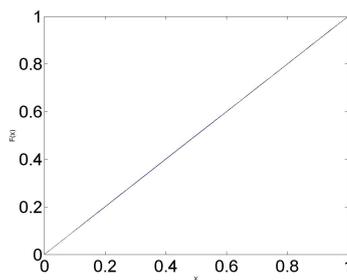
POF



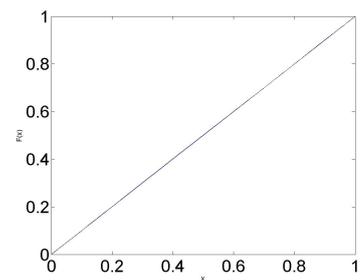
MLF



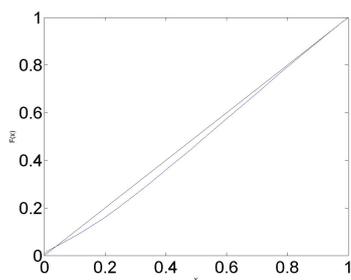
Markow



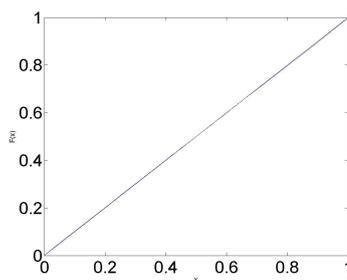
Kuiper



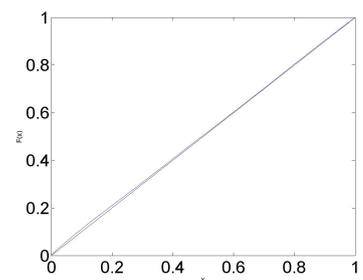
Kolmogorov



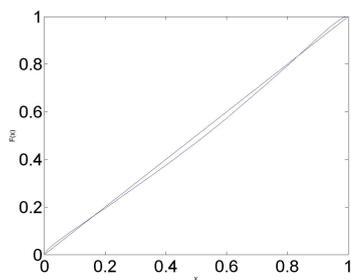
Jarque Bera



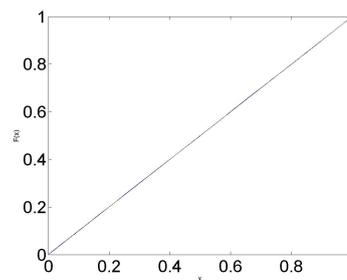
Regression



Gestutzt



Gauss gestutzt



Lugannani Rice

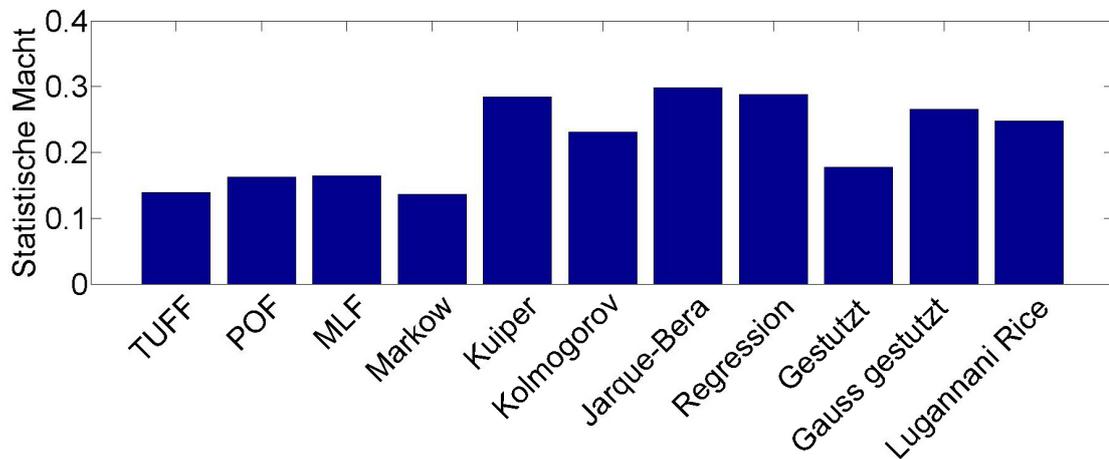
Sowohl beim Proportion of Failure-Verfahren als auch beim Markow-Test auf Unabhängigkeit zeigen sich bereits hier deutliche Schwächen, da die tatsächlichen α -Fehler deutlich von ihren spezifizierten Werten abweichen. Beim POF-Verfahren ist dies ein kleineres Problem, da die Verteilung zumindest für kleine Signifikan-

zwerte bis 5% noch recht präzise ist und somit die üblichen α -Niveaus von Hypothesentests weitestgehend realitätsnah sind. Die Verteilung des Markow-Tests weicht jedoch auch für kleine Werte bereits deutlich von einer Gleichverteilung ab. Der tatsächliche α -Fehler ist hier deutlich geringer als der erwartete α -Fehler. Dies bedeutet, dass der Markow-Test wesentlich seltener korrekte Modelle ablehnt als eigentlich zu erwarten wäre. Was zunächst sogar nach einem Vorteil klingt ist jedoch in der Praxis höchst problematisch, da in der Regel mit niedrigerem α -Fehler auch die statistische Macht eines Testverfahrens sinkt.

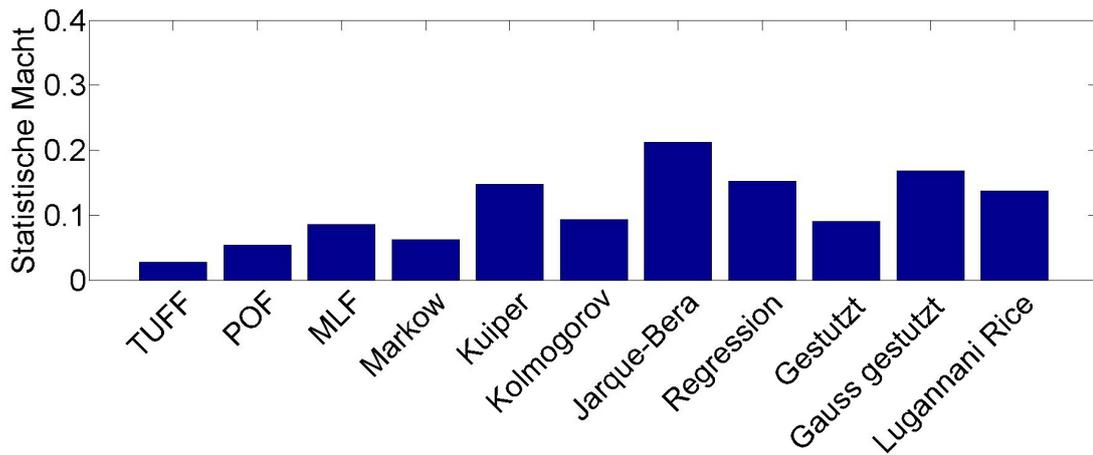
3.2. Risikomodell mit rollierenden Varianzen

Wir gehen nun von einem Risikomodell aus, das die Parameter für die zukünftige Periode prognostiziert, indem es den jeweiligen Durchschnitt der vergangenen 90 Tage berechnet. Im Vergleich zum realen Prozess passen sich die Parameter hier deutlich zu langsam an, was schnell zu Über- und Unterschätzung von Risiken führen kann.

Das Risikomodell ist eindeutig falsch und sollte daher von den Validierungsverfahren möglichst oft abgelehnt werden.



Statistische Macht bei $\alpha = 5\%$

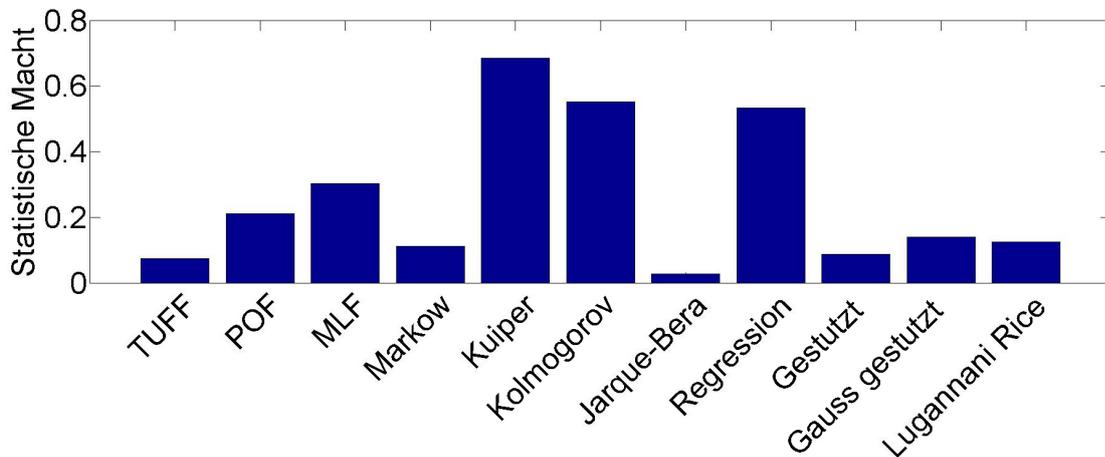


Statistische Macht bei $\alpha = 1\%$

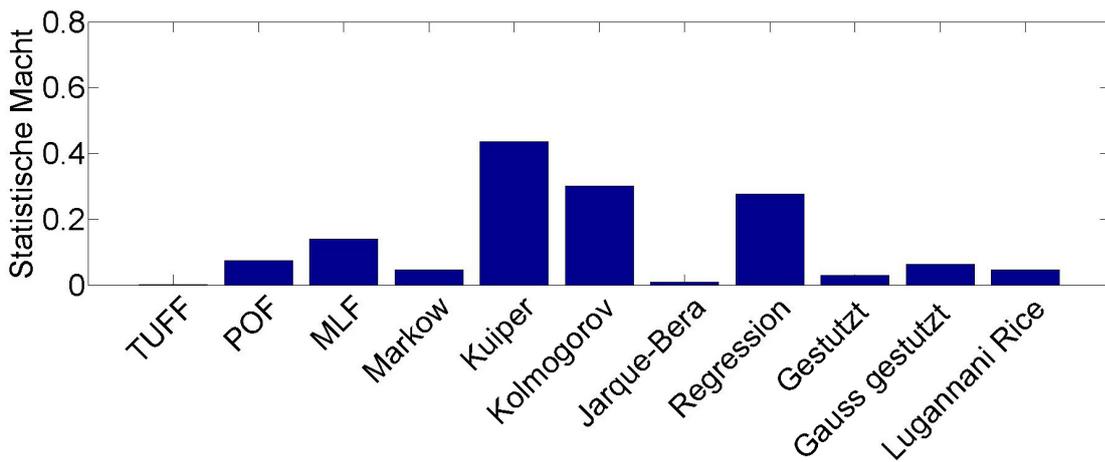
3.3. Risikomodell mit exponentiell geglätteten Varianzen

Im nun präsentierten Risikomodell erfolgen Parameterschätzungen über exponentielle Glättung mit Glättungsparameter 0,94.

