



TECHNISCHE
UNIVERSITÄT
WIEN

DIPLOMARBEIT

Modellierung und Absicherung von Langlebigkeitsrisiken: Ein Ansatz auf Basis der Skew-t-Verteilung

zur Erlangung des akademischen Grades

Diplom-Ingenieur/in
im Rahmen des Studiums

Finanz- und Versicherungsmathematik

eingereicht von

Robert Melcher, BSc
Matrikelnummer 11778952

ausgeführt am Institut für Stochastik und Wirtschaftsmathematik
der Technischen Universität Wien

unter der Anleitung von
Univ.Prof. Dipl.-Ing. Dr.techn. Stefan Gerhold

Wien, 27. Januar 2025

(Betreuer)

(Verfasser)

Kurzfassung

Die vorliegende Diplomarbeit untersucht Ansätze zur Absicherung von Langlebighkeitsrisiken, die insbesondere für Versicherungen und Pensionsplaner zunehmend an Bedeutung gewinnen. Die demografische Entwicklung hin zu einer alternden Bevölkerung erschwert präzise Prognosen erheblich. So müssen bei der Kalkulation einer Risikolebensversicherung für eine 30-jährige Person beispielsweise Mortalitätsraten über mehrere Jahrzehnte hinweg antizipiert werden. Um das Langlebighkeitsrisiko zu reduzieren, wird der Einsatz von Derivaten erörtert, die an einen Mortalitätsindex gekoppelt sind.

Bisherige Ansätze beruhen häufig auf der Annahme einer Normalverteilung, da diese mathematisch handhabbar und einfach zu implementieren ist. Die vorliegende Arbeit hinterfragt diese Annahme und untersucht alternative Modelle, die zusätzlich zur Varianz auch höhere Momente wie Schiefe (drittes Moment) und Kurtosis (viertes Moment) berücksichtigen. Es zeigt sich, dass diese erweiterten Ansätze zu präziseren Ergebnissen führen können, wodurch eine verbesserte Absicherung der Langlebighkeitsrisiken ermöglicht wird.

Abstract

This master's thesis explores approaches to hedging longevity risk, an issue of growing importance for insurers and pension planners. The demographic shift towards an aging population significantly complicates accurate forecasting. For example, when calculating a term life insurance policy for a 30-year-old individual, mortality rates must be projected decades into the future. To mitigate longevity risk, the use of derivatives linked to a mortality index is examined.

Traditional approaches often assume a normal distribution due to its mathematical simplicity and ease of application. This thesis challenges this assumption by investigating alternative models that incorporate higher moments, such as the third moment (skewness) and the fourth moment (kurtosis), in addition to variance. The findings suggest that these advanced models can yield more precise results, thereby enabling a more effective hedging of longevity risks.

Danksagung

An dieser Stelle möchte ich meinen herzlichen Dank all jenen aussprechen, die mich bei der Erstellung dieser Masterarbeit unterstützt und motiviert haben. Ohne ihre Hilfe, ihren Zuspruch und ihre Unterstützung wäre diese Arbeit in dieser Form nicht möglich gewesen.

Mein besonderer Dank gilt Univ.-Prof. Dipl.-Ing. Dr.techn. Stefan Gerhold, der mich durch seine Expertise und wertvollen Ratschläge maßgeblich bei der Ausarbeitung dieser Arbeit unterstützt hat. Seine Geduld und sein Engagement haben mich durch die Herausforderungen dieses Projekts begleitet.

Ebenso möchte ich meinen Studienkolleginnen und -kollegen danken, die mir während dieser Zeit stets eine moralische Stütze waren. Auch meinen außeruniversitären Freundinnen und Freunden, die mir durch ihre Gespräche und aufmunternden Worte immer zur Seite standen, gilt mein großer Dank.

Zuletzt richte ich meinen tiefsten Dank an meine Familie, die es mir ermöglicht hat, den Weg des Studiums zu beschreiten, und mich in allen Lebenslagen bedingungslos unterstützt. Ihre Zuversicht und ihr Vertrauen haben mich stets bestärkt, meine Ziele zu verfolgen.

Eidesstattliche Erklärung

Ich erkläre an Eides statt, dass ich die vorliegende Arbeit selbstständig und ohne fremde Hilfe verfasst, andere als die angegebenen Quellen und Hilfsmittel nicht benutzt bzw. die wörtlich oder sinngemäß entnommenen Stellen als solche kenntlich gemacht habe. Zudem bestätige ich, dass keine künstliche Intelligenz (KI) für die Verfassung der Arbeit bzw. für Teile der Arbeit zum Einsatz gekommen ist.

Wien, am 27. Jänner 2025



.....
Robert Melcher

Einverständniserklärung zur Plagiatsprüfung

Ich nehme zur Kenntnis, dass die vorgelegte Arbeit mit geeigneten und dem derzeitigen Stand der Technik entsprechenden Mitteln (Plagiat-Erkennungssoftware) elektronisch-technisch überprüft wird. Dies stellt einerseits sicher, dass bei der Erstellung der vorgelegten Arbeit die hohen Qualitätsvorgaben im Rahmen der ausgegebenen der an der TU Wien geltenden Regeln zur Sicherung guter wissenschaftlicher Praxis - „Code of Conduct“ (Mitteilungsblatt 2007, 26. Stück, Nr. 257 idgF.) an der TU Wien eingehalten wurden. Zum anderen werden durch einen Abgleich mit anderen studentischen Abschlussarbeiten Verletzungen meines persönlichen Urheberrechts vermieden.

Wien, am 27. Jänner 2025



.....
Robert Melcher

Inhaltsverzeichnis

1	Einleitung	1
2	Mortalitätsmodelle	3
2.1	Lee-Carter Modell	4
2.1.1	Das Modell	5
2.1.2	Parameterschätzung	6
2.2	Cairns-Blake-Dowd Modell	7
2.2.1	Das Modell	8
2.2.2	Parameterschätzung	8
3	Modellannahmen	10
3.1	Modellauswahl	10
3.2	Zustandsraum	10
3.3	Auswahl der Verteilung	12
3.4	Hedged liability	14
3.5	q-Forwards	15
3.5.1	LifeMetrics Index	18
3.5.2	Bepreisung von q-Forwards	20
4	Optimierungsproblem	22
4.1	Analyse der Momente	22
4.2	Polynomial goal programming	26
5	Analytische Herleitung der Ausdrücke	30
5.1	Approximierte Ausdrücke	32
5.2	Momente der Skew-t Verteilung	37
5.3	Analytische Herleitung der Momente M_t, V_t, S_t und K_t	41
6	Numerisches Beispiel	44
6.1	Dataset	44
6.2	Auswahl des Mortalitätsmodelles	44
6.3	Parameter	47
6.4	Überprüfung der Verteilungsannahmen	48
6.5	Alternative Verteilungsannahme: Skew-Normalverteilung	52
6.6	Generelle Annahmen zur Modellierung	55
6.7	Berechnung der Momente	56
6.8	Optimierungsproblem	59
6.9	Ergebnisse	60
7	Conclusio	62
	Literatur	63

Abbildungsverzeichnis

4.1	Dichtefunktionen zweier skew-t Verteilungen mit verschiedenen Schiefen	24
4.2	99,5 % Perzentil der jeweiligen Verteilungen	24
4.3	80 % Perzentil der jeweiligen Verteilungen	24
4.4	Dichtefunktionen zweier skew-t Verteilungen mit verschiedenen Kurtoses	25
4.5	99,5 % Perzentil der jeweiligen Verteilungen	26
4.6	80 % Perzentil der jeweiligen Verteilungen	26
6.1	Schätzung von $\kappa_t^{(1)}$ basierend auf den österreichischen Daten	47
6.2	Schätzung von $\kappa_t^{(2)}$ basierend auf den österreichischen Daten	48

Tabellenverzeichnis

1	Term Sheet eines einzelnen q-Forwards zur Absicherung des Langlebigkeitsrisikos.	19
2	q-Forward Vergleiche für verschiedene Ergebnisse des realisierten Referenzzinssatzes	20
3	Ergebnisse verschiedener statistischer Methoden anhand der Daten aus Österreich	46
4	Vergleich der Schätzungen der Parameter der Skew-t-Verteilung	52
5	Parameterschätzungen der Skew-Normalverteilung für die Daten aus Österreich.	54
6	Parameterschätzungen der Skew-Normalverteilung für die Daten aus Norwegen.	54
7	Parameterschätzungen der Skew-Normalverteilung für die Daten aus Kanada.	54
8	Details der vier q-Forwards in unserem Portfolio	55
9	Werte von $D_k(0, t_0 + 1)$ und $D_k(j, t_0 + 1)$	58
10	Ergebnisse des Optimierungsproblems für verschiedene Risikopräferenzen.	60

1 Einleitung

Die stetige Zunahme der Lebenserwartung stellt sowohl die Gesellschaft als auch die Versicherungs- und Pensionsindustrie vor immense Herausforderungen. Langlebighkeitsrisiken, das Risiko, dass eine Bevölkerung im Durchschnitt länger (oder auch kürzer) lebt als ursprünglich prognostiziert, beeinflussen maßgeblich die finanzielle Stabilität von Renten- und Lebensversicherungsprodukten. Ein präzises Verständnis und eine realistische Modellierung der Mortalitätsdynamiken sind daher essenziell, um die langfristige Absicherung dieser Risiken zu gewährleisten. (vgl. Biffis & Blake, 2009)

Traditionell basieren aktuariell-statistische Modelle zur Prognose von Mortalitätsveränderungen häufig auf der Annahme einer Normalverteilung. Diese Annahme kann jedoch insbesondere bei extremen Ereignissen oder asymmetrischen Verteilungen der Sterblichkeitsveränderungen zu unzureichenden oder verzerrten Ergebnissen führen. Die Nutzung einer Skew-t Verteilung, die sowohl asymmetrische als auch schwerere Ränder (Fat Tails) abbilden kann, bietet hier ein vielversprechendes alternatives Modellierungsinstrument. (vgl. Peiro, 1999)

Bis 2008 wurde das Langlebighkeitsrisiko überwiegend durch klassische Rückversicherungen abgesichert. Seitdem hat sich jedoch eine innovative Alternative über den Kapitalmarkt etabliert. Einen wesentlichen Impuls gab die Veröffentlichung des LifeMetrics-Index durch J.P. Morgan im Jahr 2007, der 2008 erstmals praktisch genutzt wurde. Als Vorreiter übertrug das britische Versicherungsunternehmen Lucida als erstes einen Teil seines Langlebighkeitsrisikos auf J.P. Morgan.

Die Grundlage für diese Transaktion bildete ein sogenannter q-Forward, ein Finanzinstrument in Form eines Swaps. Dabei wird die tatsächlich beobachtete Mortalitätsrate gegen eine im Voraus festgelegte Rate getauscht. Der Buchstabe „q“ steht hier für die in der Versicherungsmathematik gebräuchliche Sterblichkeitswahrscheinlichkeit. Präzise bezeichnet $q(x)$ die Wahrscheinlichkeit, dass eine Person im Alter von x innerhalb eines Jahres verstirbt. (vgl. Liu & Li, 2018)

Ziel dieser Arbeit ist es, ein aktuarielles Modell zur Absicherung von Langlebighkeitsrisiken zu entwickeln und zu untersuchen, bei dem die Veränderung der Sterblichkeitswahrscheinlichkeiten durch eine Skew-t Verteilung beschrieben wird. Neben der theoretischen Fundierung steht die Analyse der praktischen Anwendbarkeit und der Verbesserung der Modellgenauigkeit im Fokus.

Die Relevanz dieser Forschung ergibt sich nicht nur aus den wachsenden Heraus-

forderungen einer alternden Bevölkerung, sondern auch aus der Notwendigkeit, bestehende Modellierungsansätze kontinuierlich zu verbessern. Durch die Kombination von theoretischen und praktischen Perspektiven soll diese Arbeit dazu beitragen, ein tieferes Verständnis der Sterblichkeitsdynamiken zu schaffen und gleichzeitig innovative Ansätze für die Langlebigkeitsabsicherung bereitzustellen.

Der Aufbau dieser Arbeit ist wie folgt gegliedert: Im zweiten Kapitel werden die für das Modell relevanten Mortalitätsmodelle vorgestellt und beschrieben. Anschließend wird im dritten Kapitel das zentrale Modell detailliert erläutert, wobei ein besonderer Fokus auf die zugrunde liegenden Annahmen und die entwickelte Absicherungsstrategie gelegt wird.

Das vierte Kapitel widmet sich der Formulierung des zugrunde liegenden Optimierungsproblems. Darauf aufbauend wird im fünften Kapitel der mathematische Hintergrund beleuchtet, einschließlich der Herleitung und Approximation der für das Modell erforderlichen Ausdrücke.

Im sechsten Kapitel werden die entwickelten Ansätze anhand realistischer Datensätze veranschaulicht. Hierbei werden die Ergebnisse präsentiert und analysiert. Abschließend fasst das siebte Kapitel die wesentlichen Erkenntnisse der Arbeit zusammen und gibt einen Ausblick auf mögliche zukünftige Entwicklungen.

2 Mortalitätsmodelle

Angesichts der steigenden Lebenserwartung und der zunehmenden Alterung der Bevölkerung wurden insbesondere in den letzten Jahrzehnten sogenannte Mortalitätsmodelle entwickelt, um präzise Prognosen zur zukünftigen Entwicklung der Sterblichkeit zu erstellen. Die Mortalität selbst dient dabei als Maß für die Anzahl der Todesfälle in einer definierten Population.

Die Entwicklung dieser Modelle erfolgt überwiegend durch Fachleute aus der Versicherungsmathematik, Statistik und Demografie. Sie bilden die Grundlage für eine fundierte Analyse und ermöglichen es, die Auswirkungen demografischer Veränderungen auf soziale und finanzielle Systeme zu bewerten. (vgl. Hunt & Blake, 2020)

Obwohl die Entwicklung von Mortalitätsmodellen insbesondere in den letzten Jahrzehnten stark an Bedeutung gewonnen hat, reicht ihre Geschichte weit zurück. Bereits im Jahr 1825 veröffentlichte der britische Versicherungsmathematiker Benjamin Gompertz ein Sterblichkeitsgesetz, das als Grundlage für die ersten Mortalitätsmodelle diente. (vgl. Kirkwood, 2015)

Die Relevanz solcher Modelle hat seither stetig zugenommen, insbesondere vor dem Hintergrund der kontinuierlich steigenden Lebenserwartung in den Industrieländern seit Beginn des 19. Jahrhunderts. Dieser Anstieg, der seit dem Ende des Zweiten Weltkriegs besonders konstant ist, erfolgt häufig schneller als ursprünglich prognostiziert und unterstreicht die Notwendigkeit präziser Prognosemodelle. (vgl. Booth & Tickle, 2008)

Im Vergleich zur heutigen Praxis basierten frühere Mortalitätsmodelle häufig auf eher simplen Annahmen. Heutzutage liegt der Fokus verstärkt auf dem Einsatz moderner statistischer Methoden und insbesondere auf der fundierten Auswahl geeigneter Verteilungen.

Diese Weiterentwicklung bringt jedoch auch einen höheren Aufwand mit sich, da die zugrunde liegenden Verteilungen nicht nur sorgfältig ausgewählt, sondern ihre Parameter zudem präzise geschätzt werden müssen, um valide Ergebnisse zu gewährleisten.

Die Parameter von Mortalitätsmodellen werden in der Regel auf Basis historischer Daten geschätzt. Dabei erfolgt eine Zerlegung der Sterblichkeitsraten nach Alter, Periode und Kohorte. Die Sterblichkeitsrate beschreibt die Wahrscheinlichkeit, dass eine Person innerhalb eines definierten Zeitraums verstirbt. Basierend auf der Analyse dieser historischen Mortalitätsdaten wird ein Modell entwickelt, das die vergangene Entwicklung möglichst genau abbildet. Die geschätzten Parameter dienen

dann als Grundlage für zukünftige Prognosen. In den meisten Modellen erfolgt diese Prognose durch Extrapolation.