Das Modell sollte sich wesentlich schneller an Änderungen anpassen als das vorherige Risikomodell. Dennoch ist auch dieses Modell falsch und sollte von den Validierungsmethoden abgelehnt werden. Die relative Häufigkeit der Ablehnungen der Nullhypothese entspricht der statistischen Macht.



Statistische Macht bei $\alpha = 5\%$

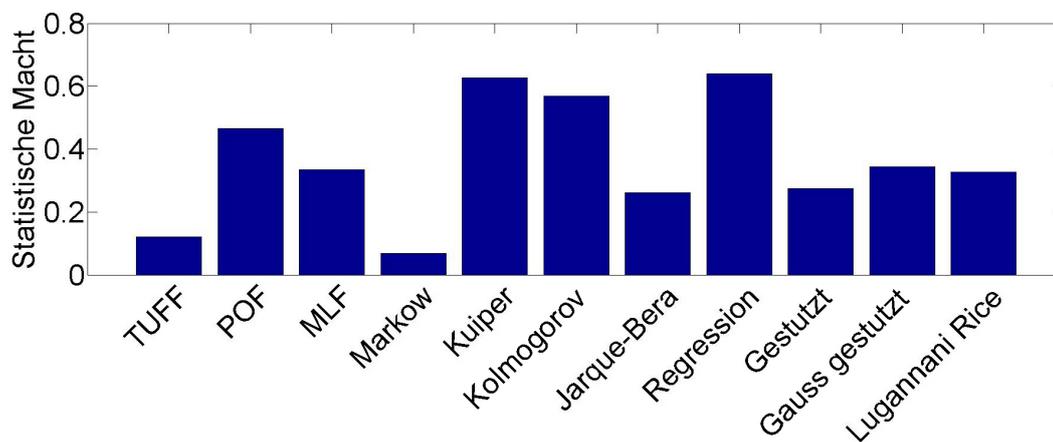


Statistische Macht bei $\alpha = 1\%$

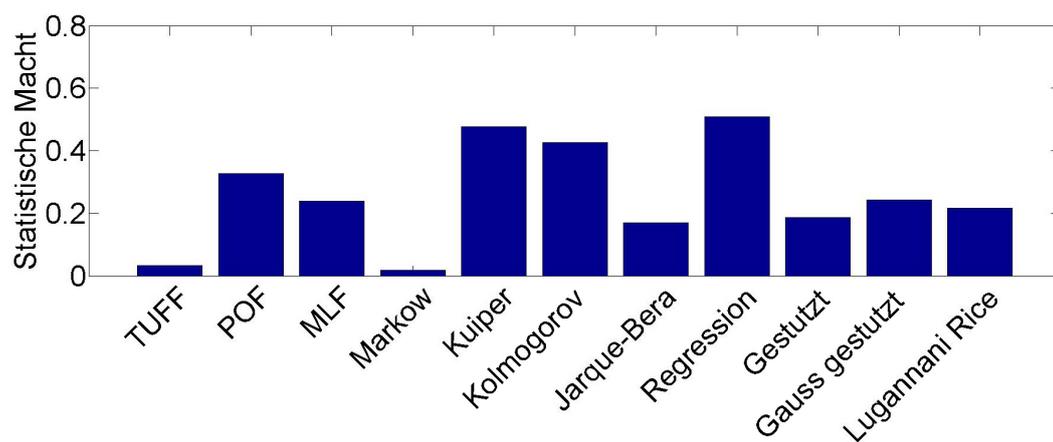
3.4. GARCH-Risikomodell

Ein GARCH-Modell entspricht bereits relativ gut dem tatsächlichen datengenerierenden Prozess. Technisch gesehen ist auch dieses Risikomodell falsch, da es im Gegensatz zum realen GJR-GARCH-Modell keine Asymmetrie von stochastischen Einflüssen einbeziehen kann. In der Praxis würde sich allerdings die Frage stellen, ob wir mit diesem Modell nicht bereits nahe genug an der Realität liegen, um es als korrektes Modell anzuerkennen. Tatsächlich wird man es niemals schaffen, den realen Prozess exakt abzubilden. Mit der GARCH-Prognose liegen wir allerdings schon relativ nahe an der Realität und es könnte deshalb wünschenswert sein, das Modell im Einsatz zu belassen.

Die nachfolgende der statistischen Macht sollte also mit Vorsicht interpretiert werden. Prinzipiell ist es zwar korrekt, das GARCH-Modell abzulehnen - in der Praxis könnte es aber dennoch sinnvoll sein, dieses Modell bewusst nicht abzulehnen.



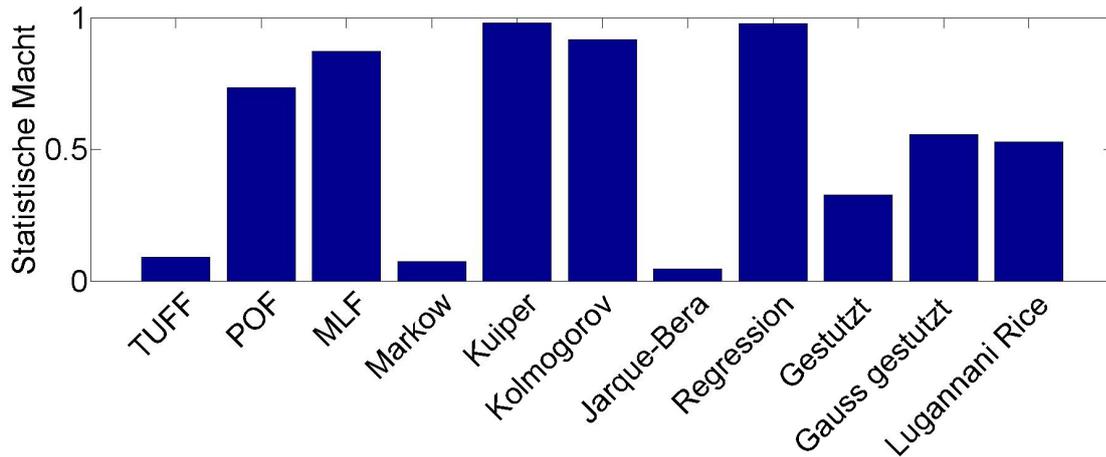
Statistische Macht bei $\alpha = 5\%$



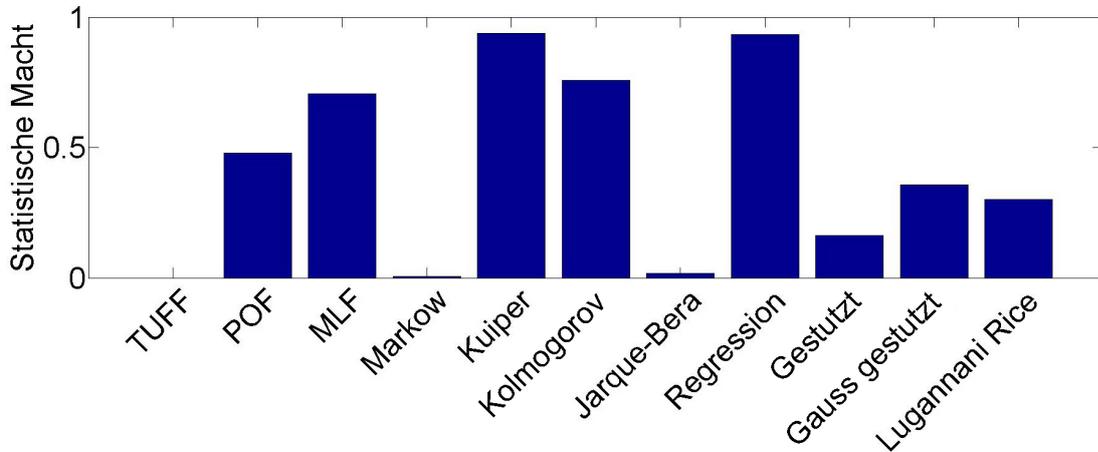
Statistische Macht bei $\alpha = 1\%$

3.5. Konstante Risikounterschätzung

Wir gehen nun dazu über, das Risiko bewusst konstant zu unterschätzen. Wir lassen Daten aus einer Normalverteilung mit Erwartungswert 0 und Varianz 1,5 generieren, gehen im Risikomodell aber von einer Normalverteilung mit Erwartungswert 0 und Varianz 1 aus.



Statistische Macht bei $\alpha = 5\%$



Statistische Macht bei $\alpha = 1\%$

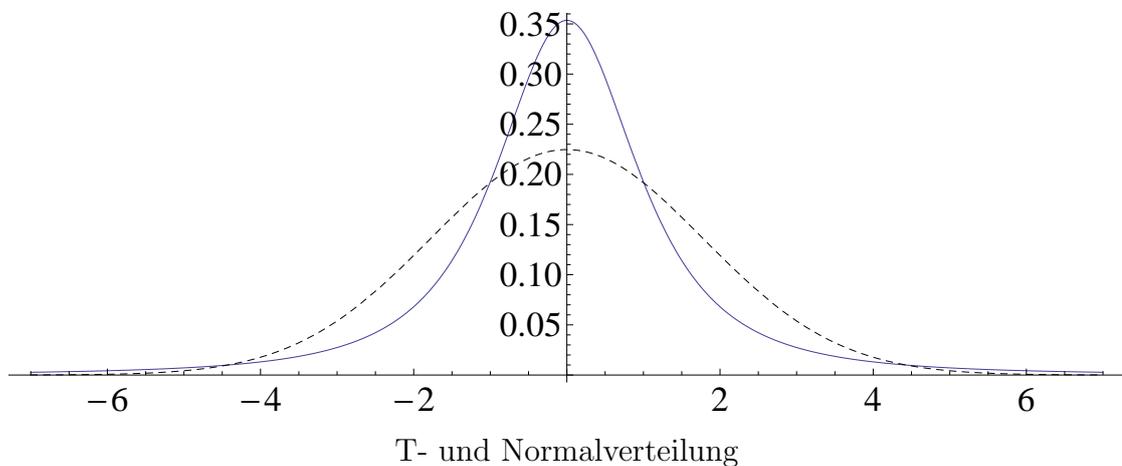
Die niedrige Macht des Markow-Verfahrens ergibt sich aus der Testkonstellation - es liegt schlicht keine Autokorrelation in den Daten vor. Das Markow-Modell hat hier also keinen Anlass, das Modell zu verwerfen.

Im ersten Moment überraschender erscheint hier das schlechte Abschneiden des Jarque Bera-Verfahrens. Im Rahmen dieser konkreten Simulationskonstellation entspricht die Serie, die durch Transformation per Wahrscheinlichkeitsintegral der

Rosenblatt-transformierten Daten hervorgeht, genau der ursprünglich generierten Serie - hier also Realisationen einer Normalverteilung mit Erwartungswert 0 und Varianz 1,5. Der Jarque Bera-Test prüft jedoch lediglich auf Schiefe und Kurtosis. Diese beiden Werte sind natürlich unabhängig von Erwartungswert und Varianz der unterliegenden Normalverteilung. Wir hätten hier also Daten mit beliebigem Erwartungswert und beliebiger Varianz generieren können - der Jarque Bera-Test hätte das Modell trotzdem in ca. $1 - \alpha$ Prozent aller Fälle nicht verworfen. Wir sehen hier sehr deutlich, dass der Jarque Bera-Test nicht ohne ergänzende Tests genutzt werden sollte, da er zwar gewisse Aspekte der Form einer Verteilung überprüft, nicht jedoch die Varianz der Daten.

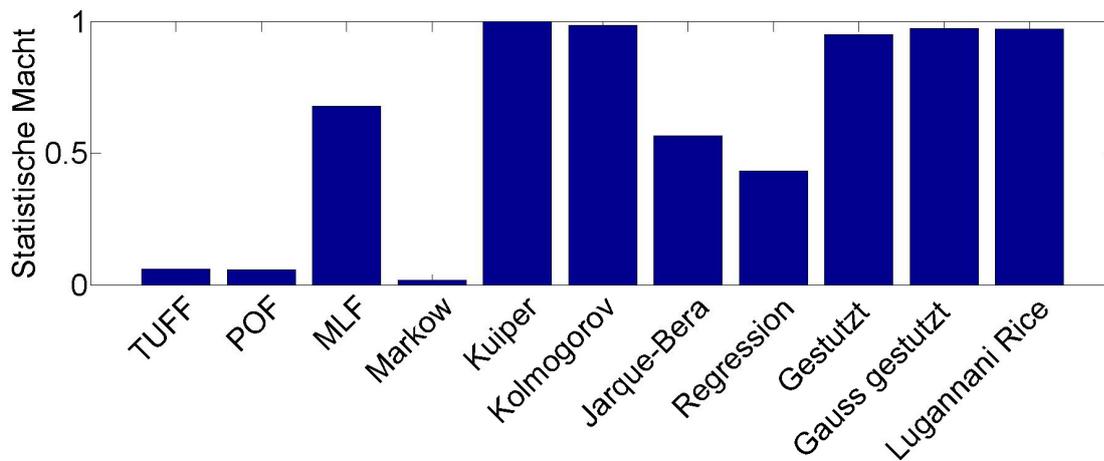
3.6. Korrekter Value at Risk, unterschätzter Expected Shortfall

Wir lassen nun Daten generieren aus einer T-Verteilung mit zwei Freiheitsgraden (blau gezeichneter Graph im Bild), gehen im Risikomodell jedoch von einer Normalverteilung mit Erwartungswert 0 und Standardabweichung 1,78 aus (gepunktete schwarze Linie im Bild):

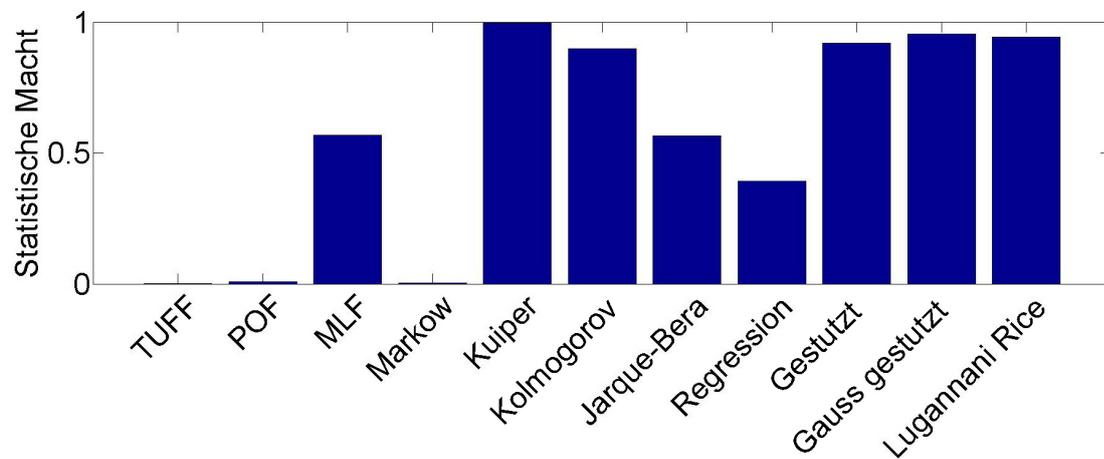


Bei beiden Verteilungen liegt das 5%-Quantil bei -2,92, sodass der Value at Risk auf 5%-Niveau identisch ist. Der Expected Shortfall der T-Verteilung beträgt 6,16, der Expected Shortfall der Normalverteilung hingegen ist 3,66. Das Risikomodell schätzt also einen korrekten Value at Risk, unterschätzt den Expected Shortfall jedoch deutlich.

Validierung von Konzepten zur Messung des Marktrisikos -
insbesondere des Value at Risk und des Expected Shortfall



Statistische Macht bei $\alpha = 5\%$



Statistische Macht bei $\alpha = 1\%$

TUFF und POF schneiden hier erwartungsgemäß schlecht ab, immerhin ist der Value at Risk ja in der Tat korrekt geschätzt. Auch der Markow-Test verwirft H_0 praktisch nie, da es keine Autokorrelation in den Daten gibt. Jarque Bera-Test sowie Regressionsverfahren können ebenfalls nicht überzeugen, da hier die Form der Daten vermutlich nicht ausreichend unterschiedlich zu einer Normalverteilung ist. Am besten schneidet in der Tat der Kuiper-Test ab, der präzise erkennt, dass die reale Verteilung unterschiedlich zur im Risikomodell angenommenen Verteilung sein muss. Das Ergebnis wäre vermutlich jedoch weniger eindeutig, wenn nur der Randbereich unterschiedlich und der Rest der Verteilung identisch mit der im Risikomodell angenommenen Verteilung wäre.

Hier zeigen vor allem der gestutzte Gauss-Test sowie das Lugannani Rice-Verfahren ihre Stärken, denn sie beinhalten nur die Value at Risk-Überschreitungen. Die

statistische Macht dieser beiden Verfahren ist im Test hervorragend, was ihre gute Eigenschaft zur Expected Shortfall-Validierung bestätigt.

3.7. Zusammenfassung der Simulationsergebnisse

Die Simulation gibt interessante Einblicke in Art und Häufigkeit von Ablehnungen der verschiedenen Validierungskonzepte. Die folgenden Aussagen können selbstverständlich nur basierend auf den in der Simulation beinhalteten Konstellationen getroffen werden.

Bei der Überprüfung, ob ein gegebener α -Fehler in einem korrekten Prognosemodell eingehalten wird zeigten manche Modelle bereits eklatante Schwächen. Speziell TUFF-, POF- und Markow-Verfahren zeigten in der Verteilung der Signifikanzniveaus deutliche Abweichungen von einer Gleichverteilung. Auch der Jarque Bera-Test sowie der gestutzte Gauss-Test wiesen, wenngleich auch geringfügigere, Abweichungen von den erwarteten α -Fehlern auf. Diese Abweichungen resultieren jedoch nicht aus Fehlern in den mathematischen Modellen, sondern schlichtweg aus der geringen Stichprobengröße. Im Versuch wurde jeweils mit einem realen Zeitraum von 250 Perioden gearbeitet, was an die Erfordernisse der Praxis angelehnt ist. Simuliert man jeweils einen weitaus längeren Zeitraum, beispielsweise 100.000 Perioden, so zeigt sich bei ausnahmslos allen Verfahren eine nahezu perfekte Gleichverteilung der Signifikanzniveaus. Derartig lange Zeiträume können in der Praxis selbstverständlich nicht umgesetzt werden. Diese erste Simulation liefert also bereits interessante Ergebnisse im Bezug auf die Praxistauglichkeit der Verfahren unter realitätsnahen Anforderungen an die verfügbare Datenmenge.

In den drei folgenden Abschnitten (rollierende Varianz, exponentiell geglättete Varianz sowie GARCH-Modell) wurden die Ablehnungsraten falsch spezifizierter Risikomodelle überprüft. Erwartungsgemäß wies das TUFF-Verfahren durchschnittlich die geringste statistische Macht auf. Auch Markow- sowie Jarque-Bera-Test können durchschnittlich nicht überzeugen. Die besten Resultate liefert der Kuiper-Test. Seine Ergebnisse sind zudem stets besser als die des Kolmogorov-Smirnov-Tests. Da beide Verfahren auf eine bestimmte Verteilungsform testen sprechen die Ergebnisse dafür, dass der Kuiper-Test das geeignetere Verfahren hierfür ist. Ebenfalls sehr gute Ergebnisse zeigt das Regressionsverfahren.

Eine Simulation konstanter Risikounterschätzung bestätigt im Wesentlichen die Einschätzung der Testverfahren und zeigt speziell nochmals die Anfälligkeit des

Jarque-Bera-Tests unter bestimmten Testkonstellationen. Als Schlussfolgerung daraus ist zu beachten, dass der Jarque-Bera-Test ausschließlich ergänzend mit anderen Verfahren genutzt werden sollte - etwa in Verbindung mit dem Regressionsverfahren.