Erweiterte Modelle könnten zusätzlich spezifische Todesursachen und deren Abhängigkeiten berücksichtigen, beispielsweise die Verbindung zwischen Lungenkrebs und Tabakkonsum. Ebenso könnten externe Schocks, wie etwa Pandemien, die die Sterblichkeitsrate signifikant beeinflussen, in die Modellierung integriert und simuliert werden. (vgl. Hunt & Blake, 2020)

Bei der Nutzung historischer Daten ist die Wahl der betrachteten Zeit- und Altersspannen entscheidend. Die Zeitspanne sollte ausreichend lang sein, um Trends und Muster zu erfassen, wobei Daten ab 1950 häufig eine sinnvolle Grundlage bilden. Dies liegt daran, dass ältere Daten je nach Land möglicherweise unvollständig sind. Für die Altersspanne liegt der Fokus vor allem auf höheren Altersgruppen, da Langlebighkeitsrisiken insbesondere für Lebensversicherungen und Leibrenten relevant sind. In den betrachteten Beispielen umfasst die Altersgruppe Personen zwischen 60 und 89 Jahren.

Im Folgenden werden zwei stochastische Mortalitätsmodelle näher betrachtet, die im weiteren Verlauf für die Anwendung des in dieser Arbeit beschriebenen Modells herangezogen werden können.

2.1 Lee-Carter Modell

Das Lee-Carter-Modell (LC) wurde erstmals 1992 von dem Wirtschaftswissenschaftler Ronald D. Lee und dem Demografen Lawrence Carter entwickelt. Grundlage war die Analyse von US-Mortalitätsdaten aus dem Zeitraum 1933 bis 1987. Es zählt zu den am häufigsten verwendeten Modellen zur Prognose von Sterblichkeitstrends und wird beispielsweise vom US Bureau of the Census als Referenzmethode eingesetzt. Das Lee-Carter-Modell findet in vielfältigen Anwendungsbereichen Verwendung. Dazu gehört unter anderem die Vorhersage von Mortalitätsreduktionsfaktoren, die eine zentrale Rolle bei der Extrapolation von Sterblichkeitstrends spielen. Ebenso wird es genutzt, um die Angemessenheit von Renteneinkommen zu bewerten.

Zu den wesentlichen Vorteilen des Lee-Carter-Modells zählen die vergleichsweise geringe Anzahl an zu schätzenden Parametern sowie die allgemeine Robustheit des Modells, die es in der Praxis besonders attraktiv macht. (vgl. Li & Chan, 2007)

2.1.1 Das Modell

Das Lee-Carter-Modell wird durch die folgende Gleichung beschrieben:

$$\log(m(x, t)) = a_x + b_x \cdot \kappa_t + \epsilon(x, t).$$

Dabei stellt $\log(m(x, t))$ die logarithmisch transformierte, altersspezifische zentrale Mortalitätsrate im Alter x im Jahr t dar. Die zentrale Mortalitätsrate $m(x, t)$ beschreibt die Wahrscheinlichkeit, dass eine Person, deren genaues Alter unbekannt ist, sich jedoch im Intervall $[x, x + 1)$ während des Zeitraums $[t - 1, t)$ befindet, innerhalb eines Jahres stirbt.

Die logarithmische Transformation der Sterbewahrscheinlichkeit erfolgt aus zwei Gründen: Erstens, um die teils sehr kleinen Werte von $m(x, t)$, insbesondere bei jüngeren Altersgruppen, besser handhabbar zu machen. Zweitens ermöglicht die Transformation eine lineare Struktur des Modells, was die Anwendung statistischer Methoden vereinfacht.

Die rechte Seite der Gleichung setzt sich aus drei Komponenten zusammen:

- a_x : eine zeitunabhängige, altersspezifische Komponente, die die mittlere Mortalitätsrate beschreibt,
- $b_x \cdot \kappa_t$: das Produkt einer altersspezifischen Komponente b_x und einer zeitabhängigen Variable κ_t , dem sogenannten Mortalitätsindex. Hierbei gibt b_x an, wie stark die Mortalitätsrate in jedem Alter variiert, wenn sich κ_t verändert.
- $\epsilon(x, t)$: ein Fehlerterm, der nicht durch das Modell erfasste, altersspezifische Einflüsse repräsentiert.

Das Lee-Carter-Modell bietet durch diese Struktur eine intuitive und zugleich statistisch gut handhabbare Möglichkeit, Mortalitätstrends zu modellieren. (vgl. Renshaw & Haberman, 2006)

Um eine verbesserte Stabilität des Lee-Carter-Modells zu gewährleisten, werden zusätzliche Bedingungen eingeführt:

$$\sum_x b_x^2 = 1 \text{ und } \sum_t \kappa_t = 0.$$

Diese Bedingungen dienen dazu, die Schätzungen der Parameter b_x und κ_t zu stabilisieren und sicherzustellen, dass das Modell eindeutig spezifiziert ist. Die erste Bedingung normiert die altersspezifischen Parameter b_x , während die zweite Bedingung den Mittelwert des Mortalitätsindex κ_t auf Null setzt, was Interpretations- und Berechnungszwecke erleichtert. (vgl. Renshaw & Haberman, 2006)

Diese zusätzlichen Nebenbedingungen können ohne Einschränkung der Allgemeinheit angenommen werden, da $a_x, b \neq 0$ und κ_t entsprechend normalisiert werden können. Die erste Bedingung kann durch die Transformation

$$(\kappa_t, a_x) \mapsto \left(\kappa_t - \frac{1}{t_{\max} - t_{\min} + 1} \sum_{t=t_{\min}}^{t_{\max}} \kappa_t, a_x + \frac{b_x}{t_{\max} - t_{\min} + 1} \sum_{t=t_{\min}}^{t_{\max}} \kappa_t \right)$$

erreicht werden. Hierbei wird κ_t so angepasst, dass dessen Mittelwert über den betrachteten Zeitraum $[t_{\min}, t_{\max}]$ gleich null ist, und a_x entsprechend korrigiert. Die zweite Bedingung wird durch die folgende Transformation realisiert:

$$(b_x, \kappa_t) \mapsto \left(\frac{b_x}{\|b\|_2}, \|b\|_2 \kappa_t \right),$$

Dabei wird b_x auf die Länge 1 normiert, wobei $\|b\|_2$ die euklidische Norm des Vektors b bezeichnet, dessen Einträge b_x die Werte für alle betrachteten Altersgruppen enthalten. Gleichzeitig wird κ_t entsprechend skaliert. (vgl. Kainhofer et al., 2006)

2.1.2 Parameterschätzung

Die Schätzung der Parameter im Lee-Carter-Modell erfolgt mithilfe der Singulärwertzerlegung (Singular Value Decomposition, SVD). In R kann diese durch die Funktion `svd()` aus dem Package `base` berechnet werden. Im ersten Schritt wird a_x , die mittlere Mortalitätsrate, ermittelt. Diese ergibt sich als Mittelwert der logarithmierten Mortalitätsraten über die betrachteten Jahre:

$$a_x = \frac{\sum_{t=1}^T \log(m(x, t))}{T},$$

wobei t die betrachteten Jahre und T das letzte betrachtete Jahr bezeichnet. Im zweiten Schritt wird die Matrix $A_{x,t}$ berechnet, die wie folgt definiert ist:

$$A_{x,t} = \log(m(x, t)) - a_x.$$

Auf die Matrix $A_{x,t}$ wird anschließend die SVD angewendet. Diese zerlegt $A_{x,t}$ in das Produkt einer unitären Matrix $U = (u_{i,j})_{i=1,\dots,x;j=1,\dots,x}$, der adjungierten unitären Matrix $V = (v_{i,j})_{i=1,\dots,t;j=1,\dots,t}$, und einer Diagonalmatrix $S = (s_{i,j})_{i=1,\dots,x;j=1,\dots,t}$, die die Singulärwerte von $A_{x,t}$ enthält, die durch das Hinzufügen von Null-Einträgen auf die Form einer Diagonalmatrix gebracht wird:

$$A_{x,t} = U \cdot S \cdot V.$$

Aus diesen Matrizen lassen sich die Parameter b_x und κ_t wie folgt extrahieren:

$$b_x = (u_{1,1}, u_{2,1}, \dots, u_{x,1}),$$
$$\kappa_t = (v_{1,1}, v_{1,2}, \dots, v_{1,t}).$$

Damit sind alle benötigten Parameter des Modells bestimmt. Um zukünftige Prognosen zu erstellen, wird ein ARMA-Modell (Autoregressive-Moving-Average-Modell) auf den Mortalitätsindex κ_t angewendet. Die Formel für die prognostizierte zukünftige Mortalitätsrate lautet:

$$m(x, t + n) = \exp(a_x + s_{1,1} \cdot \kappa_{t+n} \cdot b_x),$$

wobei κ_{t+n} durch ein ARMA-Modell geschätzt wird, also $\kappa_{t+n} = \text{ARMA}(\kappa_t, n)$. Hierbei ist n eine natürliche Zahl, die angibt, wie weit in die Zukunft die Prognose erfolgt. (vgl. Li & Chan, 2007)

2.2 Cairns-Blake-Dowd Modell

Neben dem Lee-Carter-Modell gehört das Cairns-Blake-Dowd-Modell (CBD) zu den am häufigsten verwendeten Ansätzen zur Prognose von Sterblichkeitsraten. Die in dieser Arbeit verwendete Version des CBD-Modells wurde 2009 von dem Finanzmathematiker Andrew J.G. Cairns sowie den Finanzökonom David Blake und Kevin Dowd entwickelt. Das Modell basiert auf einer linearen Regression zur Vorhersage von Sterblichkeitsraten.

Ein wesentlicher Vorteil des CBD-Modells liegt in seiner Einfachheit und gleichzeitig hohen Robustheit. Allerdings weist es auch einige Einschränkungen auf, wie beispielsweise die geringe Eignung zur Modellierung jüngerer Altersgruppen. Dieser Nachteil ist jedoch für diese Arbeit irrelevant, da lediglich Altersgruppen ab 60 Jahren betrachtet werden, für die das CBD-Modell nachweislich sehr gut geeignet ist.

Ein weiterer Aspekt ist die fehlende Berücksichtigung von Kohorteneffekten im Basis-Modell. Diese könnten zwar durch Erweiterungen des Modells einbezogen werden, doch wird in dieser Arbeit aufgrund der dadurch entstehenden Komplexität darauf verzichtet.

Das CBD-Modell wurde in der Vergangenheit erfolgreich zur Berechnung und Prognose von Sterblichkeitsraten in verschiedenen Ländern wie England und Wales, den Vereinigten Staaten, Spanien und Italien eingesetzt, was seine Vielseitigkeit und praktische Anwendbarkeit unterstreicht. (vgl. Safitri et al., 2019)

2.2.1 Das Modell

Das Cairns-Blake-Dowd-Modell (CBD) beschreibt das Logit der Mortalitätswahrscheinlichkeit und umfasst zwei zeitabhängige Parameter $\kappa_t^{(1)}$ und $\kappa_t^{(2)}$. Diese Parameter werden in vielen Anwendungen als bivariater Random Walk mit konstanter Drift modelliert, um zukünftige Prognosen zu erstellen. Die Schätzung der Parameter erfolgt jedoch mithilfe der Methode der kleinsten Quadrate (Least Squares Method).

Wie bereits erwähnt, existieren mehrere Varianten des CBD-Modells. Das in dieser Arbeit verwendete Modell wird durch die folgende Gleichung definiert:

$$\text{logit}(q(x, t)) = \log\left(\frac{q(x, t)}{1 - q(x, t)}\right) = \kappa_t^{(1)} + \kappa_t^{(2)} \cdot (x - \bar{x}) + \epsilon(x, t),$$

Dabei bezeichnet $q(x, t)$ die Sterblichkeitswahrscheinlichkeit, also die Wahrscheinlichkeit, dass eine Person im Alter x im Zeitintervall $[t - 1, t)$ verstirbt. Da $q(x, t)$ als Wahrscheinlichkeit zwischen Null und Eins liegen muss, wird eine Logit-Transformation angewandt. Die Parameter $\kappa_t^{(1)}$ und $\kappa_t^{(2)}$ sind die beiden zeitabhängigen Komponenten des Modells, die die Entwicklung der Mortalitätsraten beschreiben. \bar{x} repräsentiert das Durchschnittsalter in der betrachteten Stichprobe, während $\epsilon(x, t)$ einen Fehlerterm darstellt, der nicht durch das Modell erklärte Einflüsse erfasst, ähnlich wie beim Lee-Carter-Modell. Für die weiteren Betrachtungen wird x verwendet, um die betrachtete Altersgruppe ($x = x_a, x_{a+1}, \dots, x_b$) darzustellen, und t zur Kennzeichnung des betrachteten Zeitraums ($t = t_a, t_{a+1}, \dots, t_b$). (vgl. Siu et al., 2023)

2.2.2 Parameterschätzung

Die Schätzung der beiden zeitabhängigen Parameter im Cairns-Blake-Dowd-Modell erfolgt mithilfe der Methode der kleinsten Quadrate (Least Squares Method). Ziel dieser Methode ist es, die Summe der quadrierten Residuen zu minimieren. Für das CBD-Modell ergibt sich die Zielfunktion wie folgt:

$$J = \sum_{x=x_a}^{x_b} \sum_{t=t_a}^{t_b} \left(\log\left(\frac{q(x, t)}{1 - q(x, t)}\right) - \kappa_t^{(1)} - \kappa_t^{(2)} \cdot (x - \bar{x}) \right)^2.$$

Zur Minimierung von J werden die partiellen Ableitungen nach $\kappa_t^{(1)}$ und $\kappa_t^{(2)}$ gebildet

und gleich Null gesetzt:

$$\frac{\partial J}{\partial \kappa_t^{(1)}} = \sum_{x=x_a}^{x_b} 2 \cdot \left(\log\left(\frac{q(x,t)}{1-q(x,t)}\right) - \kappa_t^{(1)} - \kappa_t^{(2)} \cdot (x - \bar{x}) \right) = 0,$$

$$\frac{\partial J}{\partial \kappa_t^{(2)}} = \sum_{x=x_a}^{x_b} 2 \cdot (x - \bar{x}) \cdot \left(\log\left(\frac{q(x,t)}{1-q(x,t)}\right) - \kappa_t^{(1)} - \kappa_t^{(2)} \cdot (x - \bar{x}) \right) = 0.$$

Durch Umstellen dieser Gleichungen ergeben sich die geschätzten Werte für $\kappa_t^{(1)}$ und $\kappa_t^{(2)}$:

$$\hat{\kappa}_t^{(1)} = \frac{\sum_{x=x_a}^{x_b} \left(\log\left(\frac{q(x,t)}{1-q(x,t)}\right) - \kappa_t^{(2)} \cdot \sum_{x=x_a}^{x_b} (x - \bar{x}) \right)}{(b - a)},$$

$$\hat{\kappa}_t^{(2)} = \frac{(b - a) \cdot \sum_{x=x_a}^{x_b} (x - \bar{x}) \cdot \log\left(\frac{q(x,t)}{1-q(x,t)}\right) - \sum_{x=x_a}^{x_b} \log\left(\frac{q(x,t)}{1-q(x,t)}\right) \cdot \sum_{x=x_a}^{x_b} (x - \bar{x})}{(b - a) \cdot \sum_{x=x_a}^{x_b} (x - \bar{x})^2 - \left(\sum_{x=x_a}^{x_b} (x - \bar{x}) \right)^2}$$

Um die Güte der geschätzten Werte zu bewerten, kann zusätzlich der Root Mean Square Error (RMSE) berechnet werden. Der RMSE ist definiert als:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2},$$

wobei y_i die tatsächlichen und \hat{y}_i die geschätzten Werte darstellen. Der RMSE gibt an, wie gut die geschätzten Werte mit den tatsächlichen Daten übereinstimmen, und liefert somit eine Maßzahl für die Güte der Anpassung. (vgl. Safitri et al., 2019)

3 Modellannahmen

In diesem Abschnitt wird der Aufbau des in dieser Arbeit entwickelten Modells detailliert beschrieben. Dabei stehen insbesondere drei zentrale Aspekte im Fokus: die Wahl des zugrunde liegenden Mortalitätsmodells, die Auswahl der Verteilung zur Beschreibung der zeitlichen Veränderung der Mortalität sowie die Spezifikation des Hedginginstruments. Diese Elemente bilden die Grundlage für die Herleitung und Anwendung des Modells und werden im Folgenden näher erläutert.

3.1 Modellauswahl

Sowohl das Cairns-Blake-Dowd-Modell als auch das Lee-Carter-Modell liefern zuverlässige Ergebnisse bei der Prognose von Mortalitätsraten. Allerdings zeigen beide Modelle eine hohe Datenabhängigkeit, was die Wahl eines geeigneten Ansatzes beeinflussen kann. Um die Qualität der Modellanpassung zu vergleichen, stehen verschiedene Methoden zur Verfügung, darunter die Berechnung von Fehlermetriken wie dem MAE (Mean Absolute Error), dem RMSE (Root Mean Square Error) und dem MAPE (Mean Absolute Percentage Error). Zudem können Kriterien wie das Akaike Information Criterion (AIC) und das Bayesian Information Criterion (BIC) herangezogen werden, um die Modelle zu bewerten. (vgl. Hastie et al., 2000)

Für die Zwecke dieser Arbeit wird eines der beiden Modelle ausgewählt, um die Vergleichbarkeit der Ergebnisse zwischen verschiedenen Datensätzen sicherzustellen. Die Entscheidung fällt auf das Cairns-Blake-Dowd-Modell, da dieses insbesondere für ältere Altersgruppen, die im Fokus dieser Arbeit stehen, genauere Projektionen der Sterblichkeitsraten liefert, wie es im weiteren Verlauf der Arbeit bei den numerischen Beispielen illustriert wird.

3.2 Zustandsraum

Diskrete stochastische Mortalitätsmodelle, wie sie in den vorherigen Abschnitten analysiert wurden, können auch in einer Zustandsraumdarstellung formuliert werden. Diese Darstellung umfasst zwei wesentliche Gleichungen: die Beobachtungsgleichung und die Übergangsgleichung.

Die Beobachtungsgleichung verknüpft die beobachtbaren Größen mit den modellinternen, zeitabhängigen Parametern. Im Falle des Lee-Carter-Modells ist dies die logarithmierte zentrale Mortalitätsrate $\log(m(x, t))$, während im Cairns-Blake-Dowd-Modell das Logit der Sterblichkeitswahrscheinlichkeit $\log\left(\frac{q(x, t)}{1-q(x, t)}\right)$ verwendet wird.

Die Übergangsgleichung beschreibt die Dynamik und Entwicklung der zeitabhängigen Prozesse über die Zeit hinweg. Sie bildet die Grundlage für die Modellierung der stochastischen Eigenschaften, die eine realistische Prognose zukünftiger Entwicklungen ermöglichen. (vgl. Fung et al., 2017)

Die Zustandsraumdarstellung eines Mortalitätsmodells umfasst die Beschreibung der Beobachtungsgleichung und der Übergangsgleichung, die in ihrer Gesamtheit die Verbindung zwischen beobachteten Daten und den zugrunde liegenden stochastischen Prozessen herstellen.

Beobachtungsgleichung:

Die Beobachtungsgleichung wird wie folgt ausgedrückt:

$$\vec{y}(t) = \vec{d} + B\vec{\alpha}(t) + \vec{\epsilon}(t),$$

wobei die einzelnen Komponenten wie folgt definiert sind:

- $\vec{y}(t)$: Vektor der Beobachtungen für die Altersgruppen $x = x_a, \dots, x_b$,
- \vec{d} : Konstanter Vektor, der entweder die Altersparameter enthält oder ausschließlich aus Nullen besteht.
- $\vec{\alpha}(t)$: Zustandsvektor, der die zeitabhängigen Parameter umfasst, z. B. $\vec{\alpha}(t) = (\kappa_t^{(1)}, \kappa_t^{(2)})'$.
- B : Matrix, die die Beziehung zwischen den Beobachtungen und dem Zustandsvektor definiert.
- $\vec{\epsilon}(t)$: Beobachtungsfehler, der die Abweichung der Modellvorhersagen von den tatsächlichen Beobachtungen zum Zeitpunkt t beschreibt.

Die Beobachtungsgleichung stellt somit sicher, dass die beobachteten Mortalitätsraten als lineare Funktion der zugrunde liegenden Zustandsvariablen und eines Fehlerterms modelliert werden können.

Übergangsgleichung:

Die Übergangsgleichung beschreibt die zeitliche Entwicklung der Zustandsvariablen als Markov-Prozess erster Ordnung:

$$\vec{\alpha}(t) = \vec{c} + A\vec{\alpha}(t - 1) + \vec{\eta}(t),$$

mit den folgenden Komponenten:

- \vec{c} : Konstanter Vektor, der die Drift der Zustandsvariablen $(\kappa_t^{(1)}, \kappa_t^{(2)})'$ beschreibt.

- A : Übergangsmatrix, die die Abhängigkeiten zwischen den Zuständen in aufeinanderfolgenden Zeitpunkten erfasst.
- $\vec{\eta}(t)$: Zufallsvektor, der die stochastischen Veränderungen der Sterblichkeitsraten über die Zeit darstellt. Es wird angenommen, dass $\vec{\eta}(t)$ einen Mittelwertvektor von Null besitzt. Im weiteren Verlauf wird der Vektor $\vec{\eta}(t)$ auch als Innovationsprozess bzw. Innovationsvektor bezeichnet. Es sei weiters angemerkt, dass die Innovationen $\vec{\eta}(t)$ und $\vec{\eta}(s)$ für $t \neq s$ unkorreliert sind. $\vec{\eta}(t)$ repräsentiert eine unabhängige Innovation, die nichts mit den vorherigen Werten der $\vec{\alpha}(t)$ oder anderen Zeitpunkten zu tun hat.

Insbesondere $\vec{\eta}(t)$ spielt in dieser Arbeit eine zentrale Rolle, da dieser in bestehenden Modellen häufig als normalverteilt angenommen wird. Hier setzt die vorliegende Arbeit an, indem sie alternative Verteilungsannahmen untersucht. Ziel ist es, eine realistischere Modellierung der Veränderungen der Sterblichkeitsraten zu ermöglichen, da diese nicht zwangsläufig einer Normalverteilung folgen müssen. (vgl. Siu et al., 2023)

3.3 Auswahl der Verteilung

Wie im vorangegangenen Abschnitt erläutert, wird in der Literatur häufig angenommen, dass der Innovationsvektor $\vec{\eta}(t)$ einer Normalverteilung folgt. Dieser Ansatz bietet zwar eine mathematisch einfache Lösung, erweist sich jedoch als stark datenabhängig, wie im weiteren Verlauf dieser Arbeit aufgezeigt wird. Um dieser Einschränkung entgegenzuwirken, wird in dieser Arbeit ein alternativer Ansatz verfolgt: Es wird angenommen, dass $\vec{\eta}(t)$ einer Skew-t-Verteilung folgt. (vgl. Siu et al., 2023)

Die Skew-t-Verteilung ist eine erweiterte Version der Student-t-Verteilung, die zusätzlich eine Asymmetrie erlaubt. Ein zentrales Merkmal dieser Verteilung ist die Verwendung von vier Parametern: Neben Mittelwert und Varianz gibt es einen Schiefeparameter und einen Freiheitsgradparameter. Diese beiden zusätzlichen Parameter ermöglichen eine verbesserte Modellierung, da sie sowohl asymmetrische Daten als auch schwere Tails realistisch abbilden können. Dadurch bietet die Skew-t-Verteilung eine erhöhte Flexibilität im Vergleich zur Normalverteilung, die lediglich durch Mittelwert und Varianz beschrieben wird.

Eine mögliche Alternative zur Skew-t-Verteilung wäre die Skew-Normal-Verteilung. Diese erlaubt ebenfalls die Modellierung asymmetrischer Daten, verzichtet jedoch auf den Freiheitsgradparameter und ist somit weniger geeignet, um schwere Tails zu berücksichtigen. Angesichts der Zielsetzung dieser Arbeit, eine präzisere Modellierung der Innovationsprozesse zu erreichen, wurde die Skew-t-Verteilung aufgrund

ihrer größeren Anpassungsfähigkeit gewählt. (vgl. Azzalini & Capitanio, 2003)

Ein n -dimensionaler Zufallsvektor \vec{Y} , der einer multivariaten Skew-t-Verteilung folgt, wird mit

$$\vec{Y} \sim St_n(\vec{\xi}, \Omega, \vec{\zeta}, \nu)$$

bezeichnet. Hierbei stehen die Parameter für die folgenden Eigenschaften der Verteilung:

- $\vec{\xi}$: ein n -dimensionaler konstanter Vektor, der den Lageparameter beschreibt und somit die Position des Erwartungswerts bestimmt,
- Ω : eine symmetrische $n \times n$ -Matrix, die den Dispersionseffekt (bzw. die Kovarianzstruktur) zwischen den Komponenten des Vektors ausdrückt,
- $\vec{\zeta}$: ein n -dimensionaler Vektor, der den Schiefeparameter repräsentiert und die Asymmetrie der Verteilung steuert,
- ν : ein Skalar, der die Freiheitsgrade definiert und damit die Breite der Verteilung sowie das Verhalten der Tails beeinflusst.

Diese Parameter ermöglichen eine flexible Anpassung der Verteilung und machen sie insbesondere geeignet, um komplexe Datenstrukturen mit Asymmetrie und schweren Tails zu modellieren. (vgl. Azzalini & Capitanio, 2003)

Für die Schätzung der Parameter der multivariaten Skew-t-Verteilung wird die Funktion `mst.mple()` aus dem R-Package `sn` verwendet. Diese Funktion ermöglicht die Durchführung einer Maximum-Likelihood-Schätzung, um die gegebene Datenstruktur optimal an eine Skew-t-Verteilung anzupassen.

Besondere Spezialfälle dieser Verteilung ergeben sich bei spezifischen Parametereinstellungen:

- Setzt man den Schiefeparameter $\vec{\zeta} = \vec{0}$, reduziert sich die Verteilung auf eine multivariate Student-t-Verteilung.
- Lässt man zusätzlich die Freiheitsgrade $\nu \rightarrow \infty$ streben, konvergiert die Verteilung zu einer multivariaten Normalverteilung.

Diese Flexibilität erlaubt es, sowohl symmetrische als auch asymmetrische Datenverteilungen abzubilden und so verschiedene Eigenschaften der Daten zu berücksichtigen.

Die Dichtefunktion des Zufallsvektors \vec{Y} , der einer multivariaten Skew-t-Verteilung folgt, ist gegeben durch:

$$f_{\vec{Y}}(\vec{y}) = 2 \cdot t_n(\vec{y}; \nu) \cdot T_1 \left(\vec{\zeta} \Lambda^{-1}(\vec{y} - \vec{\xi}) \left(\frac{\nu + n}{(\vec{y} - \vec{\xi})' \Omega^{-1}(\vec{y} - \vec{\xi}) + \nu} \right)^{1/2}; \nu + n \right),$$

wobei

$$\Lambda = \text{diag}(\omega_{1,1}, \dots, \omega_{n,n})^{1/2},$$

mit $\omega_{i,i}$ als dem i -ten Diagonalelement von Ω . In der Gleichung beschreibt $T_1(\cdot; \nu + n)$ die Verteilungsfunktion einer univariaten Student-t-Verteilung mit $\nu + n$ Freiheitsgraden, die symmetrisch ist. Der Term $t_n(\vec{y}; \nu)$ stellt die Dichtefunktion einer multivariaten Student-t-Verteilung mit ν Freiheitsgraden dar und ist definiert als:

$$t_n(\vec{y}; \nu) = \frac{\Gamma((\nu + n)/2)}{|\Omega|^{1/2} (\pi\nu)^{n/2} \Gamma(\nu/2)} \left(1 + \frac{(\vec{y} - \vec{c})' \Omega^{-1}(\vec{y} - \vec{c})}{\nu} \right)^{-(\nu+n)/2},$$

wobei Γ die Gammafunktion bezeichnet. Die Skew-t-Verteilung kombiniert somit die Eigenschaften der multivariaten Student-t-Verteilung mit einem zusätzlichen Schiefeparameter $\vec{\zeta}$, der die Flexibilität des Modells bei der Anpassung asymmetrischer Daten erhöht. (vgl. Siu et al., 2023)

Die Momente von $\vec{\eta}(t)$ spielen in dieser Arbeit eine zentrale Rolle, da sie maßgeblich zur Optimierung der Momente der Sicherungsstrategie beitragen. Für die Berechnung dieser Momente wird ein analytischer Ansatz gewählt, da ein numerisches Vorgehen mit einem erheblichen rechnerischen Aufwand verbunden wäre.

3.4 Hedged liability

In dieser Arbeit wird bei der Betrachtung der abgesicherten Verbindlichkeit („hedged liability“) von einem Portfolio zeitlich begrenzter Leibrenten ausgegangen. Dieses Portfolio umfasst Verträge mit einer Laufzeit von T Jahren, die an Personen vergeben wurden, die zum Zeitpunkt t_0 ein Alter von x Jahren haben. Die Leibrente sieht dabei jährliche Zahlungen in Höhe von 1 \$ am Ende jedes Jahres während der Laufzeit vor, vorausgesetzt, die versicherte Person lebt noch. Für die Analyse wird angenommen, dass das Portfolio hinreichend groß ist, sodass das Mikro-Langlebigkeitsrisiko, also das Risiko durch zufällige Abweichungen bei den Mortalitätsereignissen einzelner Versicherter, vernachlässigt werden kann (vgl. Siu et al., 2023).

Im Folgenden wird der Wert der Verbindlichkeit pro Vertrag zum Zeitpunkt t_0 als L definiert. Dieser Wert entspricht der Summe der diskontierten zukünftigen Cashflows, wobei die Überlebenswahrscheinlichkeit der Versicherten berücksichtigt wird. L kann durch die folgende Formel beschrieben werden:

$$L = \sum_{u=1}^T (e^{-ru} \prod_{s=1}^u \tilde{p}(x+s, t_0+s)),$$

wobei r den Zinssatz beschreibt, mit dem die zukünftigen Cashflows abgezinst werden. Die Überlebenswahrscheinlichkeit $\tilde{p}(x, t)$ ist definiert als

$$\tilde{p}(x, t) = 1 - \tilde{q}(x, t) \approx e^{-\tilde{m}(x, t)}$$

und gibt die Wahrscheinlichkeit an, dass eine Person, die zum Zeitpunkt t das Alter $x-1$ erreicht hat, das Alter x erreicht, wobei $\tilde{q}(x, t)$ und $\tilde{m}(x, t)$ die Werte von $q(x, t)$ bzw. $m(x, t)$ repräsentieren, wenn der dazugehörige Fehlerterm $\epsilon(x, t)$ auf Null gesetzt wird.

Der Wert der Leibrente zum Zeitpunkt t_0 , basierend auf den Informationen, die bis zu einem späteren Zeitpunkt t verfügbar sind, wird mit $V_L(t)$ bezeichnet. Diese bis t bekannten Informationen umfassen sowohl die realisierten Überlebenswahrscheinlichkeiten als auch die Kenntnis darüber, ob die versicherte Person zu diesem Zeitpunkt noch am Leben ist. Entsprechend lässt sich $V_L(t)$ durch folgende Gleichung beschreiben:

$$V_L(t) = \mathbb{E}(L|\mathcal{F}_t), t \geq t_0.$$

Hierbei steht \mathcal{F}_t für die zum Zeitpunkt t verfügbare Informationsmenge, und die bedingte Erwartung $\mathbb{E}(L|\mathcal{F}_t)$ beschreibt den erwarteten Wert der zukünftigen Zahlungen L zum Zeitpunkt t_0 unter Berücksichtigung dieser Informationen. Es ist zu beachten, dass $V_L(t)$ eine Zufallsvariable darstellt, wenn sie zu einem Zeitpunkt $s < \min(t, t_0 + T)$ betrachtet wird (vgl. Siu et al., 2023).

3.5 q-Forwards

Das in dieser Arbeit verwendete Hedging-Instrument ist das sogenannte q-Forward. Der Begriff „q“ bezieht sich auf die Notation der Sterblichkeitswahrscheinlichkeit, die auch in der versicherungsmathematischen Praxis häufig verwendet wird. q-Forwards sind Finanzinstrumente, die speziell für die Übertragung von Langlebigkeits- und Sterblichkeitsrisiken entwickelt wurden. Sie sind an einen Index gebunden, der in diesem Kontext einen Langlebigkeits- oder Mortalitätsindex repräsentiert.