Eine abschließende Simulation, in welcher das Risikomodell zwar eine korrekte Einschätzung des Value at Risk, jedoch eine falsche Einschätzung des Expected Shortfall trifft, liefert interessante Resultate über diejenigen Testverfahren, die speziell einen Teilbereich der Verteilung testen. Alle drei Validierungskonzepte (Gestutzt, Gauss gestutzt, Lugannani Rice) weisen hierbei eine hervorragende statistische Macht auf. Interessanterweise zeigt sich dabei der von uns im Rahmen dieser Arbeit eingeführte gestutzte Gauss-Test als bestes Verfahren mit geringfügig höheren Ablehnungsraten als der reine gestutzte Test oder das Lugannani Rice-Verfahren. Auch der Kuiper- sowie der Kolmogorov-Smirnov-Test zeigen eine sehr hohe statistische Macht in dieser Simulationskonstellation. Dies liegt jedoch vermutlich daran, dass die zugrunde liegende T-Verteilung in ihrer Gesamtheit merklich von der genutzten Normalverteilung abweicht. Würde sie rechtsseitig des p -Quantils mit der Normalverteilung stärker übereinstimmen, so würde die Macht des Kuiper- sowie des Kolmogorov-Smirnov-Verfahrens deutlich geringer ausfallen, wohingegen die statistische Macht der drei anderen Verfahren (gestutzt, Gauss gestutzt, Lugannani Rice) nicht beeinflusst würde. Die letztgenannten Testverfahren sind also tatsächlich sehr effektiv, um Abweichungen im linken Randbereich der Verteilung festzustellen und dienen somit hervorragend zur Validierung des Expected Shortfall.

4. Fazit

Nach einer kurzen Einführung in relevante statistische Verfahren sowie eine einleitende Erklärung zum Marktrisikomanagement im Allgemeinen betrachteten alle gängigen Verfahren zur Validierung von Marktrisikomodellen.

Wir erläuterten zuerst Verfahren, welche die Güte eines Risikomodells anhand Überschreitungen des Value at Risk beurteilen. TUFF, POF, das Ampelkonzept, die MLF sowie der Markow-Test auf Unabhängigkeit zählen zu diesen Verfahren.

Weiterführend zur bisherigen Literatur zeigten wir dabei einige zusätzliche Aspekte einiger Testverfahren auf. Beim TUFF-Verfahren fanden wir erhebliche Inkonsistenzen bei der statistischen Macht sowie bei der Einhaltung des α -Fehlers. Wir zeigten ebenfalls, dass eine mangelnde Ablehnungsfähigkeit risikounterschätzender Modelle erst später auftritt als bisher von Kupiec (1998) angenommen.

Vergleichend zu diesem Verfahren simulierten wir auch beim TUFF-Test α - und β -Fehler und fanden heraus, dass es auch bei ersterem kleinere Abweichungen gibt, letzterer sich jedoch wesentlich konstanter als beim POF-Verfahren entwickelt.

Die MLF modifizierten wir im Unterschied zur bisherigen Praxis um eine Korrektur der Standardabweichung und berechneten kritische Werte für eine Normalverteilung.

Anschließend betrachteten wir Verfahren, die die gesamte Verteilung validieren. Dazu zählt die Rosenblatt-Transformation und darauf basierend die Kuiper-Statistik mit Worry Function und der Kolmogorov-Smirnov-Test mit Worry Function sowie, basierend auf dem Wahrscheinlichkeitsintegral zur Normalverteilung, der Jarque-Bera-Test und eine Regression auf Mittelwert, Varianz und Autokorrelation.

Mit dem Kolmogorov-Smirnov-Test führten wir ein neues Verfahren ein, welches sich jedoch im Bezug auf seine statistische Macht im Rahmen der getesteten Konstellationen später im Vergleich zur Kuiper-Statistik als geringfügig unterlegen herausstellte. Zu beiden Verfahren simulierten wir mit einer großen Datenmenge kritische Werte unter Einbeziehung der Worry Function.

Zuletzt erläuterten wir Verfahren, welche speziell für die Validierung des Expected Shortfall von Interesse sind. Dazu zählt der gestutzte Likelihood-Quotienten-Test, der Gauss-Test des Expected Shortfall bei gestutzter Verteilung sowie die Sattelpunktnäherung mittels Lugannani Rice-Formel. Beim Erstgenannten wichen

wir von der bisherigen Literatur zum Thema ab, indem wir statt einer zensierten Verteilung eine gestutzte Verteilung betrachteten und die entsprechenden Formeln herleiteten. Mit dem Gauss-Test bei gestutzter Verteilung führten wir ein gänzlich neues Verfahren ein.

Abschließend simulierten wir für alle angeführten Testverfahren die Signifikanzniveaus sowie die statistische Macht unter verschiedenen Konstellationen. Wir zeigten Probleme einiger Modelle bei der Einhaltung des α -Fehlers, die Aufschluss über die Praxistauglichkeit der jeweiligen Verfahren bei realitätsnahen Datenmengen geben. Des Weiteren belegten wir, dass Verfahren, die nur auf Überschreitungen des Value at Risk basieren, in der Tat eine recht geringe statistische Macht aufweisen. Wir sahen, dass eine Validierung der gesamten Verteilung wesentlich erfolgreicher war und führten ebenfalls aus, dass die Verfahren, die auf einem Teilbereich der Verteilung basieren, extrem effizient bei der Ablehnung von Modellen mit falsch spezifiziertem Expected Shortfall arbeiten.

Die Validierung von Marktrisikomodellen wird sicherlich weiterhin ein relevantes Thema in der Forschung und in der Finanzindustrie bleiben. Welche der vorgestellten Modelle tatsächlich in der Praxis genutzt werden sollten obliegt der Entscheidung des Risikomanagements. Es empfiehlt sich in den meisten Fällen, mehrere Verfahren zu kombinieren. Die Eigenschaften der Validierungsmodelle sollten dabei abgewogen und individuell auf die Prioritäten sowie die Charakteristika des Portfolios abgestimmt werden. Die Resultate dieser Arbeit dürften eine Hilfe dabei sein, konkrete Stärken und Schwächen der Modelle unter verschiedenen Konstellationen einzuschätzen.

Literaturverzeichnis

- Berkowitz, Jeremy. 2001. "Testing Density Forecasts, With Applications to Risk Management." *Journal of Business and Economic Statistics* p. 465.
- Campbell, Sean D. 2006/2007. "A review of backtesting and backtesting procedures." *The Journal of Risk* p. 1.
- Christoffersen, Peter. 2003. *Elements of Financial Risk Management*. Academic Press.
- Cremers, Heinz. 2011. "Risikomanagement." Vorlesungsskript. SS 2011.
- Crnkovic, Cedomir & Jordan Drachman. 1996. "Quality Control." *Risk* 9:138–142.
- DasGupta, Anirban. 2008. *Asymptotic Theory of Statistics and Probability*. Springer.
- Dowd, Kevin. 2004. "A modified Berkowitz back-test." *Risk Magazine* .
- Goethe Universität. 2004. "Mathematisch-Statistische Verfahren des Risikomanagements." Präsentation. SS 2004.
- Jost, Julian, Thorsten Spree, Katarina Hapke, Fabian Mehmke & Niclas Huck. 2011. "Hypothesentests.". Seminararbeit Quantitative Methoden der Ökonometrie.
- Kupiec, Paul H. 1998. "Techniques for Verifying the Accuracy of Risk Measurement Models." *Journal of Derivatives* pp. 73–84.
- Lopez, Jose A. 1998. Methods for Evaluating Value-at-Risk Estimates. Technical report Federal Reserve Bank of New York.
- Ruppert, David. 2010. *Statistics and Data Analysis for Financial Engineering (Springer Texts in Statistics)*. Springer.
- Sanddorf-Köhle, Walter. 2011. "Zeitreihenanalyse Sommersemester 2011." Vorlesungsskript. Foliensatz Nr. 6.
- Schröder, Frank & Hermann Schulte-Mattler. 1997. "CD-Verfahren als Alternative zum Baseler Backtesting." *Die Bank* 7.

Suhov, Yuri & Mark Kelbert. 2005. *Probability and Statistics by Example: Volume 1, Basic Probability and Statistics (v. 1)*. Cambridge University Press.

Supervisory framework for the use of "Backtesting" in conjunction with the internal models approach to market risk capital requirements. 1998. Technical report Bank for International Settlements.

van der Vaart, A. W. 2000. *Asymptotic Statistics (Cambridge Series in Statistical and Probabilistic Mathematics)*. Cambridge University Press.

Wong, Woon K. 2008. "Backtesting trading risk of commercial banks using expected shortfall." *Journal of Banking and Finance* 32:1404–1415.

Nomenklatur

α	Risikowahrscheinlichkeit eines Hypothesentests. Je nach Zusammenhang auch Parameter in einem APGARCH-Modell.
β	Fehler zweiter Art eines Hypothesentests. Je nach Zusammenhang auch Parameter in einem APGARCH-Modell.
χ_d^2	Chi-Quadrat-Verteilung mit Freiheitsgrad d
δ	Parameter in einem APGARCH-Modell
ϵ_t	Residuum einer Regression
λ	Glättungsparameter bei exponentiell geglätteter Varianz
$\Lambda(\cdot)$	Teststatistik eines allgemeinen Likelihood-Quotienten-Tests
\ln	Natürlicher Logarithmus
$\mathbb{B}(p)$	Bernoulliverteilung mit Parameter p .
$\mathbb{B}(T, p)$	Binomialverteilung mit Parametern T und p .
$\mathbb{G}(p)$	Geometrische Verteilung mit Parameter p .
\mathbb{N}	Natürliche Zahlen
$\mathbb{N}(\mu, \sigma)$	Eine Normalverteilung mit Erwartungswert μ und Varianz σ^2
$\mathbb{N}(\mu, \sigma^2, \mathcal{T})$	Normalverteilung mit Parametern μ, σ^2 , die an der Stelle \mathcal{T} gestutzt ist.
\mathbb{R}	Die reellen Zahlen
\mathbb{S}	Menge aller Parameterkombinationen einer Simulation
\mathbb{U}	Gleichverteilung auf dem Intervall $(0, 1)$.

\mathcal{O}	Landau-Symbol
Π	Spaltenvektor aller q-Werte, $(q_1, \dots, q_T)^T$.
\mathcal{T}	Stützungsstelle einer Verteilung.
\mathfrak{T}_{ij}	Absolute Häufigkeit, in der auf eine Value at Risk-Überschreitung ($i = 1$) bzw. einer Nicht-Überschreitung ($i = 0$) in einer Stichprobe eine Überschreitung ($j = 1$) oder eine Nicht-Überschreitung ($j = 0$) folgt.
μ	Erwartungswert
ω	Parameter in einem APGARCH-Modell
Ω_t	Zustandsraum zum Zeitpunkt t
$\bar{\cdot}$	Mittelwert, z.B. wäre \bar{q} der Mittelwert aller q_1, \dots, q_T
$\Phi(\cdot)$	Verteilungsfunktion einer Standardnormalverteilung
$\phi(\cdot)$	Dichtefunktion einer Standardnormalverteilung
$\Phi^{-1}(\cdot)$	Inverse Verteilungsfunktion einer Standardnormalverteilung
Π	Transitionsmatrix eines Markow-Tests
π	Kreiszahl
π_{ij}	Elemente der Transitionsmatrix Π
Π_{ML}	Maximum Likelihood-Schätzung einer Transitionsmatrix
Π_U	Transitionsmatrix unter Unabhängigkeit
$\mathcal{P}(\cdot)$	Wahrscheinlichkeitsoperator
ρ	Korrelationskoeffizient / Autokorrelationsparameter bei einer linearen Regression
σ	Standardabweichung
σ^2	Varianz
τ	Die Menge aller Möglichen Zeitindizes $\{1, \dots, T\}$.