Die Bindung an einen Index bietet mehrere Vorteile. Neben der vergleichsweise ein-

fachen Bewertung und der geringeren Kostenstruktur zählen diese Eigenschaften zu den wesentlichen Vorzügen von q-Forwards im Vergleich zu anderen langlebigkeitsgebundenen Wertpapieren, die auf dem Markt gehandelt werden (vgl. Coughlan et al., 2007).

Ein q-Forward basiert auf dem Austausch der realisierten Sterblichkeitsrate einer bestimmten Bevölkerung zu einem späteren Zeitpunkt gegen eine im Voraus festgelegte Sterblichkeitsrate. Dieses Instrument ähnelt strukturell einem forward rate agreement, wobei der variable Teil (Floating-Leg) durch die tatsächlich realisierte Sterblichkeitsrate und der feste Teil (Fixed-Leg) durch die zuvor vereinbarte Sterblichkeitsrate repräsentiert wird.

In der Praxis werden q-Forwards vor allem von Lebensversicherern und Pensionsplannern genutzt. Für Pensionspläne bieten sie eine effektive Möglichkeit, sich gegen die steigende Lebenserwartung ihrer Mitglieder abzusichern. Lebensversicherer hingegen können q-Forwards einsetzen, um das Risiko eines unerwartet starken Anstiegs der Sterblichkeit bei ihren Versicherungsnehmern zu begrenzen (vgl. Coughlan et al., 2007).

Um das Langlebigkeitsrisiko zu reduzieren, sollten Hedger die Position des Festzins-Empfängers bei einem q-Forward einnehmen. In dieser Rolle erhalten sie Nettzahlungen vom Festzins-Zahler, wenn die zukünftigen Sterblichkeitsraten niedriger ausfallen als erwartet. Diese Zahlungen können dazu verwendet werden, die höheren Verbindlichkeiten aus längeren Rentenzahlungen zu kompensieren.

Ein prominentes Beispiel für einen solchen Swap ist der Handel zwischen J.P. Morgan und Canada Life. In diesem Fall erhielt Canada Life von J.P. Morgan Zahlungen, die für die Deckung der Rentenzahlungen an Versicherte verwendet wurden. Im Gegenzug leistete Canada Life eine Reihe von festen Zahlungen an J.P. Morgan. Ein weiteres Beispiel stellt der Handel zwischen Lucida und J.P. Morgan dar (vgl. Liu & Li, 2018).

Um das mathematische Konzept hinter den q-Forwards zu erläutern, nehmen wir an, dass unser Hedge-Portfolio aus $m \in \mathbb{N}$ q-Forwards besteht, die alle zum Zeitpunkt t_I emittiert werden, wobei $t_I \geq t_0$ gilt. Hierbei ist t_0 der Startzeitpunkt der abzusichernden Verbindlichkeit. Das Referenzalter des j -ten q-Forwards wird mit x_j für $j = 1, \dots, m$ bezeichnet, während die Zeit bis zur Fälligkeit mit T_j gegeben ist. An dieser Stelle ist wichtig zu erwähnen, dass die Laufzeit T der zeitlich begrenzten Leibrenten in unserem Portfolio nicht mit der Laufzeit der j verschiedenen q-Forwards übereinstimmen muss, wodurch eine Unterscheidung notwendig ist und eine Indexierung für die Laufzeiten der q-Forwards vorgenommen wird. Die vereinbarte Sterblichkeitswahrscheinlichkeit des j -ten q-Forwards wird durch $q^f(x_j, t_I + T_j)$

beschrieben. Für einen Nominalwert von einer Geldeinheit ergibt sich die Auszahlung des j -ten q-Forwards aus Sicht des Festzins-Empfängers zum Zeitpunkt t_0 :

$$H(j, t_I) = e^{-r \cdot (t_I - t_0 + T_j)} (q^f(x_j, t_I + T_j) - \tilde{q}(x_j, t_I + T_j)),$$

wobei r den Zinssatz zur Diskontierung der Cashflows darstellt und $\tilde{q}(x_j, t_I + T_j)$ für die tatsächlich realisierte Sterblichkeitswahrscheinlichkeit $q(x_j, t_I + T_j)$ steht, unter der Annahme, dass der Fehlerterm $\epsilon(x_j, t_I + T_j)$ gleich Null ist. Es ist zu beachten, dass aus Konsistenzgründen die Werte der q-Forwards zum Zeitpunkt t_0 betrachtet werden, da zu diesem Zeitpunkt auch die Rentenverbindlichkeiten betrachtet werden. Der Wert des j -ten q-Forwards zum Zeitpunkt t_0 gegeben jenen Informationen, die zum Zeitpunkt t verfügbar sind (also die bis dahin realisierten Sterblichkeitsraten) wird mit

$$V_H(t; j, t_I) = \mathbb{E}(H(j, t_I) | \mathcal{F}_t), t \geq t_I,$$

bezeichnet, wobei \mathcal{F}_t die verfügbare Information bis zum Zeitpunkt t repräsentiert. Für $t \geq t_I + T_j$ ist $V_H(t; j, t_I)$ keine Zufallsvariable mehr, da die Sterblichkeitsrate bereits realisiert wurde. Ebenso gilt, dass für einen Zeitpunkt $s \geq t$ (\mathcal{F}_s mit $s \geq t$) V_H keine Zufallsvariable darstellt. In allen anderen Fällen, also für \mathcal{F}_s mit $s < \min(t, t_I + T_j)$ ist $V_H(t; j, t_I)$ eine Zufallsvariable. (vgl. Siu et al., 2023)

Die in dieser Arbeit vorgestellten Hedging-Strategien berücksichtigen auch die Kosten der Absicherung. Dabei wird die von der Gegenpartei (hier: dem Festzinsszahler) verlangte Risikoprämie in die Forward-Mortalitätsrate integriert. Zur Bestimmung der Forward-Mortalitätsrate für den j -ten q-Forward in unserem Portfolio verwenden wir die folgende Formel:

$$q^f(x_j, t_I + T_j) = (1 - T_j \cdot \ell \cdot v(x_j)) \cdot \hat{q}(x_j, t_I + T_j),$$

wobei $\hat{q}(x_j, t_I + T_j)$ den besten Schätzwert für $q(x_j, t_I + T_j)$ darstellt. Dieser Schätzwert ergibt sich unter der Annahme, dass die Fehlerterme $\epsilon(x, s)$ und $\tilde{\eta}(s)$ für alle x und $s = t_I + 1, \dots, t_I + T_j$ gleich Null sind.

Der Parameter $v(x_j)$ beschreibt die geschätzte Volatilität der jährlichen Änderungen der Mortalitätsraten im Alter x_j . Der Faktor $\ell > 0$ repräsentiert den Marktpreis des Risikos und gibt die Vergütung an, die der Festzinsszahler für die Übernahme des Langlebighkeitsrisikos an den Festzinsempfänger leistet. Der Marktpreis des Risikos wirkt sich direkt auf die Forward-Mortalitätsrate aus, indem er die erwarteten Kosten für die Absicherung im Rahmen des Hedging-Prozesses reflektiert. (vgl. Siu et al., 2023)

3.5.1 LifeMetrics Index

Wie bereits erwähnt, sind q-Forwards an einen Index gekoppelt. Die Indizes für Langlebigkeits- und Mortalitätsderivate können dabei entweder angepasst oder standardisiert sein.

Angepasste Indizes repräsentieren die tatsächlichen Erfahrungen spezifischer Gruppen, die mit einer bestimmten Risikoposition verbunden sind. Dies können beispielsweise die Versicherungsnehmer eines Lebensversicherungsvertrags oder die Mitglieder eines leistungsorientierten Pensionsplans sein.

Im Gegensatz dazu spiegeln standardisierte Indizes, wie der LifeMetrics Index, die Erfahrungen einer größeren, breiter gefassten Population wider. Standardisierte Indizes bieten den Vorteil, dass sie eine allgemeine Vergleichsbasis darstellen und häufig als Grundlage für den Handel von Langlebigkeits- und Mortalitätsderivaten verwendet werden. (vgl. Coughlan et al., 2007)

Der LifeMetrics Index wurde im Jahr 2007 von J.P. Morgan eingeführt. Er basiert auf altersspezifischen Sterblichkeitswahrscheinlichkeiten bestimmter Bevölkerungsgruppen und dient als Referenzinstrument für Langlebigkeits- und Mortalitätsderivate. Im Jahr 2008 fand die erste Transaktion unter Verwendung des LifeMetrics Index statt. Seitdem hat sich der Index als Standard etabliert und wird von verschiedenen Investmentbanken genutzt, darunter namhafte Akteure wie Credit Suisse und Goldman Sachs. (vgl. Siu et al., 2023)

Die folgende Tabelle zeigt ein Beispiel für die Konditionen eines einzelnen q-Forwards, der zur Absicherung des Langlebigkeitsrisikos verwendet wird und den LifeMetrics Index als Referenz enthält:

Notional Amount	GBP 50,000,000
Trade Date	31 Dec 2006
Effective Date	31 Dec 2006
Maturity Date	31 Dec 2016
Reference year	2015
Fixed Rate	1.2000 %
Fixed Amount Payer	JPMorgan
Fixed Amount	Notional Amount x Ficed Rate x 100
Reference Rate	LifeMetrics graduated initial mortality rate for 65-year-old males in the reference year for England & Wales natioanl population Bloomberg ticker: LMQMEW65 Index <GO>
Floating Amount Payer	XYZ Pension
Floating Amount	Notional Amount x Reference Rate x 100
Settlement	Net settlement = Fixed amount - Floating amount

Tabelle 1: Term Sheet eines einzelnen q-Forwards zur Absicherung des Langlebigkeitsrisikos.

Das in Tabelle 1 dargestellte Beispiel beschreibt einen 10-Jahres-q-Forward-Vertrag, der am 31. Dezember 2006 begann und am 31. Dezember 2016 endete. Dieser Vertrag ist Teil einer langfristigen Absicherungsstrategie eines britischen Pensionsplans. Als Referenzpopulation dient die Gruppe der 65-jährigen Männer in England und Wales, wobei der LifeMetrics Index für diese Population als Grundlage für die Auszahlung verwendet wird.

Bei Fälligkeit des Vertrags zahlt der Festzinsszahler einen Betrag an den Pensionsplan, der proportional zu einer festgelegten Sterblichkeitsrate von 1,2000 % berechnet wird. Im Gegenzug leistet der Pensionsplan eine Zahlung an den Sicherungsgeber, die auf dem Referenzzinssatz basiert, welcher die tatsächlich realisierte Sterblichkeitsrate für die betrachtete Population widerspiegelt. Dieser Wert wird durch den aktuellen Stand des LifeMetrics Index zum Fälligkeitszeitpunkt repräsentiert.

Aufgrund der Verzögerung von etwa zehn Monaten bei der Veröffentlichung offizieller Daten erfolgt die Abrechnung am 31. Dezember 2016 auf Basis des LifeMetrics-Indexwerts für das Referenzjahr 2015. Dieses Vertragsmodell ermöglicht eine effektive Absicherung gegen unerwartete Veränderungen der Mortalitätsraten. (vgl. Coughlan et al., 2007)

Die Abwicklung des q-Forward-Vertrags erfolgt bei Fälligkeit auf Basis des Nettobetrags, der proportional zur Differenz zwischen der festgelegten Sterblichkeitsrate (dem Transaktionszinssatz) und der realisierten Referenzsterblichkeitsrate berechnet wird. Eine positive Differenz führt zu einer Zahlung vom Festzinsszahler an den Floating-Zahler (z. B. den Pensionsplan), während eine negative Differenz umge-

kehrt zu einer Zahlung vom Floating-Zahler an den Festzinsszahler führt. Die folgende Tabelle veranschaulicht mögliche Szenarien für verschiedene Werte des realisierten Referenzzinssatzes:

Realized Rate	Fixed Rate	Notional	Settlement
1.0000 %	1.2000 %	50,000,000	10,000,000
1.1000 %	1.2000 %	50,000,000	5,000,000
1.2000 %	1.2000 %	50,000,000	0
1.3000 %	1.2000 %	50,000,000	-5,000,000

Tabelle 2: q-Forward Vergleiche für verschiedene Ergebnisse des realisierten Referenzzinssatzes

Bei einer geringeren Sterblichkeit, das heißt, wenn der realisierte Referenzzinssatz im Referenzjahr unterhalb des festgelegten Zinssatzes liegt, ergibt sich ein positiver Abwicklungsbetrag. In diesem Fall erhält der Pensionsplan eine Ausgleichszahlung vom Hedge-Anbieter. Diese Zahlung dient dazu, den Anstieg des Wertes seiner Verbindlichkeiten aufgrund der niedrigeren Sterblichkeitsrate zu kompensieren.

Umgekehrt, bei einer höheren Sterblichkeit – wenn der Referenzzinssatz den festen Zinssatz übersteigt – ist die Abwicklung negativ. In diesem Szenario leistet der Pensionsplan eine Zahlung an den Hedge-Anbieter. Dieser Abfluss wird jedoch durch den Rückgang des Wertes der Verbindlichkeiten des Pensionsplans infolge der höheren Sterblichkeit ausgeglichen.

Die Struktur dieser Zahlungen sorgt dafür, dass der Pensionsplan sowohl bei steigender als auch bei sinkender Sterblichkeit abgesichert ist, was die Stabilität seiner finanziellen Verpflichtungen verbessert (vgl. Coughlan et al., 2007).

3.5.2 Bepreisung von q-Forwards

Die Bepreisung von q-Forwards stellt einen zentralen Aspekt in der Bewertung von Langlebigkeitsrisiken dar. q-Forwards sind Finanzinstrumente, deren Wert auf Überlebenswahrscheinlichkeiten basiert, und sie dienen vor allem dazu, Langlebigkeitsrisiken für Rentenversicherer und Pensionskassen abzusichern. Der Preis eines q-Forwards hängt sowohl von den erwarteten Überlebenswahrscheinlichkeiten als auch von einer Risikoprämie ab, die das Langlebigkeitsrisiko widerspiegelt.

Die allgemeine Formel für die Bewertung eines q-Forwards basiert auf der Forward-Mortalitätsrate ($q^f(x_j, t_I + T_j)$) und wird durch den Einfluss einer Risikoprämie modifiziert: (vgl. Bauer et al., 2010)

$$F_0 = e^{-rT_j} \cdot q^f(x_j, t_I + T_j) \cdot e^{-\int_0^T \lambda(s) ds},$$

wobei:

- F_0 der Preis des q-Forwards ist,
- e^{-rT_j} der Diskontfaktor bis zur Laufzeit T_j ist,
- $q^f(x_j, t_I + T_j)$ die Forward-Mortalitätsrate für ein Individuum im Alter x_j zum Zeitpunkt T_j ist,
- $\lambda(s)$ den Marktpreis des Langlebigkeitsrisikos beschreibt.

q^f bildet dabei die Basisgröße, während die Risikoprämie und Diskontierung zusätzliche Anpassungen zur Bewertung von F_0 darstellen. Die Risikoprämie wird benötigt, da Versicherer oder andere Marktteilnehmer für die Übernahme von Langlebigkeitsrisiken eine Entschädigung verlangen.

Die Risikoprämie ($\lambda(s)$) kann auf verschiedene Weisen modelliert werden, beispielsweise durch den Sharpe-Ratio-Ansatz oder durch einen Wang-Transform-Ansatz. Beim Sharpe-Ratio-Ansatz wird die Risikoprämie proportional zur Volatilität der Langlebigkeitsrisiken modelliert, ähnlich wie bei der Bewertung von Aktien. Dabei wird eine konstante Sharpe-Ratio verwendet, die aus anderen Märkten kalibriert werden kann. Der Ansatz liefert eine proportional zur Unsicherheit wachsende Risikoprämie, wodurch eine intuitive und konsistente Bewertung ermöglicht wird. (vgl. Bayraktar et al., 2009)

In der Praxis fehlt es häufig an Marktdaten für Langlebigkeitsderivate, weshalb die Kalibrierung der Modelle eine Herausforderung darstellt. Um die Parameter zu schätzen, könnten Daten aus Annuity-Märkten verwendet werden. Empirische Analysen zeigen jedoch, dass Risikoprämien für Annuities möglicherweise höher sind als für Langlebigkeitsderivate, da erstere auch nicht-systematische Risiken (z. B. durch kleine Portfolios) enthalten. (vgl. Lin & Cox, 2005)

4 Optimierungsproblem

In der Versicherungsbranche stellt die Unsicherheit über den Wert der Verbindlichkeiten innerhalb eines Jahres eine der größten Herausforderungen dar. Die Eigenkapitalanforderungen eines Versicherers werden in vielen Rechtssystemen durch ein Risikomaß bestimmt, das sich aus der Verteilung dieser Verbindlichkeiten über ein Jahr ergibt. Im Rahmen von Solvency II beispielsweise sind alle Versicherer in der Europäischen Union verpflichtet, ausreichend Solvenzkapital vorzuhalten, das mindestens dem Value-at-Risk (VaR) der Verbindlichkeiten bei einem Konfidenzniveau von 99,5 % entspricht. Der Schweizer Solvenzttest, der sich eng an das europäische Modell anlehnt, verwendet hingegen den Tail-VaR (auch als conditional tail expectation bekannt) bei einem Konfidenzniveau von 99 %.

Daher besteht unser Ziel darin, die Unsicherheit im Zusammenhang mit dem Wert der Rentenverbindlichkeiten innerhalb eines Jahres zu minimieren. Um übermäßige Notation zu vermeiden, nehmen wir an, dass die Absicherung gegen Langlebigkeitsrisiken zum Zeitpunkt t beginnt und dass die dafür eingesetzten q-Forwards ebenfalls zu diesem Zeitpunkt emittiert werden. Insgesamt gehen wir von m q-Forwards aus. Unser Ziel erfordert, dass der Hedger einen Nominalbetrag $N_{j,t}$ des j -ten q-Forwards für $j = 1, \dots, m$ erwirbt, um die Unsicherheit der Größe

$$V_P(t+1, \vec{N}_t) = V_L(t+1) - \sum_{j=1}^m N_{j,t} V_H(t+1; j, t)$$

im Vergleich zur Unsicherheit von $V_L(t+1)$ zu reduzieren. Dabei bezeichnet $\vec{N}_t = (N_{1,t}, \dots, N_{m,t})'$ den Vektor der Nominalbeträge. (vgl. Siu et al., 2023)

4.1 Analyse der Momente

Zur Risikominderung betrachten wir die ersten vier Momente von $V_P(t+1, \vec{N}_t)$, die im Rahmen der Optimierung minimiert werden sollen. Diese Momente sind wie folgt definiert:

1. Erwartungswert (Mittelwert):

$$M^{(1)}(\vec{N}_t) = \mathbb{E}[V_P(t+1; \vec{N}_t) | \mathcal{F}_t],$$

welcher den Mittelwert der abgesicherten Position beschreibt. Die Minimierung dieses Moments hat das Ziel, die durchschnittlichen Kosten der Absicherung

gering zu halten.

2. Varianz:

$$M^{(2)}(\vec{N}_t) = \mathbb{E}[(V_P(t+1; \vec{N}_t) - M^{(1)}(\vec{N}_t))^2 | \mathcal{F}_t],$$

welche die Volatilität der abgesicherten Position quantifiziert. Eine geringe Varianz reduziert die Schwankungen der Verbindlichkeiten und sorgt für eine stabilere Absicherung.

3. Schiefe (Skewness):

$$M^{(3)}(\vec{N}_t) = \frac{\mathbb{E}[(V_P(t+1; \vec{N}_t) - M^{(1)}(\vec{N}_t))^3 | \mathcal{F}_t]}{M^{(2)}(\vec{N}_t)^{3/2}},$$

welche die Asymmetrie der abgesicherten Position misst. Eine Minimierung dieses Moments verringert das Risiko von stark unausgewogenen Extremen, insbesondere der Wahrscheinlichkeit von sehr hohen Verbindlichkeiten.

4. Kurtosis (Wölbung):

$$M^{(4)}(\vec{N}_t) = \frac{\mathbb{E}[(V_P(t+1; \vec{N}_t) - M^{(1)}(\vec{N}_t))^4 | \mathcal{F}_t]}{M^{(2)}(\vec{N}_t)^2},$$

welche die Dicke der Verteilungstails beschreibt. Eine Minimierung der Kurtosis sorgt für schmalere Verteilungstails, wodurch extreme Ausreißer in den Verbindlichkeiten weniger wahrscheinlich werden.

Die gezielte Optimierung dieser Momente ermöglicht eine umfassende Risikominde-
rung, die nicht nur die Kosten und die Volatilität berücksichtigt, sondern auch die
Wahrscheinlichkeit und das Ausmaß extremer Ereignisse reduziert. (vgl. Siu et al.,
2023)

Die Minimierung der Schiefe ($M^{(3)}(\vec{N}_t)$) und der Kurtosis ($M^{(4)}(\vec{N}_t)$) ist entschei-
dend für eine ausgewogene Absicherung der Verbindlichkeiten. Während die Reduk-
tion des Mittelwerts und der Varianz direkt Kosten und Schwankungen kontrolliert,
zielt die Verringerung der Schiefe darauf ab, extreme positive Abweichungen in den
Verbindlichkeiten zu reduzieren. Eine geringe Kurtosis sorgt dafür, dass die Vertei-
lungstails schmaler werden, wodurch extreme Werte unwahrscheinlicher werden.

Um die Bedeutung der Schiefe graphisch zu veranschaulichen, vergleichen wir zwei
Dichtefunktionen von skew-t-verteiltern Zufallsvariablen. Beide Zufallsvariablen ha-
ben einen Mittelwert von 0 und eine Varianz von 1, jedoch unterschiedliche Schiefe-
Werte ($-0,45$ bzw. $-0,93$). Die erste Dichtefunktion mit einer moderaten Schiefe
($-0,45$) zeigt eine relativ symmetrische Verteilung mit leichtem Übergewicht in

den positiven Tails. Im Vergleich dazu weist die zweite Dichtefunktion mit stärkerer negativer Schiefe ($-0,93$) eine deutlichere Asymmetrie auf, was das Risiko von extremen Abweichungen verstärkt. (vgl. Siu et al., 2023)

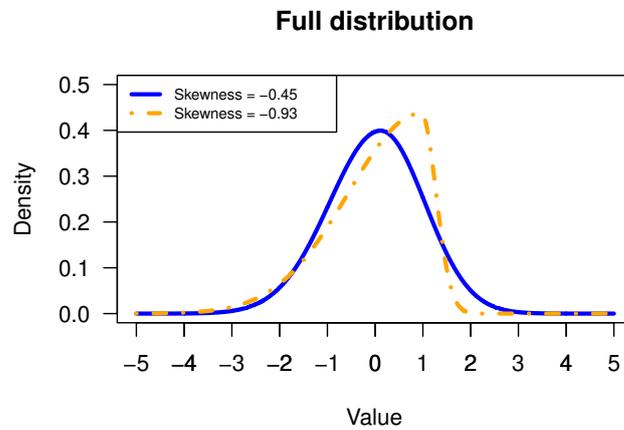


Abbildung 4.1: Dichtefunktionen zweier skew-t Verteilungen mit verschiedenen Schiefen

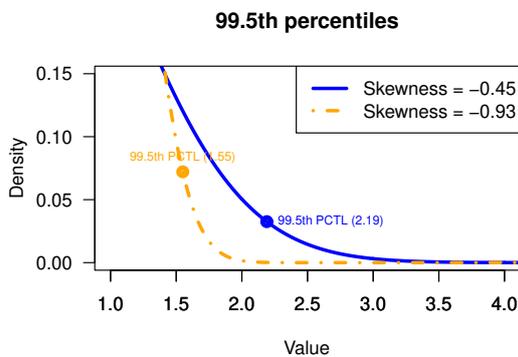


Abbildung 4.2: 99,5 % Percentil der jeweiligen Verteilungen

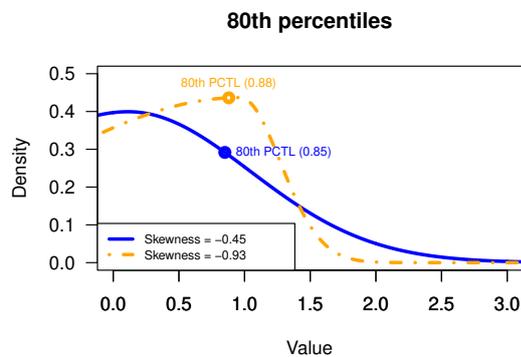


Abbildung 4.3: 80 % Percentil der jeweiligen Verteilungen

Die Auswirkungen einer Änderung der Schiefe auf die Dichtefunktion und deren praktische Relevanz für die Risikominderung lassen sich anhand der Abbildungen 4.1 bis 4.3 verdeutlichen:

1. Abbildung 4.1 (Veränderung der Dichtefunktion): Diese Abbildung zeigt, wie sich eine Änderung der Schiefe auf die Form der Verteilung auswirkt. Eine Schiefe näher der Null führt zu einer symmetrischeren Verteilung, während eine betragsmäßig höhere Schiefe die Dichtefunktion asymmetrisch macht und die Wahrscheinlichkeit für extreme Werte in einem Tail erhöht.

2. Abbildung 4.2 (Auswirkungen auf das 99,5 %-Perzentil): In Abbildung 4.2 wird deutlich, dass eine Verteilung mit geringerer Schiefe ein niedrigeres 99,5 %-Perzentil aufweist. Dies ist aus Absicherungssicht vorteilhaft, da es darauf hinweist, dass die Wahrscheinlichkeit extremer negativer Ereignisse reduziert wird. Für einen Versicherer bedeutet dies eine geringere Kapitalbindung zur Deckung der Verbindlichkeiten bei seltenen Extremszenarien.
3. Abbildung 4.3 (Betrachtung des 80 %-Perzentils): Abbildung 4.3 verdeutlicht jedoch auch, dass eine geringere Schiefe nicht immer von Vorteil ist. Beim Vergleich des 80 %-Perzentils zeigt sich, dass die Verteilung mit höherer Schiefe in diesem Bereich bessere Eigenschaften aufweist, da die Wahrscheinlichkeit für moderate negative Abweichungen geringer ist. Dies könnte in Fällen vorteilhaft sein, in denen solche Abweichungen für den Versicherer weniger kritisch sind.

Ein ähnlicher Zusammenhang wie bei der Schiefe zeigt sich auch bei der Kurtosis. Die in dieser Arbeit verwendeten Verteilungen wurden mit dem R-Package `sn` von Adelchi Azzalini generiert, wobei die Funktion `dst` zur Erstellung der Dichtefunktion einer skew-t-Verteilung herangezogen wurde. Diese Funktion erfordert vier Parameter: einen Lageparameter, einen Varianzparameter, einen Parameter für die Schiefe sowie einen Parameter für die Freiheitsgrade (degrees of freedom).

Ein direkter Parameter zur gezielten Einstellung eines spezifischen Kurtosis-Wertes existiert jedoch nicht. Stattdessen kann die Kurtosis durch die Wahl der Freiheitsgrade beeinflusst werden. Niedrige Werte der Freiheitsgrade führen zu einer höheren Kurtosis. Mit zunehmendem Wert der Freiheitsgrade nähert sich die Verteilungsfunktion hingegen einer Normalverteilung an, die eine Kurtosis von 3 aufweist.

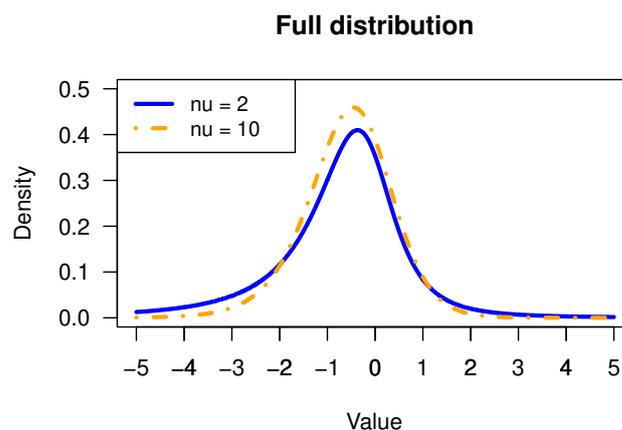


Abbildung 4.4: Dichtefunktionen zweier skew-t-Verteilungen mit verschiedenen Kurtosen

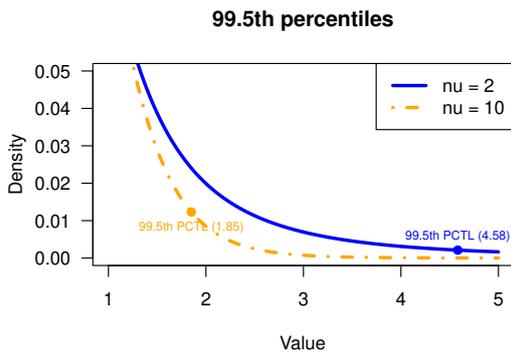


Abbildung 4.5: 99,5 % Perzentil der jeweiligen Verteilungen

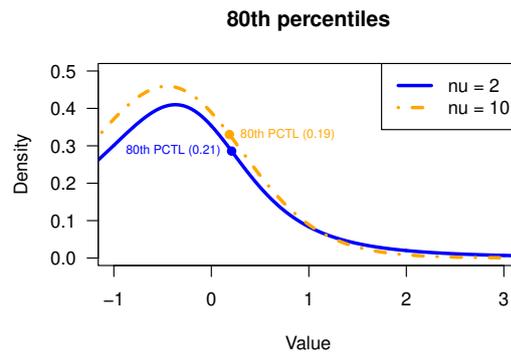


Abbildung 4.6: 80 % Perzentil der jeweiligen Verteilungen

In obigen Abbildungen werden erneut zwei Dichtefunktionen von skew-t-verteiltern Zufallsvariablen dargestellt, die sich ausschließlich in den Freiheitsgraden unterscheiden. Es ist ersichtlich, dass eine höhere Kurtosis zu einem deutlich größeren 99,5%-Perzentil führt. Betrachtet man jedoch das 80-%-Perzentil, unterscheiden sich die jeweiligen Werte nur geringfügig. (vgl. Siu et al., 2023)

Ähnlich wie bei der Schiefe ist eine Minimierung der Kurtosis daher von Vorteil, insbesondere wenn der Fokus auf den äußeren Bereichen der rechten Verteilungstails liegt. Die Reduktion der Kurtosis kann die Wahrscheinlichkeit extremer Werte senken und trägt somit dazu bei, die Unsicherheit in den Verbindlichkeiten besser zu kontrollieren.

4.2 Polynomial goal programming

Polynomial Goal Programming (PGP) ist eine Technik des Zielprogrammierens, die bei der Lösung von Entscheidungsproblemen eingesetzt wird, wenn mehrere, häufig konkurrierende Ziele verfolgt werden und die Beziehungen zwischen diesen Zielen oder den Entscheidungsvariablen nicht-linear sind. Ziel solcher Modelle ist es, die Summe der gewichteten Abweichungen von einer Reihe vorab definierter und hierarchisch geordneter Ziele zu minimieren.

Ein zentrales Element von PGP-Modellen ist die Gewichtung der Ziele, da sie die Präferenzen des Hedgers widerspiegelt. Die Wahl der Gewichte beeinflusst daher maßgeblich die Ergebnisse der Optimierung und deren praktischen Nutzen. Das Lösen von PGP-Modellen erfordert den Einsatz spezialisierter Software. In dieser Arbeit wurden sämtliche Optimierungsverfahren in der Programmiersprache R implementiert. (vgl. Deckro & Hebert, 1988)

Um die Rahmenbedingungen unseres Optimierungsproblems zu definieren, legen wir zunächst fest, dass

$$N_{j,t} \geq 0, j = 1, \dots, m.$$

Diese Nebenbedingung stellt sicher, dass der Hedger die Position des festverzinslichen Empfängers einnimmt und somit keine Leerverkäufe von q-Forwards zulässig sind.

Zusätzlich fordern wir, dass

$$\sum_{j=1}^m N_{j,t} \mathbb{E}[V_H(t+1; j, t) | \mathcal{F}_t] \leq C \cdot \mathbb{E}[V_L(t+1) | \mathcal{F}_t].$$

Diese Nebenbedingung berücksichtigt ein Budget C , das dem Hedger für die Absicherung zur Verfügung steht. Der Wert von C wird als ein Bruchteil des Wertes der Rentenverbindlichkeiten definiert, wodurch sichergestellt wird, dass die Absicherung im Verhältnis zu den verfügbaren finanziellen Mitteln erfolgt.