Θ	$\Theta_0 \cup \Theta_1$
θ	Parametervektor
Θ_0	Parameterraum einer parametrischen Nullhypothese
Θ_1	Parameterraum einer parametrischen Alternativhypothese
$\Upsilon(\cdot)$	Teststatistik des Gauss-Tests bei gestutzter Verteilung.
ϑ	Ein unbekannter Verteilungsparameter
$\hat{\cdot}$	Empirische Schätzungen bestimmter Werte. Zum Beispiel $\hat{\mu}, \hat{\sigma}$ etc.
A	Ablehnungsbereich eines Hypothesentests
D	Kolmogorov-Smirnov-Teststatistik
D_w	Mit Worry Function gewichtete Kolmogorov-Smirnov-Statistik
$E(\cdot)$	Erwartungswert-Operator
$ES_t(p)$	Expected Shortfall zum Zeitpunkt t für Value at Risk-Vertrauenswahrscheinlichkeit $1 - p$.
$f_X(x; \Theta)$	Dichtefunktion der Variable X an der Stelle x mit Parametervektor θ . Auch ohne X und θ geschrieben, sofern deren Bedeutungen aus dem Zusammenhang deutlich werden.
$F_X(x; \theta)$	Verteilungsfunktion der Zufallsvariable X an der Stelle x . Auch ohne X und θ geschrieben, sofern deren Bedeutungen aus dem Zusammenhang deutlich werden.
H_0	Nullhypothese
H_1	Alternativhypothese
I	Summe der Überschreitungsfunktionen zu jedem Zeitpunkt in τ .
I_t	Value at Risk-Überschreitungsfunktion zum Zeitpunkt t .
JB	Jarque-Bera-Teststatistik
K	Kurtosis einer Verteilung beim Jarque-Bera-Test

- $K(\cdot)$ Kumulantenerzeugende Funktion
- $L(\cdot)$ Likelihood-Funktion
- $M(\cdot)$ Momentenerzeugende Funktion
- N Umfang einer Stichprobe
- $n(x)$ Anzahl aller Werte in einer Stichprobe z_1, \dots, z_T , die kleiner als x sind.
- p Tatsächliche, unbekannte Wahrscheinlichkeit, dass der Value at Risk überschritten wird.
- p^* Modellwahrscheinlichkeit, mit der der Value at Risk annahmegemäß überschritten werden sollte.
- Q Kuiper-Statistik
- Q^* Zufallsvariable wie Q , jedoch nur definiert für Werte kleinergleich $F_Q^{-1}(p)$, d.h. also Stützung von Q .
- Q_t Per Wahrscheinlichkeitsintegral transformierte Zufallsvariable Z_t
- Q_w Mit Worry Function gewichtete Kuiper-Statistik
- r_t Diskrete Rendite zum Zeitpunkt t
- S Schiefe einer Verteilung beim Jarque-Bera-Test
- T Die Anzahl der beobachteten Zeitpunkte. Je nach Kontext auch Teststatistik eines Hypothesentests.
- $T(\mathbf{x})$ Testfunktion eines Hypothesentests, wobei $x \in \mathbb{R}^n$
- T^* Anzahl aller Realisationen von Q_1, \dots, Q_T , die kleinergleich $F_Q^{-1}(p)$ sind.
- $Var(\cdot)$ Varianz-Operator
- $VaR_t(p)$ Value at Risk für den Zeitpunkt t mit Vertrauenswahrscheinlichkeit $1-p$
- $W(x)$ Worry Function
- X Eine Zufallsvariable

Z_t	Eine Zufallsvariable, die der Rosenblatt-Transformation der Zufallsvariable t entspricht.
\mathbf{s}	Vektor, der Werte einer Teststatistik zu verschiedenen Stichproben beinhaltet.
d	Unterschied der Dimensionen zwischen Θ und Θ_0 .
$G(\cdot)$	Gütefunktion
$gT(\cdot)$	Ganzzahliger Teil einer Zahl.
ML	Maximum Likelihood
MLF	Magnitude of Loss-Function
POF	Proportion of Failure
TLA	Traffic Light Approach, entspricht dem Ampelkonzept.
TUFF	Time Until First Failure

A. Anhang

A.1. Tabellen

	$\alpha=0,1\%$	1%	2%	3%	4%
$p^* = 1\%$	<1 852.31	1.88 610.51	3.02 536.86	4.11 493.52	5.16 462.64
2%	<1 424.26	1.23 303.96	1.78 267.33	2.32 245.77	2.83 230.4
3%	<1 281.57	1.05 201.78	1.38 177.48	1.73 163.18	2.06 152.99
4%	<1 210.22	<1 150.68	1.19 132.55	1.44 121.88	1.68 114.28
5%	<1 167.41	<1 120.03	1.08 105.6	1.27 97.1	1.46 91.05
6%	<1 138.86	<1 99.59	1.02 87.62	1.16 80.58	1.31 75.57
7%	<1 118.47	<1 84.98	<1 74.78	1.09 68.78	1.21 64.51
8%	<1 103.18	<1 74.03	<1 65.15	1.04 59.93	1.14 56.21
9%	<1 91.28	<1 65.51	<1 57.66	1.01 53.04	1.09 49.75
10%	<1 81.76	<1 58.69	<1 51.67	<1 47.53	1.05 44.59
11%	<1 73.97	<1 53.12	<1 46.76	<1 43.03	1.02 40.36
12%	<1 67.48	<1 48.47	<1 42.68	<1 39.27	1 36.84
13%	<1 61.98	<1 44.53	<1 39.22	<1 36.09	<1 33.86
14%	<1 57.27	<1 41.16	<1 36.25	<1 33.36	<1 31.3
15%	<1 53.18	<1 38.23	<1 33.68	<1 31	<1 29.09
16%	<1 49.61	<1 35.67	<1 31.43	<1 28.93	<1 27.15
17%	<1 46.45	<1 33.41	<1 29.44	<1 27.1	<1 25.44
18%	<1 43.65	<1 31.4	<1 27.67	<1 25.48	<1 23.91
19%	<1 41.14	<1 29.6	<1 26.09	<1 24.03	<1 22.55
20%	<1 38.88	<1 27.99	<1 24.67	<1 22.72	<1 21.33

Tabelle A.1.: TUFF-Test, kritische Werte für unterschiedliche p^*

Validierung von Konzepten zur Messung des Marktrisikos -
insbesondere des Value at Risk und des Expected Shortfall

	$\alpha=5\%$	6%	7%	8%	9%	10%
$p^* = 1\%$	6.2 438.6	7.2 418.9	8.2 402.2	9.1 387.7	10.1 374.8	11 363.3
2%	3.3 218.4	3.8 208.6	4.3 200.3	4.8 193.1	5.3 186.7	5.7 181
3%	2.4 145.1	2.7 138.6	3 133	3.4 128.3	3.7 124	4 120.2
4%	1.9 108.4	2.2 103.5	2.4 99.4	2.6 95.8	2.9 92.7	3.1 89.8
5%	1.7 86.3	1.8 82.5	2 79.2	2.2 76.4	2.4 73.8	2.6 71.6
6%	1.5 71.7	1.6 68.5	1.8 65.7	1.9 63.4	2.1 61.3	2.2 59.4
7%	1.3 61.2	1.5 58.4	1.6 56.1	1.7 54.1	1.9 52.3	2 50.8
8%	1.3 53.3	1.4 50.9	1.5 48.9	1.6 47.2	1.7 45.6	1.8 44.2
9%	1.2 47.2	1.3 45.1	1.4 43.3	1.5 41.8	1.6 40.4	1.7 39.2
10%	1.1 42.3	1.2 40.4	1.3 38.8	1.4 37.4	1.5 36.2	1.5 35.1
11%	1.1 38.3	1.2 36.6	1.2 35.1	1.3 33.9	1.4 32.8	1.5 31.8
12%	1.1 34.9	1.1 33.4	1.2 32.1	1.2 30.9	1.3 29.9	1.4 29
13%	1 32.1	1.1 30.7	1.1 29.5	1.2 28.4	1.3 27.5	1.3 26.7
14%	1 29.7	1.1 28.4	1.1 27.3	1.2 26.3	1.2 25.5	1.3 24.7
15%	<1 27.6	1 26.4	1.1 25.3	1.1 24.5	1.2 23.7	1.2 22.9
16%	<1 25.8	1 24.6	1 23.7	1.1 22.8	1.1 22.1	1.2 21.4
17%	<1 24.1	1 23.1	1 22.2	1.1 21.4	1.1 20.7	1.2 20.1
18%	<1 22.7	<1 21.7	1 20.9	1 20.1	1.1 19.5	1.1 18.9
19%	<1 21.4	<1 20.5	1 19.7	1 19	1.1 18.4	1.1 17.8
20%	<1 20.2	<1 19.4	<1 18.6	1 17.9	1 17.4	1.1 16.9