Das vollständige Optimierungsproblem kann unter Berücksichtigung der beschriebenen Zielsetzungen und Nebenbedingungen vorerst wie folgt formuliert werden:

minimiere: $M^{(i)}(\vec{N}_t), i = 1, 2, 3, 4$

Nebenbedingungen: $N_{j,t} \geq 0, j = 1, \dots, m$
 $\sum_{j=1}^m N_{j,t} \mathbb{E}[V_H(t+1; j, t) | \mathcal{F}_t] \leq C \cdot \mathbb{E}[V_L(t+1) | \mathcal{F}_t]$

Um das Optimierungsproblem mithilfe von Polynomial Goal Programming (PGP) zu lösen, werden die vier Ziele in einer aggregierten Zielfunktion zusammengefasst.

Um die Zielfunktion für das Polynomial Goal Programming (PGP) zu formulieren, definieren wir zunächst einige zentrale Größen:

1. Referenzlösung für jedes Moment: $\mathcal{M}^{(i)}$ bezeichnet die optimale Lösung des Teilproblems, das sich ausschließlich auf das i -te Moment konzentriert, für $i = 1, 2, 3, 4$. Diese Referenzwerte dienen als Benchmark zur Bewertung der Abweichungen.
2. Abweichungsmetrik: Um die Abweichung zwischen dem Wert $M^{(i)}(\vec{N}_t)$ eines Moments und der entsprechenden Referenzlösung $\mathcal{M}^{(i)}$ zu messen, verwenden wir den normierten Ausdruck:

$$\left| \frac{M^{(i)}(\vec{N}_t) - \mathcal{M}^{(i)}}{\mathcal{M}^{(i)}} \right|.$$

Diese Metrik gibt die relative Abweichung an und ermöglicht eine vergleichbare Bewertung über die verschiedenen Momente hinweg.

3. Gewichte: Die Gewichtungsfaktoren $\lambda_i, i = 1, 2, 3, 4$, spiegeln die Präferenzen des Hedgers wider. Ein höherer Wert von λ_i zeigt an, dass das entsprechende Moment für den Hedger von größerer Bedeutung ist. Diese Werte werden vor Beginn der Optimierung festgelegt und sind entscheidend für die Priorisierung der Ziele.
4. Zielvariablen: Mit g_i bezeichnen wir die Zielvariable, die die Abweichung von $M^{(i)}(\vec{N}_t)$ im Vergleich zu $\mathcal{M}^{(i)}$ misst. Diese Größe beschreibt, um wie viel größer der Wert von $M^{(i)}(\vec{N}_t)$ im Vergleich zur Benchmark ist.

Mit diesen Definitionen lässt sich eine PGP-Zielfunktion formulieren, die alle vier Momente berücksichtigt, während die Gewichtungen λ_i die Präferenzen steuern. Die Zielgrößen g_i ermöglichen eine direkte Messung und Steuerung der Abweichungen, sodass das Optimierungsproblem sowohl flexibel als auch auf die spezifischen Anforderungen des Hedgers zugeschnitten ist.

Das beschriebene Optimierungsproblem wird im Rahmen des Polynomial Goal Programming (PGP) folgendermaßen formuliert:

$$\begin{aligned} \text{minimiere:} \quad & \lambda_1 \left| \frac{g_1}{\mathcal{M}^{(1)}} \right| + \lambda_2 \left| \frac{g_2}{\mathcal{M}^{(2)}} \right| + \lambda_3 \left| \frac{g_3}{\mathcal{M}^{(3)}} \right| + \lambda_4 \left| \frac{g_4}{\mathcal{M}^{(4)}} \right| \\ \text{Nebenbedingungen:} \quad & N_{j,t} \geq 0, j = 1, \dots, m \\ & M^{(i)}(\vec{N}_t) - g_i = \mathcal{M}^{(i)}, i = 1, 2, 3, 4 \\ & \sum_{j=1}^m N_{j,t} \mathbb{E}[V_H(t+1; j, t) | \mathcal{F}_t] \leq C \cdot \mathbb{E}[V_L(t+1) | \mathcal{F}_t] \\ & g_i \geq 0, i = 1, 2, 3, 4 \end{aligned}$$

Die Zielfunktion kombiniert die gewichteten Abweichungen der vier Momente $M^{(i)}(\vec{N}_t)$ von den jeweiligen Referenzwerten $\mathcal{M}^{(i)}$. Die Gewichtungen λ_i steuern dabei die Priorität, die der Hedger jedem Moment zuweist.

Die Nebenbedingungen lassen sich folgendermaßen erläutern:

1. Die erste Nebenbedingung stellt sicher, dass der Hedger die Rolle des Empfängers des festen Zinses übernimmt, indem die Nominalbeträge der q-Forwards nicht negativ sind.
2. Die zweite Bedingung koppelt die Abweichung g_i mit dem Moment $M^{(i)}(\vec{N}_t)$ und dessen Benchmark $\mathcal{M}^{(i)}$.
3. Die dritte Nebenbedingung begrenzt die Kosten des Hedge-Portfolios durch das verfügbare Budget C , das als ein Bruchteil des Wertes der Rentenverbindlichkeiten definiert ist.

4. Die letzte Bedingung stellt sicher, dass die Abweichungen g_i nicht negativ sind, sodass die Referenzlösungen $\mathcal{M}^{(i)}$ die bestmöglichen Werte darstellen.

Die Lösung dieses Optimierungsproblems liefert die optimalen Nominalbeträge $\hat{N}_{j,t}$ für $j = 1, \dots, m$, die in das Hedge-Portfolio aufgenommen werden sollen, um die Unsicherheit der Rentenverbindlichkeiten effektiv zu reduzieren.

5 Analytische Herleitung der Ausdrücke

In diesem Abschnitt wird der mathematische Hintergrund der bisher definierten Ausdrücke näher beleuchtet. Dazu gehören sowohl die Approximation verschiedener Terme, beispielsweise mithilfe von Simulationen, als auch die analytische Herleitung, wenn der Aufwand für Simulationen unverhältnismäßig hoch ist. Ziel ist es, die verwendeten Methoden und deren Anwendung im Kontext des Modells systematisch darzustellen.

Zunächst betrachten wir den Vektor

$$(V_L(t+1), V_H(t+1; 1, t), \dots, V_H(t+1; m, t))',$$

welcher die Zufallsvariablen für den Wert der Rentenverbindlichkeiten und der Hedging-Instrumente beschreibt. Für diesen Vektor definieren wir die Momente bis zur vierten Ordnung wie folgt:

$$\begin{aligned} M_t &= \{\mu_{i,t} : i = 1, \dots, m+1\}, && \text{(Erwartungswerte),} \\ V_t &= \{\sigma_{ij,t} : i, j = 1, \dots, m+1\}, && \text{(Varianzen und Kovarianzen),} \\ S_t &= \{s_{iju,t} : i, j, u = 1, \dots, m+1\}, && \text{(dritte Momente),} \\ K_t &= \{k_{ijuv,t} : i, j, u, v = 1, \dots, m+1\}, && \text{(vierte Momente).} \end{aligned}$$

Diese Momente beschreiben die zentralen statistischen Eigenschaften der betrachteten Zufallsvariablen und sind für die Berechnung der Zielgrößen sowie die Optimierung des Hedge-Portfolios essenziell.

Mithilfe der Momente M_t, V_t, S_t, K_t lassen sich nun die Formeln für die Momente $M^{(i)}(\vec{N}_t), i = 1, 2, 3, 4$ ableiten. Diese beschreiben die Eigenschaften der abgesicherten Position $V_P(t+1; \vec{N}_t)$:

$$\begin{aligned} M^{(1)}(\vec{N}_t) &= \mathbb{E} \left[V_P(t+1; \vec{N}_t) \mid \mathcal{F}_t \right] = \sum_{i=1}^{m+1} w_{i,t} \mu_{i,t}, \\ M^{(2)}(\vec{N}_t) &= \mathbb{E} \left[\left(V_P(t+1; \vec{N}_t) - M^{(1)}(\vec{N}_t) \right)^2 \mid \mathcal{F}_t \right] = \sum_{i=1}^{m+1} \sum_{j=1}^{m+1} w_{i,t} w_{j,t} \sigma_{ij,t}, \end{aligned}$$

$$\begin{aligned}
 M^{(3)}(\vec{N}_t) &= \frac{\mathbb{E} \left[\left(V_P(t+1; \vec{N}_t) - M^{(1)}(\vec{N}_t) \right)^3 \mid \mathcal{F}_t \right]}{\left(M^{(2)}(\vec{N}_t) \right)^{3/2}} \\
 &= \frac{\sum_{i=1}^{m+1} \sum_{j=1}^{m+1} \sum_{u=1}^{m+1} \omega_{i,t} \omega_{j,t} \omega_{u,t} s_{iju,t}}{\left(\sum_{i=1}^{m+1} \sum_{j=1}^{m+1} \omega_{i,t} \omega_{j,t} \sigma_{ij,t} \right)^{3/2}}
 \end{aligned}$$

und schließlich

$$\begin{aligned}
 M^{(4)}(\vec{N}_t) &= \frac{\mathbb{E} \left[\left(V_P(t+1; \vec{N}_t) - M^{(1)}(\vec{N}_t) \right)^4 \mid \mathcal{F}_t \right]}{\left(M^{(2)}(\vec{N}_t) \right)^2} \\
 &= \frac{\sum_{i=1}^{m+1} \sum_{j=1}^{m+1} \sum_{u=1}^{m+1} \sum_{v=1}^{m+1} \omega_{i,t} \omega_{j,t} \omega_{u,t} \omega_{v,t} k_{ijuv,t}}{\left(\sum_{i=1}^{m+1} \sum_{j=1}^{m+1} \omega_{i,t} \omega_{j,t} \sigma_{ij,t} \right)^2},
 \end{aligned}$$

Hierbei sind die Gewichte $\omega_{i,t}$ definiert als:

$$\omega_{1,t} = 1 \text{ und } \omega_{i,t} = -N_{i-1} \text{ f\"ur } i = 2, \dots, m+1.$$

Die oben angegebenen Formeln verdeutlichen, dass die Werte der Momente M_t, V_t, S_t, K_t eine entscheidende Rolle bei der Minimierung der Zielgrößen $M^{(i)}(\vec{N}_t)$, $i = 1, 2, 3, 4$ spielen. Eine Möglichkeit, diese Werte zu bestimmen, besteht darin, Simulationen durchzuführen. Allerdings bringt diese Methode erhebliche Herausforderungen mit sich, da die Simulationen verschachtelt durchgeführt werden müssen.

Konkret erfordert die Simulationsmethode zunächst die Generierung einer großen Anzahl von Simulationen für die Sterbewahrscheinlichkeiten $q(x, t)$, was in unserem Modell die Simulation verschiedener Werte von $\vec{\eta}(t)$ umfasst. Für jede dieser $\vec{\eta}(t)$ -Simulationen müsste anschließend erneut eine große Anzahl an Simulationen durchgeführt werden, um die Momente M_t, V_t, S_t, K_t zu berechnen. Der damit verbundene Rechenaufwand wäre enorm und könnte die praktische Anwendbarkeit des Modells erheblich einschränken.

Um diesen hohen Aufwand zu vermeiden, wird in dieser Arbeit ein analytischer Ansatz gewählt. Ziel ist es, geschlossene Formeln für die Werte von M_t, V_t, S_t, K_t herzuleiten. Dieser Ansatz erlaubt eine effiziente Berechnung der Momente, ohne auf zeitaufwändige Simulationen angewiesen zu sein. Dabei werden die mathematischen Eigenschaften des Modells genutzt, um präzise und verlässliche Ausdrücke für die benötigten Werte zu entwickeln. (vgl. Siu et al., 2023)

5.1 Approximierte Ausdrücke

Für die Berechnungen der Momente M_t, V_t, S_t, K_t werden approximierte Ausdrücke für L und $H(j, t)$ verwendet, um die Komplexität der Berechnungen zu reduzieren. Für die Herleitung der Ausdrücke wird auf Ergebnisse von Siu et al., 2023 verwiesen, welche in diesem Abschnitt nochmals zusammengefasst werden. Dabei werden die zugrunde liegenden Definitionen der beiden Größen erneut in Erinnerung gerufen:

- L repräsentiert die Summe der zum Zeitpunkt t_0 diskontierten Cashflows pro Vertrag in unserem Portfolio. Diese Größe gibt den aktuellen Wert der zukünftigen Rentenzahlungen basierend auf den Überlebenswahrscheinlichkeiten der Versicherten wieder.
- $H(j, t)$ bezeichnet die Auszahlung des j -ten q-Forwards aus der Perspektive des Empfängers mit festverzinslichem Zinssatz zum Zeitpunkt t_0 .

Um L zu approximieren, verwenden wir eine Taylorentwicklung erster Ordnung, die um die Innovationsvektoren $\vec{\eta}(t+1), \dots, \vec{\eta}(t_0+T)$ zwischen den Zeitpunkten $t+1$ und dem Endzeitpunkt unserer Rentenverbindlichkeiten t_0+T expandiert wird. Die Approximation von L ergibt sich somit zu:

$$L \approx l = \hat{L} + \sum_{s=t+1}^{t_0+T} \left(\frac{\partial L}{\partial \vec{\eta}(s)} \right)' \cdot \left(\vec{\eta}(s) + \vec{\hat{\eta}}(s) \right),$$

wobei die einzelnen Terme wie folgt definiert sind:

- $\vec{\hat{\eta}} = \mathbb{E}(\vec{\eta}(s) | \mathcal{F}_t) = \vec{0}$ für $s = t+1, \dots, t_0+T$,
- \hat{L} ist der Wert von L , der berechnet wird, indem die realisierten Werte von $\vec{\eta}(t_0), \dots, \vec{\eta}(t)$ verwendet werden und angenommen wird, dass $\vec{\eta}(s) = \vec{\hat{\eta}}(s) = \vec{0}$ für $s = t+1, \dots, t_0+T$,
- $\partial L / \partial \vec{\eta}(s)$ bezeichnet die partielle Ableitung von L nach $\vec{\eta}(s)$, berechnet unter der Verwendung der realisierten Werte von $\vec{\eta}(t_0), \dots, \vec{\eta}(t)$ und der Annahme $\vec{\eta}(u) = \vec{\hat{\eta}}(u) = \vec{0}$ für $u = t+1, \dots, t_0+T$.

Der Wert von L gegeben jenen Information bis zu einem Zeitpunkt $t+1$ lässt sich unter Verwendung der Taylor-Approximation und der bedingten Erwartung wie folgt darstellen:

$$V_L(t+1) = \mathbb{E}(L | \mathcal{F}_{t+1}) \approx V_l(t+1) := \mathbb{E}(l | \mathcal{F}_{t+1}) = \hat{L} + \left(\frac{\partial L}{\partial \vec{\eta}(t+1)} \right)' \vec{\eta}(t+1).$$

Hierbei wird ausschließlich auf $\vec{\eta}(t+1)$ Bezug genommen, da die bedingte Erwartung von $\vec{\eta}(s)$ für $s > t + 1$ ein Nullvektor ist und $\vec{\eta}(t + 1)$ zum Zeitpunkt $t + 1$ bereits bekannt ist. Die Schwierigkeit liegt dabei in der Herleitung des Ausdrucks $\frac{\partial L}{\partial \vec{\eta}(t+1)}$. Zur Erinnerung: L beschreibt die abgezinste Summe der zukünftigen Cashflows eines Rentenportfolios und wird wie folgt definiert:

$$L = \sum_{u=1}^T (e^{-ru} \prod_{s=1}^u \tilde{p}(x + s, t_0 + s)),$$

wobei $\tilde{p}(x + s, t_0 + s)$ die Überlebenswahrscheinlichkeit repräsentiert.

Es ist dabei entscheidend, dass $\tilde{p}(x + s, t_0 + s)$ nur von $\vec{\eta}(t + 1)$ abhängt, wenn $s < t - t_0 + 1$. Daraus ergibt sich für $\frac{\partial L}{\partial \vec{\eta}(t+1)}$:

$$\frac{\partial L}{\partial \vec{\eta}(t+1)} = \sum_{u=t-t_0+1}^T e^{-ru} \frac{\partial}{\partial \vec{\eta}(t+1)} \left(\prod_{s=1}^u \tilde{p}(x + s - 1, t_0 + s) \right).$$

Die Ableitung des Produktterms lautet (Produktregel):

$$\begin{aligned} & \frac{\partial}{\partial \vec{\eta}(t+1)} \left(\prod_{s=1}^u \tilde{p}(x + s - 1, t_0 + s) \right) \\ &= \sum_{s=t-t_0+1}^u \frac{\partial \tilde{p}(x + s - 1, t_0 + s)}{\partial \vec{\eta}(t+1)} \prod_{\substack{w=1 \\ w \neq s}}^u \tilde{p}(x + w - 1, t_0 + w). \end{aligned}$$

Setzt man die Ableitung von $\tilde{p}(x, t)$ nach $\vec{\eta}(t + 1)$ ein, ergibt sich:

$$\begin{aligned} & \frac{\partial}{\partial \vec{\eta}(t+1)} \left(\prod_{s=1}^u \tilde{p}(x + s - 1, t_0 + s) \right) \\ &= \sum_{s=t-t_0+1}^u \frac{\partial \tilde{p}(x + s - 1, t_0 + s)}{\partial \vec{y}(x + s - 1, t_0 + s)} \frac{\partial \vec{y}(x + s - 1, t_0 + s)}{\partial \vec{\eta}(t+1)} \prod_{\substack{w=1 \\ w \neq s}}^u \tilde{p}(x + w - 1, t_0 + w). \end{aligned}$$

Hierbei ist zu beachten:

- $y(x, t)$ ist das Alter- x -bezogene Element des Beobachtungsvektors $\vec{y}(t)$ aus der Beobachtungsgleichung.
- $\tilde{y}(x, t)$ bezeichnet den Wert von $y(x, t)$, wenn der Innovationsvektorfehler $\epsilon(x, t)$ auf null gesetzt wird.

Die Herleitung der partiellen Ableitungen aus den Modellen und ihrer Struktur erfordert präzise Ausdrücke, um die Berechnungen konsistent zu gestalten. Zunächst

betrachten wir die partielle Ableitung von

$$\frac{\partial \tilde{p}(x+s-1, t_0+s)}{\partial \tilde{y}(x+s-1, t_0+s)}.$$

Diese Ableitung ist modellabhängig und ergibt sich beispielsweise:

1. Im Lee-Carter-Modell ($y(x, t) = \log(m(x, t))$):

$$\frac{\partial \tilde{p}(x+s-1, t_0+s)}{\partial \tilde{y}(x+s-1, t_0+s)} = -\tilde{p}(x+s-1, t_0+s) \cdot \tilde{m}(x+s-1, t_0+s),$$

wobei $\tilde{m}(x, t)$ die Sterblichkeitsrate beschreibt.

2. Im Cairns-Blake-Dowd-Modell ($y(x, t) = \log(q(x, t)/(1 - q(x, t)))$):

$$\frac{\partial \tilde{p}(x+s-1, t_0+s)}{\partial \tilde{y}(x+s-1, t_0+s)} = -\tilde{p}(x+s-1, t_0+s) \cdot \tilde{q}(x+s-1, t_0+s),$$

wobei $\tilde{q}(x, t)$ die Sterbewahrscheinlichkeit bezeichnet.

Für die partielle Ableitung

$$\frac{\partial \tilde{y}(x+s-1, t_0+s)}{\partial \tilde{\eta}(t+1)}$$

betrachten wir die Struktur von $\tilde{y}(x, t)$, welche im Zustandsraummodell dargestellt werden kann als:

$$\tilde{y}(x+s-1, t_0+s) = d(x+s-1) + B(x+s-1, \cdot) \cdot \vec{\alpha}(t_0+s).$$

Die Zustände $\vec{\alpha}(t_0+s)$ lassen sich über die Übergangsgleichung ausdrücken:

$$\vec{\alpha}(t_0+s) = A^{t_0+s-t} \vec{\alpha}(t) + \left(\sum_{w=1}^{t_0+s-t} A^{t_0+s-t-w} (\vec{c} + \vec{\eta}(t+w)) \right),$$

wobei A die Übergangsmatrix und \vec{c} der Driftterm ist.

Setzt man dies nun in die Gleichung für \tilde{y} ein, ergibt sich:

$$\tilde{y}(x+s-1, t_0+s) = d(x+s-1) + B(x+s-1, \cdot) \left(A^{t_0+s-t} \vec{\alpha}(t) + \sum_{w=1}^{t_0+s-t} A^{t_0+s-t-w} \vec{\eta}(t+w) \right),$$

wobei $d(x+s-1)$ unabhängig von $\vec{\eta}(t+1)$ ist.

Daraus folgt für die partielle Ableitung:

$$\frac{\partial \tilde{y}(x+s-1, t_0+s)}{\partial \tilde{\eta}(t+1)} = \frac{\partial B(x+s-1, \cdot) \left(\sum_{w=1}^{t_0+s-t} A^{t_0+s-t-w} \vec{\eta}(t+w) \right)}{\partial \tilde{\eta}(t+1)}.$$

Da es sich hierbei um eine Zeilen-Vektor-Multiplikation handelt, ergibt sich der transponierte Ausdruck:

$$\frac{\partial \tilde{y}(x + s - 1, t_0 + s)}{\partial \vec{\eta}(t + 1)} = (B(x + s - 1, \cdot)A^{t_0+s-t-1})'.$$

Damit gilt für $\frac{\partial L}{\partial \vec{\eta}(t+1)}$:

$$\begin{aligned} \frac{\partial L}{\partial \vec{\eta}(t + 1)} = & \sum_{u=t-t_0+1}^T e^{-ru} \left(\sum_{s=t-t_0+1}^u -\tilde{p}(x + s - 1, t_0 + s) \cdot \tilde{m}(x + s - 1, t_0 + s) \right. \\ & \left. \cdot (B(x + s - 1, \cdot)A^{t_0+s-t-1})' \prod_{\substack{w=1 \\ w \neq s}}^u \tilde{p}(x + w - 1, t_0 + w) \right), \end{aligned}$$

wenn das Lee-Carter-Modell verwendet wird und

$$\begin{aligned} \frac{\partial L}{\partial \vec{\eta}(t + 1)} = & \sum_{u=t-t_0+1}^T e^{-ru} \left(\sum_{s=t-t_0+1}^u -\tilde{p}(x + s - 1, t_0 + s) \cdot \tilde{q}(x + s - 1, t_0 + s) \right. \\ & \left. \cdot (B(x + s - 1, \cdot)A^{t_0+s-t-1})' \prod_{\substack{w=1 \\ w \neq s}}^u \tilde{p}(x + w - 1, t_0 + w) \right), \end{aligned}$$

wenn auf das Cairns-Blake-Dowd-Modell zurückgegriffen wird.

Der Wert $H(j, t)$, welcher die Auszahlung des j -ten q-Forwards zum Zeitpunkt t_0 in unserem Hedge-Portfolio darstellt, kann für einen Zeitpunkt $t \geq t_0$ mithilfe einer Approximation über eine Taylorreihe erster Ordnung dargestellt werden. Die Approximation erfolgt um die Innovationsvektoren $\vec{\eta}(t + 1), \dots, \vec{\eta}(t + T_j)$, die den Zeitraum zwischen $t + 1$ und der Fälligkeit $t + T_j$ des j -ten q-Forwards umfassen (T_j beschreibt wieder die Maturität des j -ten q-Forwards). Dies führt zu folgender Näherung:

$$H(j, t) \approx h(j, t) := \hat{H}(j, t) + \sum_{s=t+1}^{t+T_j} \left(\frac{\partial H(j, t)}{\partial \vec{\eta}(s)} \right)' (\vec{\eta}(s) - \vec{\eta}(s)),$$

für $j = 1, \dots, m$.

Dabei gilt:

1. Erwartungswert der Innovationsvektoren:

Für $s = t + 1, \dots, t + T_j$ ist die Erwartung der Innovationsvektoren $\vec{\eta}(s)$ definiert als:

$$\vec{\eta}(s) = \mathbb{E}(\vec{\eta}(s) | \mathcal{F}_t) = \vec{0}.$$

2. Basiswert $\hat{H}(j, t)$:

Der Wert $\hat{H}(j, t)$ ist der Wert von $H(j, t)$, wenn alle zukünftigen Innovationsvektoren gleich ihrem erwarteten Wert sind, d. h., wenn $\vec{\eta}(s) = \vec{\hat{\eta}}(s) = \vec{0}$ für $s = t + 1, \dots, t + T_j$.

3. Partielle Ableitungen:

Die Terme $\partial H(j, t) / \partial \vec{\eta}(s)$ stellen die partiellen Ableitungen von $H(j, t)$ nach den Innovationsvektoren $\vec{\eta}(s)$ dar, berechnet unter der Annahme, dass $\vec{\eta}(u) = \vec{\hat{\eta}}(u) = \vec{0}$ für $u = t + 1, \dots, t + T_j$.

Für den Wert des j -ten q-Forwards im Hedge-Portfolio gegeben jenen Information bis zu einem Zeitpunkt $t + 1$ ergibt sich folgende Darstellung:

$$\begin{aligned} V_H(t + 1; j, t) &= \mathbb{E}(H(j, t) | \mathcal{F}_{t+1}) \\ &\approx V_h(t + 1; j, t) \\ &:= \mathbb{E}(h(j, t) | \mathcal{F}_{t+1}) \\ &= \hat{H}(j, t) + \left(\frac{\partial H(j, t)}{\partial \vec{\eta}(t + 1)} \right)' \vec{\eta}(t + 1). \end{aligned}$$

Hierbei wird ausschließlich auf den Innovationsvektor $\vec{\eta}(t + 1)$ Bezug genommen. Dies liegt daran, dass die Erwartung von $\vec{\eta}(s)$ für $s > t + 1$ ein Nullvektor ist und $\vec{\eta}(t + 1)$ zum Zeitpunkt $t + 1$ bereits bekannt ist.

Die Herausforderung bei der Berechnung liegt im Ausdruck $\frac{\partial H(j, t)}{\partial \vec{\eta}(t + 1)}$. Zur Herleitung sei die Formel für $H(j, t)$ wie folgt in Erinnerung gerufen:

$$H(j, t) = e^{-r \cdot (t - t_0 + T_j)} (q^f(x_j, t + T_j) - \tilde{q}(x_j, t + T_j)),$$

wobei $q^f(x_j, t + T_j)$ die Forward-Mortalitätsrate, $\tilde{q}(x_j, t + T_j)$ die geschätzte Mortalitätsrate darstellt (also wenn der Fehlerterm vernachlässigt wird) und x_j beschreibt wieder das Referenzalter des j -ten q-Forwards.

Die partielle Ableitung von $H(j, t)$ nach $\vec{\eta}(t + 1)$ ergibt sich als:

$$\frac{\partial H(j, t)}{\partial \vec{\eta}(t + 1)} = \frac{\partial}{\partial \vec{\eta}(t + 1)} \left(e^{-r \cdot (t - t_0 + T_j)} (q^f(x_j, t + T_j) - \tilde{q}(x_j, t + T_j)) \right).$$

Da $q^f(x_j, t + T_j)$ unabhängig von $\vec{\eta}(t + 1)$ ist, bleibt:

$$\frac{\partial H(j, t)}{\partial \vec{\eta}(t + 1)} = -e^{-r \cdot (t - t_0 + T_j)} \cdot \frac{\partial \tilde{q}(x_j, t + T_j)}{\partial \vec{\eta}(t + 1)}.$$

Die Ableitung von $\tilde{q}(x_j, t + T_j)$ wird durch die Kettenregel weiter zerlegt:

$$\frac{\tilde{q}(x_j, t + T_j)}{\partial \vec{\eta}(t + 1)} = \frac{\partial \tilde{q}(x_j, t + T_j)}{\partial \tilde{y}(x_j, t + T_j)} \cdot \frac{\partial \tilde{y}(x_j, t + T_j)}{\partial \vec{\eta}(t + 1)}.$$

Die partielle Ableitung von $\tilde{q}(x_j, t + T_j)$ nach $\tilde{y}(x_j, t + T_j)$ hängt vom gewählten Mortalitätsmodell ab:

1. Im Lee-Carter-Modell ($y(x, t) = \log(m(x, t))$):

$$\frac{\partial \tilde{q}(x_j, t + T_j)}{\partial \tilde{y}(x_j, t + T_j)} = \tilde{p}(x_j, t + T_j) \cdot \tilde{m}(x_j, t + T_j),$$

2. Im Cairns-Blake-Dowd-Modell ($y(x, t) = \log(q(x, t)/(1 - q(x, t)))$):

$$\frac{\partial \tilde{q}(x_j, t + T_j)}{\partial \tilde{y}(x_j, t + T_j)} = \tilde{p}(x_j, t + T_j) \cdot \tilde{q}(x_j, t + T_j).$$

Um die zweite partielle Ableitung $\frac{\partial \tilde{y}(x_j, t + T_j)}{\partial \vec{\eta}(t + 1)}$ zu bestimmen, nutzen wir die Darstellung von $\tilde{y}(x_j, t + T_j)$ in Abhängigkeit von $\vec{\eta}(t + 1)$:

$$\tilde{y}(x_j, t + T_j) = d(x_j) + B(x_j, \cdot) \left(A^{t_0 + s - t} \vec{\alpha}(t) + \sum_{w=1}^{T_j} A^{T_j - w} \vec{\eta}(t + w) \right).$$

Die Ableitung ergibt:

$$\frac{\partial \tilde{y}(x_j, t + T_j)}{\partial \vec{\eta}(t + 1)} = B(x_j, \cdot) A^{T_j - 1}.$$

Der Ausdruck für $\frac{\partial H(j, t)}{\partial \vec{\eta}(t + 1)}$ lautet damit:

$$\frac{\partial H(j, t)}{\partial \vec{\eta}(t + 1)} = -e^{-r \cdot (t - t_0 + T_j)} \cdot \frac{\partial \tilde{q}(x_j, t + T_j)}{\partial \tilde{y}(x_j, t + T_j)} \cdot B(x_j, \cdot) A^{T_j - 1},$$

wobei die partielle Ableitung auf der rechten Seite der Gleichung vom zugrunde liegenden Mortalitätsmodell abhängig ist.

5.2 Momente der Skew-t Verteilung

Um die Werte von M_t, V_t, S_t, K_t herzuleiten, ist es ebenfalls notwendig, die Momente der multivariaten Skew-t-Verteilung zu berechnen. Diese Verteilung wurde von Azzalini und Capitanio eingeführt. Zur Erinnerung verwenden wir die Notation

$$\vec{Y} \sim St_n(\vec{\xi}, \Omega, \vec{\zeta}, \nu),$$

um einen n -dimensionalen Zufallsvektor zu beschreiben, der einer multivariaten Skew-t-Verteilung folgt. Dabei sind:

- $\vec{\xi}$: der Location-Parameter,
- Ω : der Dispersionsparameter,
- $\vec{\zeta}$: der Schiefeparameter,
- ν : die Freiheitsgrade der Verteilung.

Obwohl Azzalini und Capitanio die Darstellung der Momente beschrieben haben, gehen sie dabei von $\vec{\xi} = \vec{0}$ aus. In unserem Kontext ist diese Annahme jedoch nicht zwingend erfüllt. Daher verwenden wir zunächst die Berechnungen von Azzalini und Capitanio für die Momente unter der Voraussetzung $\vec{\xi} = \vec{0}$, um anschließend die entsprechenden Momente für $\vec{\xi} \neq \vec{0}$ herzuleiten (vgl. Azzalini & Capitanio, 2003).