Tabelle A.2.: Fortsetzung TUFF-Test, kritische Werte für unterschiedliche p^*

FRANKFURT SCHOOL / HFB – WORKING PAPER SERIES

No.	Author/Title	Year
191.	Tinschert, Jonas / Cremers, Heinz Fixed Income Strategies for Trading and for Asset Management	2012
190.	Schultz, André / Kozlov, Vladimir / Libman, Alexander Roving Bandits in Action: Outside Option and Governmental Predation in Autocracies	2012
189.	Börner, René / Goeken, Matthias / Rabhi, Fethi SOA Development and Service Identification – A Case Study on Method Use, Context and Success Factors	2012
188.	Mas, Ignacio / Klein, Michael A Note on Macro-financial implications of mobile money schemes	2012
187.	Dietmar Harhoff, Elisabeth Müller, John Van Reenen What are the Channels for Technology Sourcing? Panel Data Evidence from German Companies	2012
186.	Decarolis, Francesco/ Klein, Michael Auctions that are too good to be true	2012
185.	Klein, Michael Infrastructure Policy: Basic Design Options	2012
184.	Eaton, Sarah / Kostka, Genia Does Cadre Turnover Help or Hinder China’s Green Rise? Evidence from Shanxi Province	2012
183.	Behley, Dustin / Leyer, Michael Evaluating Concepts for Short-term Control in Financial Service Processes	2011
182.	Herrmann-Pillath, Carsten Naturalizing Institutions: Evolutionary Principles and Application on the Case of Money	2011
181.	Herrmann-Pillath, Carsten Making Sense of Institutional Change in China: The Cultural Dimension of Economic Growth and Modernization	2011
180.	Herrmann-Pillath, Carsten Hayek 2.0: Grundlinien einer naturalistischen Theorie wirtschaftlicher Ordnungen	2011
179.	Braun, Daniel / Allgeier, Burkhard / Cremres, Heinz Ratingverfahren: Diskriminanzanalyse versus Logistische Regression	2011
178.	Kostka, Genia / Moslener, Ulf / Andreas, Jan G. Barriers to Energy Efficiency Improvement: Empirical Evidence from Small- and-Medium-Sized Enterprises in China	2011
177.	Löchel, Horst / Xiang Li, Helena Understanding the High Profitability of Chinese Banks	2011
176.	Herrmann-Pillath, Carsten Neuroökonomik, Institutionen und verteilte Kognition: Empirische Grundlagen eines nicht-reduktionistischen naturalistischen Forschungsprogramms in den Wirtschaftswissenschaften	2011
175.	Libman, Alexander/ Mendelski, Martin History Matters, but How? An Example of Ottoman and Habsburg Legacies and Judicial Performance in Romania	2011
174.	Kostka, Genia Environmental Protection Bureau Leadership at the Provincial Level in China: Examining Diverging Career Backgrounds and Appointment Patterns	2011
173.	Durst, Susanne / Leyer, Michael Bedürfnisse von Existenzgründern in der Gründungsphase	2011
172.	Klein, Michael Enrichment with Growth	2011
171.	Yu, Xiaofan A Spatial Interpretation of the Persistency of China’s Provincial Inequality	2011
170.	Leyer, Michael Stand der Literatur zur operativen Steuerung von Dienstleistungsprozessen	2011
169.	Libman, Alexander / Schultz, André Tax Return as a Political Statement	2011
168.	Kostka, Genia / Shin, Kyoung Energy Service Companies in China: The Role of Social Networks and Trust	2011
167.	Andriani, Pierpaolo / Herrmann-Pillath, Carsten Performing Comparative Advantage: The Case of the Global Coffee Business	2011
166.	Klein, Michael / Mayer, Colin Mobile Banking and Financial Inclusion: The Regulatory Lessons	2011

165.	Cremers, Heinz / Hewicker, Harald Modellierung von Zinsstrukturkurven	2011
164.	Roßbach, Peter / Karlow, Denis The Stability of Traditional Measures of Index Tracking Quality	2011
163.	Libman, Alexander / Herrmann-Pillath, Carsten / Yarav, Gaudav Are Human Rights and Economic Well-Being Substitutes? Evidence from Migration Patterns across the Indian States	2011
162.	Herrmann-Pillath, Carsten / Andriani, Pierpaolo Transactional Innovation and the De-commodification of the Brazilian Coffee Trade	2011
161.	Christian Büchler, Marius Buxkaemper, Christoph Schalast, Gregor Wedell Incentivierung des Managements bei Unternehmenskäufen/Buy-Outs mit Private Equity Investoren – eine empirische Untersuchung –	2011
160.	Herrmann-Pillath, Carsten Revisiting the Gaia Hypothesis: Maximum Entropy, Kauffman’s “Fourth Law” and Physiosemeiosis	2011
159.	Herrmann-Pillath, Carsten A ‘Third Culture’ in Economics? An Essay on Smith, Confucius and the Rise of China	2011
158.	Boeing, Philipp / Sandner, Philipp The Innovative Performance of China’s National Innovation System	2011
157.	Herrmann-Pillath, Carsten Institutions, Distributed Cognition and Agency: Rule-following as Performative Action	2011
156.	Wagner, Charlotte From Boom to Bust: How different has microfinance been from traditional banking?	2010
155.	Libman Alexander / Vinokurov, Evgeny Is it really different? Patterns of Regionalisation in the Post-Soviet Central Asia	2010
154.	Libman, Alexander Subnational Resource Curse: Do Economic or Political Institutions Matter?	2010
153.	Herrmann-Pillath, Carsten Meaning and Function in the Theory of Consumer Choice: Dual Selves in Evolving Networks	2010
152.	Kostka, Genia / Hobbs, William Embedded Interests and the Managerial Local State: Methanol Fuel-Switching in China	2010
151.	Kostka, Genia / Hobbs, William Energy Efficiency in China: The Local Bundling of Interests and Policies	2010
150.	Umber, Marc P. / Grote, Michael H. / Frey, Rainer Europe Integrates Less Than You Think. Evidence from the Market for Corporate Control in Europe and the US	2010
149.	Vogel, Ursula / Winkler, Adalbert Foreign banks and financial stability in emerging markets: evidence from the global financial crisis	2010
148.	Libman, Alexander Words or Deeds – What Matters? Experience of Decentralization in Russian Security Agencies	2010
147.	Kostka, Genia / Zhou, Jianghua Chinese firms entering China’s low-income market: Gaining competitive advantage by partnering governments	2010
146.	Herrmann-Pillath, Carsten Rethinking Evolution, Entropy and Economics: A triadic conceptual framework for the Maximum Entropy Principle as applied to the growth of knowledge	2010
145.	Heidorn, Thomas / Kahlert, Dennis Implied Correlations of iTraxx Tranches during the Financial Crisis	2010
144.	Fritz-Morgenthal, Sebastian G. / Hach, Sebastian T. / Schalast, Christoph M&A im Bereich Erneuerbarer Energien	2010
143.	Birkmeyer, Jörg / Heidorn, Thomas / Rogalski, André Determinanten von Banken-Spreads während der Finanzmarktkrise	2010
142.	Bannier, Christina E. / Metz, Sabrina Are SMEs large firms en miniature? Evidence from a growth analysis	2010
141.	Heidorn, Thomas / Kaiser, Dieter G. / Voinea, André The Value-Added of Investable Hedge Fund Indices	2010
140.	Herrmann-Pillath, Carsten The Evolutionary Approach to Entropy: Reconciling Georgescu-Roegen’s Natural Philosophy with the Maximum Entropy Framework	2010
139.	Heidorn, Thomas / Löw, Christian / Winker, Michael Funktionsweise und Replikationstil europäischer Exchange Traded Funds auf Aktienindices	2010
138.	Libman, Alexander Constitutions, Regulations, and Taxes: Contradictions of Different Aspects of Decentralization	2010

137.	Herrmann-Pillath, Carsten / Libman, Alexander / Yu, Xiaofan State and market integration in China: A spatial econometrics approach to ‘local protectionism’	2010
136.	Lang, Michael / Cremers, Heinz / Hentze, Rainald Ratingmodell zur Quantifizierung des Ausfallrisikos von LBO-Finanzierungen	2010
135.	Bannier, Christina / Feess, Eberhard When high-powered incentive contracts reduce performance: Choking under pressure as a screening device	2010
134.	Herrmann-Pillath, Carsten Entropy, Function and Evolution: Naturalizing Peircian Semiosis	2010
133.	Bannier, Christina E. / Behr, Patrick / Güttler, Andre Rating opaque borrowers: why are unsolicited ratings lower?	2009
132.	Herrmann-Pillath, Carsten Social Capital, Chinese Style: Individualism, Relational Collectivism and the Cultural Embeddedness of the Institutions-Performance Link	2009
131.	Schäffler, Christian / Schmaltz, Christian Market Liquidity: An Introduction for Practitioners	2009
130.	Herrmann-Pillath, Carsten Dimensionen des Wissens: Ein kognitiv-evolutionärer Ansatz auf der Grundlage von F.A. von Hayeks Theorie der „Sensory Order“	2009
129.	Hankir, Yassin / Rauch, Christian / Umber, Marc It’s the Market Power, Stupid! – Stock Return Patterns in International Bank M&A	2009
128.	Herrmann-Pillath, Carsten Outline of a Darwinian Theory of Money	2009
127.	Cremers, Heinz / Walzner, Jens Modellierung des Kreditrisikos im Portfoliofall	2009
126.	Cremers, Heinz / Walzner, Jens Modellierung des Kreditrisikos im Einwertpapierfall	2009
125.	Heidorn, Thomas / Schmaltz, Christian Interne Transferpreise für Liquidität	2009
124.	Bannier, Christina E. / Hirsch, Christian The economic function of credit rating agencies - What does the watchlist tell us?	2009
123.	Herrmann-Pillath, Carsten A Neurolinguistic Approach to Performativity in Economics	2009
122.	Winkler, Adalbert / Vogel, Ursula Finanzierungsstrukturen und makroökonomische Stabilität in den Ländern Südosteuropas, der Türkei und in den GUS-Staaten	2009
121.	Heidorn, Thomas / Rupprecht, Stephan Einführung in das Kapitalstrukturmanagement bei Banken	2009
120.	Roszbach, Peter Die Rolle des Internets als Informationsbeschaffungsmedium in Banken	2009
119.	Herrmann-Pillath, Carsten Diversity Management und diversitätsbasiertes Controlling: Von der „Diversity Scorecard“ zur „Open Balanced Scorecard“	2009
118.	Hölscher, Luise / Clasen, Sven Erfolgsfaktoren von Private Equity Fonds	2009
117.	Bannier, Christina E. Is there a hold-up benefit in heterogeneous multiple bank financing?	2009
116.	Roßbach, Peter / Gießbamer, Dirk Ein eLearning-System zur Unterstützung der Wissensvermittlung von Web-Entwicklern in Sicherheitsthemen	2009
115.	Herrmann-Pillath, Carsten Kulturelle Hybridisierung und Wirtschaftstransformation in China	2009
114.	Schalast, Christoph: Staatsfonds – „neue“ Akteure an den Finanzmärkten?	2009
113.	Schalast, Christoph / Alram, Johannes Konstruktion einer Anleihe mit hypothekarischer Besicherung	2009
112.	Schalast, Christoph / Bolder, Markus / Radünz, Claus / Siepmann, Stephanie / Weber, Thorsten Transaktionen und Servicing in der Finanzkrise: Berichte und Referate des Frankfurt School NPL Forums 2008	2009
111.	Werner, Karl / Moormann, Jürgen Efficiency and Profitability of European Banks – How Important Is Operational Efficiency?	2009

110.	Herrmann-Pillath, Carsten Moralische Gefühle als Grundlage einer wohlstandschaffenden Wettbewerbsordnung: Ein neuer Ansatz zur erforschung von Sozialkapital und seine Anwendung auf China	2009
109.	Heidorn, Thomas / Kaiser, Dieter G. / Roder, Christoph Empirische Analyse der Drawdowns von Dach-Hedgefonds	2009
108.	Herrmann-Pillath, Carsten Neuroeconomics, Naturalism and Language	2008
107.	Schalast, Christoph / Benita, Barten Private Equity und Familienunternehmen – eine Untersuchung unter besonderer Berücksichtigung deutscher Maschinen- und Anlagenbauunternehmen	2008
106.	Bannier, Christina E. / Grote, Michael H. Equity Gap? – Which Equity Gap? On the Financing Structure of Germany's Mittelstand	2008
105.	Herrmann-Pillath, Carsten The Naturalistic Turn in Economics: Implications for the Theory of Finance	2008
104.	Schalast, Christoph (Hrsg.) / Schanz, Kay-Michael / Scholl, Wolfgang Aktionärschutz in der AG falsch verstanden? Die Leica-Entscheidung des LG Frankfurt am Main	2008
103.	Bannier, Christina E./ Müsch, Stefan Die Auswirkungen der Subprime-Krise auf den deutschen LBO-Markt für Small- und MidCaps	2008
102.	Cremers, Heinz / Vetter, Michael Das IRB-Modell des Kreditrisikos im Vergleich zum Modell einer logarithmisch normalverteilten Verlustfunktion	2008
101.	Heidorn, Thomas / Pleißner, Mathias Determinanten Europäischer CMBS Spreads. Ein empirisches Modell zur Bestimmung der Risikoaufschläge von Commercial Mortgage-Backed Securities (CMBS)	2008
100.	Schalast, Christoph (Hrsg.) / Schanz, Kay-Michael Schaeffler KG/Continental AG im Lichte der CSX Corp.-Entscheidung des US District Court for the Southern District of New York	2008
99.	Hölscher, Luise / Haug, Michael / Schweinberger, Andreas Analyse von Steueramnestiedaten	2008
98.	Heimer, Thomas / Arend, Sebastian The Genesis of the Black-Scholes Option Pricing Formula	2008
97.	Heimer, Thomas / Hölscher, Luise / Werner, Matthias Ralf Access to Finance and Venture Capital for Industrial SMEs	2008
96.	Böttger, Marc / Guthoff, Anja / Heidorn, Thomas Loss Given Default Modelle zur Schätzung von Recovery Rates	2008
95.	Almer, Thomas / Heidorn, Thomas / Schmaltz, Christian The Dynamics of Short- and Long-Term CDS-spreads of Banks	2008
94.	Barthel, Erich / Wollersheim, Jutta Kulturunterschiede bei Mergers & Acquisitions: Entwicklung eines Konzeptes zur Durchführung einer Cultural Due Diligence	2008
93.	Heidorn, Thomas / Kunze, Wolfgang / Schmaltz, Christian Liquiditätsmodellierung von Kreditzusagen (Term Facilities and Revolver)	2008
92.	Burger, Andreas Produktivität und Effizienz in Banken – Terminologie, Methoden und Status quo	2008
91.	Löchel, Horst / Pecher, Florian The Strategic Value of Investments in Chinese Banks by Foreign Financial Insitutions	2008
90.	Schalast, Christoph / Morgenschweis, Bernd / Sprengetter, Hans Otto / Ockens, Klaas / Stachuletz, Rainer / Safran, Robert Der deutsche NPL Markt 2007: Aktuelle Entwicklungen, Verkauf und Bewertung – Berichte und Referate des NPL Forums 2007	2008
89.	Schalast, Christoph / Stralkowski, Ingo 10 Jahre deutsche Buyouts	2008
88.	Bannier, Christina E./ Hirsch, Christian The Economics of Rating Watchlists: Evidence from Rating Changes	2007
87.	Demidova-Menzel, Nadeshda / Heidorn, Thomas Gold in the Investment Portfolio	2007
86.	Hölscher, Luise / Rosenthal, Johannes Leistungsmessung der Internen Revision	2007
85.	Bannier, Christina / Hänsel, Dennis Determinants of banks' engagement in loan securitization	2007

84.	Bannier, Christina "Smoothing" versus "Timeliness" - Wann sind stabile Ratings optimal und welche Anforderungen sind an optimale Berichtsregeln zu stellen?	2007
83.	Bannier, Christina E. Heterogeneous Multiple Bank Financing: Does it Reduce Inefficient Credit-Renegotiation Incidences?	2007
82.	Cremers, Heinz / Löhr, Andreas Deskription und Bewertung strukturierter Produkte unter besonderer Berücksichtigung verschiedener Marktszenarien	2007
81.	Demidova-Menzel, Nadeshda / Heidorn, Thomas Commodities in Asset Management	2007
80.	Cremers, Heinz / Walzner, Jens Risikosteuerung mit Kreditderivaten unter besonderer Berücksichtigung von Credit Default Swaps	2007
79.	Cremers, Heinz / Traugber, Patrick Handlungsalternativen einer Genossenschaftsbank im Investmentprozess unter Berücksichtigung der Risikotragfähigkeit	2007
78.	Gerdemeier, Dieter / Roffia, Barbara Monetary Analysis: A VAR Perspective	2007
77.	Heidorn, Thomas / Kaiser, Dieter G. / Muschiol, Andrea Portfoliooptimierung mit Hedgefonds unter Berücksichtigung höherer Momente der Verteilung	2007
76.	Jobe, Clemens J. / Ockens, Klaas / Safran, Robert / Schalast, Christoph Work-Out und Servicing von notleidenden Krediten – Berichte und Referate des HfB-NPL Servicing Forums 2006	2006
75.	Abrar, Kamyar / Schalast, Christoph Fusionskontrolle in dynamischen Netzsektoren am Beispiel des Breitbandkabelsektors	2006
74.	Schalast, Christoph / Schanz, Kay-Michael Wertpapierprospekte: Markteinführungspublizität nach EU-Prospektverordnung und Wertpapierprospektgesetz 2005	2006
73.	Dickler, Robert A. / Schalast, Christoph Distressed Debt in Germany: What's Next? Possible Innovative Exit Strategies	2006
72.	Belke, Ansgar / Polleit, Thorsten How the ECB and the US Fed set interest rates	2006
71.	Heidorn, Thomas / Hoppe, Christian / Kaiser, Dieter G. Heterogenität von Hedgefondsindizes	2006
70.	Baumann, Stefan / Löchel, Horst The Endogeneity Approach of the Theory of Optimum Currency Areas - What does it mean for ASEAN + 3?	2006
69.	Heidorn, Thomas / Trautmann, Alexandra Niederschlagsderivate	2005
68.	Heidorn, Thomas / Hoppe, Christian / Kaiser, Dieter G. Möglichkeiten der Strukturierung von Hedgefondsportfolios	2005
67.	Belke, Ansgar / Polleit, Thorsten (How) Do Stock Market Returns React to Monetary Policy ? An ARDL Cointegration Analysis for Germany	2005
66.	Daynes, Christian / Schalast, Christoph Aktuelle Rechtsfragen des Bank- und Kapitalmarktsrechts II: Distressed Debt - Investing in Deutschland	2005
65.	Gerdemeier, Dieter / Polleit, Thorsten Measures of excess liquidity	2005
64.	Becker, Gernot M. / Harding, Perham / Hölscher, Luise Financing the Embedded Value of Life Insurance Portfolios	2005
63.	Schalast, Christoph Modernisierung der Wasserwirtschaft im Spannungsfeld von Umweltschutz und Wettbewerb – Braucht Deutschland eine Rechtsgrundlage für die Vergabe von Wasserversorgungskonzessionen? –	2005
62.	Bayer, Marcus / Cremers, Heinz / Kluß, Norbert Wertsicherungsstrategien für das Asset Management	2005
61.	Löchel, Horst / Polleit, Thorsten A case for money in the ECB monetary policy strategy	2005
60.	Richard, Jörg / Schalast, Christoph / Schanz, Kay-Michael Unternehmen im Prime Standard - „Staying Public“ oder „Going Private“? - Nutzenanalyse der Börsennotiz -	2004
59.	Heun, Michael / Schlink, Torsten Early Warning Systems of Financial Crises - Implementation of a currency crisis model for Uganda	2004
58.	Heimer, Thomas / Köhler, Thomas Auswirkungen des Basel II Akkords auf österreichische KMU	2004
57.	Heidorn, Thomas / Meyer, Bernd / Pietrowiak, Alexander Performanceeffekte nach Directors' Dealings in Deutschland, Italien und den Niederlanden	2004

56.	Gerdemeier, Dieter / Roffia, Barbara The Relevance of real-time data in estimating reaction functions for the euro area	2004
55.	Barthel, Erich / Gierig, Rauno / Kühn, Ilmhart-Wolfram Unterschiedliche Ansätze zur Messung des Humankapitals	2004
54.	Anders, Dietmar / Binder, Andreas / Hesdahl, Ralf / Schalast, Christoph / Thöne, Thomas Aktuelle Rechtsfragen des Bank- und Kapitalmarktrechts I : Non-Performing-Loans / Faule Kredite - Handel, Work-Out, Outsourcing und Securitisation	2004
53.	Polleit, Thorsten The Slowdown in German Bank Lending – Revisited	2004
52.	Heidorn, Thomas / Siragusano, Tindaro Die Anwendbarkeit der Behavioral Finance im Devisenmarkt	2004
51.	Schütze, Daniel / Schalast, Christoph (Hrsg.) Wider die Verschleuderung von Unternehmen durch Pfandversteigerung	2004
50.	Gerhold, Mirko / Heidorn, Thomas Investitionen und Emissionen von Convertible Bonds (Wandelanleihen)	2004
49.	Chevalier, Pierre / Heidorn, Thomas / Krieger, Christian Temperaturderivate zur strategischen Absicherung von Beschaffungs- und Absatzrisiken	2003
48.	Becker, Gernot M. / Seeger, Norbert Internationale Cash Flow-Rechnungen aus Eigner- und Gläubigersicht	2003
47.	Boenkost, Wolfram / Schmidt, Wolfgang M. Notes on convexity and quanto adjustments for interest rates and related options	2003
46.	Hess, Dieter Determinants of the relative price impact of unanticipated Information in U.S. macroeconomic releases	2003
45.	Cremers, Heinz / Kluß, Norbert / König, Markus Incentive Fees. Erfolgsabhängige Vergütungsmodelle deutscher Publikumsfonds	2003
44.	Heidorn, Thomas / König, Lars Investitionen in Collateralized Debt Obligations	2003
43.	Kahlert, Holger / Seeger, Norbert Bilanzierung von Unternehmenszusammenschlüssen nach US-GAAP	2003
42.	Beiträge von Studierenden des Studiengangs BBA 012 unter Begleitung von Prof. Dr. Norbert Seeger Rechnungslegung im Umbruch - HGB-Bilanzierung im Wettbewerb mit den internationalen Standards nach IAS und US-GAAP	2003
41.	Overbeck, Ludger / Schmidt, Wolfgang Modeling Default Dependence with Threshold Models	2003
40.	Balthasar, Daniel / Cremers, Heinz / Schmidt, Michael Portfoliooptimierung mit Hedge Fonds unter besonderer Berücksichtigung der Risikokomponente	2002
39.	Heidorn, Thomas / Kantwill, Jens Eine empirische Analyse der Spreadunterschiede von Festsatzanleihen zu Floatern im Euroraum und deren Zusammenhang zum Preis eines Credit Default Swaps	2002
38.	Böttcher, Henner / Seeger, Norbert Bilanzierung von Finanzderivaten nach HGB, EstG, IAS und US-GAAP	2003
37.	Moormann, Jürgen Terminologie und Glossar der Bankinformatik	2002
36.	Heidorn, Thomas Bewertung von Kreditprodukten und Credit Default Swaps	2001
35.	Heidorn, Thomas / Weier, Sven Einführung in die fundamentale Aktienanalyse	2001
34.	Seeger, Norbert International Accounting Standards (IAS)	2001
33.	Moormann, Jürgen / Stehling, Frank Strategic Positioning of E-Commerce Business Models in the Portfolio of Corporate Banking	2001
32.	Sokolovsky, Zbynek / Strohhecker, Jürgen Fit für den Euro, Simulationsbasierte Euro-Maßnahmenplanung für Dresdner-Bank-Geschäftsstellen	2001
31.	Roßbach, Peter Behavioral Finance - Eine Alternative zur vorherrschenden Kapitalmarkttheorie?	2001
30.	Heidorn, Thomas / Jaster, Oliver / Willeitner, Ulrich Event Risk Covenants	2001

29.	Biswas, Rita / Löchel, Horst Recent Trends in U.S. and German Banking: Convergence or Divergence?	2001
28.	Eberle, Günter Georg / Löchel, Horst Die Auswirkungen des Übergangs zum Kapitaldeckungsverfahren in der Rentenversicherung auf die Kapitalmärkte	2001
27.	Heidorn, Thomas / Klein, Hans-Dieter / Siebrecht, Frank Economic Value Added zur Prognose der Performance europäischer Aktien	2000
26.	Cremers, Heinz Konvergenz der binomialen Optionspreismodelle gegen das Modell von Black/Scholes/Merton	2000
25.	Löchel, Horst Die ökonomischen Dimensionen der ‚New Economy‘	2000
24.	Frank, Axel / Moormann, Jürgen Grenzen des Outsourcing: Eine Exploration am Beispiel von Direktbanken	2000
23.	Heidorn, Thomas / Schmidt, Peter / Seiler, Stefan Neue Möglichkeiten durch die Namensaktie	2000
22.	Böger, Andreas / Heidorn, Thomas / Graf Waldstein, Philipp Hybrides Kernkapital für Kreditinstitute	2000
21.	Heidorn, Thomas Entscheidungsorientierte Mindestmargenkalkulation	2000
20.	Wolf, Birgit Die Eigenmittelkonzeption des § 10 KWG	2000
19.	Cremers, Heinz / Robé, Sophie / Thiele, Dirk Beta als Risikomaß - Eine Untersuchung am europäischen Aktienmarkt	2000
18.	Cremers, Heinz Optionspreisbestimmung	1999
17.	Cremers, Heinz Value at Risk-Konzepte für Marktrisiken	1999
16.	Chevalier, Pierre / Heidorn, Thomas / Rütze, Merle Gründung einer deutschen Strombörse für Elektrizitätsderivate	1999
15.	Deister, Daniel / Ehrlicher, Sven / Heidorn, Thomas CatBonds	1999
14.	Jochum, Eduard Hoshin Kanri / Management by Policy (MbP)	1999
13.	Heidorn, Thomas Kreditderivate	1999
12.	Heidorn, Thomas Kreditrisiko (CreditMetrics)	1999
11.	Moormann, Jürgen Terminologie und Glossar der Bankinformatik	1999
10.	Löchel, Horst The EMU and the Theory of Optimum Currency Areas	1998
09.	Löchel, Horst Die Geldpolitik im Währungsraum des Euro	1998
08.	Heidorn, Thomas / Hund, Jürgen Die Umstellung auf die Stückaktie für deutsche Aktiengesellschaften	1998
07.	Moormann, Jürgen Stand und Perspektiven der Informationsverarbeitung in Banken	1998
06.	Heidorn, Thomas / Schmidt, Wolfgang LIBOR in Arrears	1998
05.	Jahresbericht 1997	1998
04.	Ecker, Thomas / Moormann, Jürgen Die Bank als Betreiberin einer elektronischen Shopping-Mall	1997
03.	Jahresbericht 1996	1997
02.	Cremers, Heinz / Schwarz, Willi Interpolation of Discount Factors	1996
01.	Moormann, Jürgen Lean Reporting und Führungsinformationssysteme bei deutschen Finanzdienstleistern	1995

FRANKFURT SCHOOL / HFB – WORKING PAPER SERIES
CENTRE FOR PRACTICAL QUANTITATIVE FINANCE

No.	Author/Title	Year
33.	Schmidt, Wolfgang Das Geschäft mit Derivaten und strukturierten Produkten: Welche Rolle spielt die Bank?	2012
32.	Hübsch, Arnd / Walther, Ursula The impact of network inhomogeneities on contagion and system stability	2012
31.	Scholz, Peter Size Matters! How Position Sizing Determines Risk and Return of Technical Timing Strategies	2012
30.	Detering, Nils / Zhou, Qixiang / Wystup, Uwe Volatilität als Investment. Diversifikationseigenschaften von Volatilitätsstrategien	2012
29.	Scholz, Peter / Walther, Ursula The Trend is not Your Friend! Why Empirical Timing Success is Determined by the Underlying's Price Characteristics and Market Efficiency is Irrelevant	2011
28.	Beyna, Ingo / Wystup, Uwe Characteristic Functions in the Cheyette Interest Rate Model	2011
27.	Detering, Nils / Weber, Andreas / Wystup, Uwe Return distributions of equity-linked retirement plans	2010
26.	Veiga, Carlos / Wystup, Uwe Ratings of Structured Products and Issuers' Commitments	2010
25.	Beyna, Ingo / Wystup, Uwe On the Calibration of the Cheyette. Interest Rate Model	2010
24.	Scholz, Peter / Walther, Ursula Investment Certificates under German Taxation. Benefit or Burden for Structured Products' Performance	2010
23.	Esquivel, Manuel L. / Veiga, Carlos / Wystup, Uwe Unifying Exotic Option Closed Formulas	2010
22.	Packham, Natalie / Schlögl, Lutz / Schmidt, Wolfgang M. Credit gap risk in a first passage time model with jumps	2009
21.	Packham, Natalie / Schlögl, Lutz / Schmidt, Wolfgang M. Credit dynamics in a first passage time model with jumps	2009
20.	Reiswich, Dimitri / Wystup, Uwe FX Volatility Smile Construction	2009
19.	Reiswich, Dimitri / Tompkins, Robert Potential PCA Interpretation Problems for Volatility Smile Dynamics	2009
18.	Keller-Ressel, Martin / Kilin, Fiodar Forward-Start Options in the Barndorff-Nielsen-Shephard Model	2008
17.	Griebsch, Susanne / Wystup, Uwe On the Valuation of Fader and Discrete Barrier Options in Heston's Stochastic Volatility Model	2008
16.	Veiga, Carlos / Wystup, Uwe Closed Formula for Options with Discrete Dividends and its Derivatives	2008
15.	Packham, Natalie / Schmidt, Wolfgang Latin hypercube sampling with dependence and applications in finance	2008
14.	Hakala, Jürgen / Wystup, Uwe FX Basket Options	2008
13.	Weber, Andreas / Wystup, Uwe Vergleich von Anlagestrategien bei Riesterrenten ohne Berücksichtigung von Gebühren. Eine Simulationsstudie zur Verteilung der Renditen	2008
12.	Weber, Andreas / Wystup, Uwe Riesterrente im Vergleich. Eine Simulationsstudie zur Verteilung der Renditen	2008
11.	Wystup, Uwe Vanna-Volga Pricing	2008
10.	Wystup, Uwe Foreign Exchange Quanto Options	2008

09.	Wystup, Uwe Foreign Exchange Symmetries	2008
08.	Becker, Christoph / Wystup, Uwe Was kostet eine Garantie? Ein statistischer Vergleich der Rendite von langfristigen Anlagen	2008
07.	Schmidt, Wolfgang Default Swaps and Hedging Credit Baskets	2007
06.	Kilin, Fiodar Accelerating the Calibration of Stochastic Volatility Models	2007
05.	Gribsch, Susanne/ Kühn, Christoph / Wystup, Uwe Instalment Options: A Closed-Form Solution and the Limiting Case	2007
04.	Boenkost, Wolfram / Schmidt, Wolfgang M. Interest Rate Convexity and the Volatility Smile	2006
03.	Becker, Christoph/ Wystup, Uwe On the Cost of Delayed Currency Fixing Announcements	2005
02.	Boenkost, Wolfram / Schmidt, Wolfgang M. Cross currency swap valuation	2004
01.	Wallner, Christian / Wystup, Uwe Efficient Computation of Option Price Sensitivities for Options of American Style	2004

HfB – SONDERARBEITSBERICHTE DER HfB - BUSINESS SCHOOL OF FINANCE & MANAGEMENT

No.	Author/Title	Year
01.	Nicole Kahmer / Jürgen Moormann Studie zur Ausrichtung von Banken an Kundenprozessen am Beispiel des Internet (Preis: € 120,-)	2003

Printed edition: € 25.00 + € 2.50 shipping

Download:

Working Paper: http://www.frankfurt-school.de/content/de/research/publications/list_of_publication/list_of_publication
CPQF: http://www.frankfurt-school.de/content/de/cpqf/research_publications.html

Order address / contact

Frankfurt School of Finance & Management
Sonnemannstr. 9–11 ▪ D–60314 Frankfurt/M. ▪ Germany
Phone: +49 (0) 69 154 008–734 ▪ Fax: +49 (0) 69 154 008–728
eMail: e.lahdensuu@fs.de
Further information about Frankfurt School of Finance & Management
may be obtained at: <http://www.fs.de>