Die Berechnung der Momente der multivariaten Skew-t-Verteilung für den Fall $\vec{\xi} \neq \vec{0}$ wird hier aus Platzgründen in verkürzter Form dargestellt. Die detaillierten Herleitungen und Vereinfachungen sind in Appendix A bei Siu et al. (2023) angegeben. Unter der Annahme $\vec{\xi} = \vec{0}$ kann der Zufallsvektor \vec{Y} wie folgt dargestellt werden:

$$\vec{Y} = \vec{\xi} + W^{-1/2} \vec{Z},$$

wobei

- $W = \chi_\nu^2 / \nu$ eine Zufallsvariable mit einer χ^2 -Verteilung und ν Freiheitsgraden ist,
- \vec{Z} ein von W unabhängiger Zufallsvektor ist, der einer multivariaten Skew-Normalverteilung folgt mit
 - Location-Parameter $\vec{\xi} = 0$
 - Dispersionsparameter Ω
 - Schiefeparameter $\vec{\zeta}$.

Sei nun $\mathbb{E}^{(m)}(\vec{X})$ das m -te Moment eines generischen Zufallsvektors \vec{X} . Für den Fall $\vec{\xi} = 0$ lassen sich die Momente von \vec{Y} durch folgende Beziehung ausdrücken:

$$\mathbb{E}^{(m)}(\vec{Y}) = \mathbb{E}(W^{-m/2}) \mathbb{E}^{(m)}(\vec{Z}),$$

wobei

$$\mathbb{E}(W^{-m/2}) = \frac{(\nu/2)^{m/2} \Gamma((\nu - m)/2)}{\Gamma(\nu/2)}$$

und Γ die Gammafunktion bezeichnet.

Für die ersten vier Momente von \vec{Z} gilt:

- Das erste Moment:

$$\mathbb{E}^{(1)}(\vec{Z}) = \mathbb{E}(\vec{Z}) = \sqrt{\frac{2}{\pi}} \vec{\delta}.$$

- Das zweite Moment:

$$\mathbb{E}^{(2)}(\vec{Z}) = \mathbb{E}(\vec{Z}\vec{Z}') = \Omega.$$

- Das dritte Moment:

$$\mathbb{E}^{(3)}(\vec{Z}) = \mathbb{E}(\vec{Z}\vec{Z}' \otimes \vec{Z}') = \sqrt{\frac{2}{\pi}} \left(\vec{\delta} \otimes \Omega + \text{vec}(\Omega)\vec{\delta}' + (I_n \otimes \vec{\delta})\Omega - \vec{\delta} \otimes \vec{\delta}' \otimes \vec{\delta} \right).$$

- Das vierte Moment:

$$\mathbb{E}^{(4)}(\vec{Z}) = \mathbb{E}(\vec{Z}\vec{Z}' \otimes \vec{Z}' \otimes \vec{Z}') = \text{vec}(\Omega)' \otimes \Omega + \Omega(I_n \otimes \text{vec}(\Omega)) + \Omega I_n (\text{vec}(\Omega)' \otimes I_n) (I_n \otimes U).$$

Hierbei ist

$$\vec{\delta} = \frac{\Lambda \bar{\Omega} \vec{\xi}}{(1 + \vec{\zeta}' \bar{\Omega} \vec{\zeta})^{1/2}},$$

wobei $\bar{\Omega}$ die Korrelationsmatrix von Ω darstellt. Das (i, j) -te Element von $\bar{\Omega}$ wird durch $\omega_{i,j} / \sqrt{\omega_{i,i}\omega_{j,j}}$ berechnet und mit \otimes wird das Kronecker-Produkt bezeichnet.

Weitere verwendete Größen sind:

- I_n : die $n \times n$ -Identitätsmatrix,
- U : eine $n^2 \times n^2$ -Matrix, die die Permutationsmatrix, die einer $n \times n$ Matrix zugeordnet ist, darstellt.

Die Permutationsmatrix U ist definiert durch die Eigenschaft, dass für eine generische Matrix X gilt:

$$\text{vec}(X') = U \text{vec}(X).$$

Diese Momente dienen als Grundlage für die Herleitung der entsprechenden Momente von \vec{Y} im Fall $\vec{\xi} \neq 0$. (vgl. Azzalini & Capitanio, 2003)

Nachdem wir die Momente für den Fall $\vec{\xi} = 0$ dargestellt haben, sollen nun die Momente für $\vec{\xi} \neq 0$ hergeleitet werden. Für das erste Moment ergibt sich:

$$\mathbb{E}^{(1)}(\vec{Y}) = \vec{\xi} + \mathbb{E}(W^{-1/2})\mathbb{E}(\vec{Z}).$$

Durch Einsetzen der entsprechenden Ausdrücke erhalten wir:

$$\mathbb{E}^{(1)}(\vec{Y}) = \vec{\xi} + \sqrt{\frac{\nu}{\pi}} \frac{\Gamma((\nu - 1)/2)}{\Gamma(\nu/2)} \vec{\delta}.$$

Anschließend betrachten wir das zweite Moment von \vec{Y} im Fall $\vec{\xi} \neq 0$:

$$\mathbb{E}^{(2)}(\vec{Y}) = \mathbb{E} \left[\left(\vec{Y} - \mathbb{E}(\vec{Y}) \right) \left(\vec{Y} - \mathbb{E}(\vec{Y}) \right)' \right].$$

Durch schrittweises Ersetzen und Umformen ergibt sich:

$$\mathbb{E}^{(2)}(\vec{Y}) = \mathbb{E}(W^{-1})\mathbb{E}(\vec{Z}\vec{Z}') - \mathbb{E}(W^{-1/2})^2\mathbb{E}(\vec{Z})\mathbb{E}(\vec{Z})'.$$

Durch Einsetzen der Momente für $\vec{\xi} = 0$ lautet das Ergebnis:

$$\mathbb{E}^{(2)}(\vec{Y}) = \frac{\nu}{\nu - 2}\Omega - \frac{\nu}{\pi} \left(\frac{\Gamma((\nu - 1)/2)}{\Gamma(\nu/2)} \right)^2 \frac{\Lambda\bar{\Omega}\vec{\zeta}\vec{\zeta}'\bar{\Omega}\Lambda}{1 + \vec{\zeta}'\bar{\Omega}\vec{\zeta}}.$$

Für das dritte Moment von \vec{Y} ergibt sich:

$$\mathbb{E}^{(3)} \left(\vec{Y} - \mathbb{E}(\vec{Y}) \right) = \mathbb{E} \left[\left(\vec{Y} - \mathbb{E}(\vec{Y}) \right) \left(\vec{Y} - \mathbb{E}(\vec{Y}) \right)' \otimes \left(\vec{Y} - \mathbb{E}(\vec{Y}) \right)' \right].$$

Nach Einsetzen und Vereinfachungen erhalten wir:

$$\begin{aligned} \mathbb{E}^{(3)}(\vec{Y}) &= \mathbb{E} \left(W^{-\frac{3}{2}} \right) \mathbb{E} \left(\vec{Z}\vec{Z}' \otimes \vec{Z}' \right) + 2 \left(\mathbb{E} \left(W^{-\frac{1}{2}} \right) \right)^3 \left(\mathbb{E}(\vec{Z})\mathbb{E}(\vec{Z})' \otimes \mathbb{E}(\vec{Z})' \right) \\ &\quad - \mathbb{E} \left(W^{-1} \right) \mathbb{E} \left(W^{-\frac{1}{2}} \right) \left(\mathbb{E} \left(\mathbb{E}(\vec{Z})\vec{Z}' \otimes \vec{Z}' \right) + \mathbb{E} \left(\vec{Z}\mathbb{E}(\vec{Z})' \otimes \vec{Z}' \right) + \mathbb{E} \left(\vec{Z}\vec{Z}' \otimes \mathbb{E}(\vec{Z})' \right) \right). \end{aligned}$$

Im weiteren Verlauf werden zusätzliche Terme benötigt, um das dritte Moment von \vec{Y} auszudrücken. Dabei gelten die folgenden Relationen:

$$\begin{aligned} \mathbb{E} \left(\mathbb{E}(\vec{Z})\vec{Z}' \otimes \vec{Z}' \right) &= \mathbb{E}(\vec{Z}) \otimes \text{vec} \left(\mathbb{E} \left(\vec{Z}\vec{Z}' \right) \right)', \\ \mathbb{E} \left(\vec{Z}\mathbb{E}(\vec{Z})' \otimes \vec{Z}' \right) &= \mathbb{E} \left(\vec{Z}' \right) \otimes \mathbb{E} \left(\vec{Z}\vec{Z}' \right), \\ \mathbb{E} \left(\vec{Z}\vec{Z}' \otimes \mathbb{E}(\vec{Z})' \right) &= \mathbb{E} \left(\vec{Z}\vec{Z}' \right) \otimes \mathbb{E}(\vec{Z})'. \end{aligned}$$

Mit diesen Zusammenhängen lässt sich das dritte Moment von \vec{Y} schließlich wie folgt formulieren:

$$\begin{aligned} \mathbb{E}^{(3)}(\vec{Y} - \mathbb{E}(\vec{Y})) &= \mathbb{E} \left(W^{-\frac{3}{2}} \right) \mathbb{E} \left(\vec{Z}\vec{Z}' \otimes \vec{Z}' \right) + 2 \left(\mathbb{E} \left(W^{-\frac{1}{2}} \right) \right)^3 \left(\mathbb{E}(\vec{Z})\mathbb{E}(\vec{Z})' \otimes \mathbb{E}(\vec{Z})' \right) \\ &\quad - \mathbb{E} \left(W^{-1} \right) \mathbb{E} \left(W^{-\frac{1}{2}} \right) \left(\mathbb{E}(\vec{Z}) \otimes \text{vec} \left(\mathbb{E} \left(\vec{Z}\vec{Z}' \right) \right)' \right) \\ &\quad + \mathbb{E} \left(\vec{Z}' \right) \otimes \mathbb{E} \left(\vec{Z}\vec{Z}' \right) + \mathbb{E} \left(\vec{Z}\vec{Z}' \right) \otimes \mathbb{E}(\vec{Z})'. \end{aligned}$$

Die Terme $\mathbb{E}(W^{-\frac{1}{2}})$, $\mathbb{E}(W^{-1})$ und $\mathbb{E}(W^{-\frac{3}{2}})$ stammen dabei aus den bereits bekannten Formeln für die Momente der Verteilung von \vec{Y} im Fall $\vec{\xi} = 0$.

Abschließend wird das vierte Moment von \vec{Y} für den Fall $\vec{\xi} \neq 0$ betrachtet. Die Herleitung erfolgt analog zur Berechnung des dritten Moments, weshalb hier auf eine ausführliche Darstellung der Zwischenschritte verzichtet wird. Das vierte Moment ergibt sich zu

$$\begin{aligned} \mathbb{E}^{(4)}(\vec{Y} - \mathbb{E}(\vec{Y})) &= \mathbb{E} \left[(\vec{Y} - \mathbb{E}(\vec{Y}))(\vec{Y} - \mathbb{E}(\vec{Y}))' \otimes (\vec{Y} - \mathbb{E}(\vec{Y}))' \otimes (\vec{Y} - \mathbb{E}(\vec{Y}))' \right] \\ &= \mathbb{E}(W^{-2}) \mathbb{E}^{(4)}(\vec{Z}) - \mathbb{E}(W^{-\frac{3}{2}}) \mathbb{E}(W^{-\frac{1}{2}}) K_1 \\ &\quad + \mathbb{E}(W^{-1}) \left(\mathbb{E}(W^{-\frac{1}{2}}) \right)^2 K_2 - 3 \left(\mathbb{E}(W^{-\frac{1}{2}}) \right)^4 K_3, \end{aligned}$$

wobei die Terme K_1, K_2 und K_3 wie folgt definiert sind:

$$\begin{aligned} K_1 &= \mathbb{E}(\vec{Z}) \text{vec} \left(\mathbb{E}^{(3)}(\vec{Z}) \right)' + \mathbb{E}(\vec{Z})' \otimes \mathbb{E}^{(3)}(\vec{Z}) + \mathbb{E}^{(3)}(\vec{Z}) \left(I_d \otimes \mathbb{E}(\vec{Z})' \otimes I_d \right) \\ &\quad + \mathbb{E}^{(3)}(\vec{Z}) \otimes \mathbb{E}(\vec{Z})', \\ K_2 &= \mathbb{E}(\vec{Z}) \mathbb{E}(\vec{Z})' \otimes \text{vec} \left(\mathbb{E}(\vec{Z}\vec{Z}') \right)' + \left(\mathbb{E}(\vec{Z}) \mathbb{E}(\vec{Z})' \right) \left(\mathbb{E}(\vec{Z}' \otimes I_d \otimes \vec{Z}') \right) \\ &\quad + \text{vec} \left(\mathbb{E}(\vec{Z}) \mathbb{E}(\vec{Z})' \right)' \otimes \mathbb{E}(\vec{Z}\vec{Z}') + \mathbb{E}(\vec{Z}) \otimes \text{vec} \left(\mathbb{E}(\vec{Z}\vec{Z}') \right)' \otimes \mathbb{E}(\vec{Z})' \\ &\quad + \mathbb{E}(\vec{Z}\vec{Z}') \left(\mathbb{E}(\vec{Z})' \otimes I_d \otimes \mathbb{E}(\vec{Z})' \right) + \mathbb{E}(\vec{Z}\vec{Z}') \otimes \mathbb{E}(\vec{Z})' \otimes \mathbb{E}(\vec{Z})', \\ K_3 &= \mathbb{E}(\vec{Z}) \mathbb{E}(\vec{Z})' \otimes \mathbb{E}(\vec{Z})' \otimes \mathbb{E}(\vec{Z})'. \end{aligned}$$

Die Terme $\mathbb{E}(W^{-\frac{1}{2}}), \mathbb{E}(W^{-1}), \mathbb{E}(W^{-\frac{3}{2}})$ und $\mathbb{E}(W^{-2})$ sowie $\mathbb{E}(\vec{Z}), \mathbb{E}(\vec{Z}\vec{Z}'), \mathbb{E}(\vec{Z}^{(3)})$ und $\mathbb{E}(\vec{Z}^{(4)})$ werden mithilfe der oben hergeleiteten Formeln für den Fall $\vec{\xi} = \vec{0}$ berechnet.

5.3 Analytische Herleitung der Momente M_t, V_t, S_t und K_t

In diesem Abschnitt werden die analytischen Formeln für M_t, V_t, S_t und K_t hergeleitet. Hierzu betrachten wir $V_l(t+1)$ und $V_h(t+1; j, t)$ als Näherungen für $V_L(t+1)$ bzw. $V_H(t+1; j, t)$ für $j = 1, \dots, m$. Der betrachtete Zufallsvektor lautet:

$$(V_l(t+1), V_h(t+1; 1, t), \dots, V_h(t+1; m, t))'$$

Unter der Annahme, dass das zugrunde liegende Mortalitätsmodell n Innovationskomponenten besitzt (entspricht der Länge des Innovationsvektors $\vec{\eta}(t)$ zu einem

Zeitpunkt t), lassen sich $V_l(t+1)$ und $V_h(t+1; j, t)$ wie folgt darstellen:

$$V_l(t+1) = \hat{L} + \sum_{k=1}^n (D_k(0, t+1) \cdot \eta_k(t+1))$$

und

$$V_h(t+1; j, t) = \hat{H}(j, t) + \sum_{k=1}^n (D_k(j, t+1) \cdot \eta_k(t+1)), \quad j = 1, \dots, m,$$

wobei die Ableitungen definiert sind als:

$$D_k(0, t+1) = \frac{\partial L}{\partial \eta_k(t+1)}$$

und

$$D_k(j, t+1) = \frac{\partial H(j, t)}{\partial \eta_k(t+1)}, \quad j = 1, \dots, m.$$

Wie in Kapitel 3 erläutert, wird für die Verteilung der Innovationsvektoren von einer Skew-t-Verteilung ausgegangen, das heißt, $\eta(t+1) \sim St_n(\vec{\xi}, \Omega, \vec{\zeta}, \nu)$. Mithilfe dieser Verteilungsannahme und der obigen Gleichungen können die ersten vier Momente des Zufallsvektors $(V_l(t+1), V_h(t+1; 1, t), \dots, V_h(t+1; m, t))'$ berechnet werden. Dies ermöglicht die Bestimmung der Werte für M_t, V_t, S_t und K_t .

Für die Bestimmung von M_t betrachten wir die bedingten Erwartungswerte $\mathbb{E}[V_l(t+1)|\mathcal{F}_t]$ und $\mathbb{E}[V_h(t+1; j, t)|\mathcal{F}_t]$. Basierend auf diesen bedingten Erwartungswerten wird M_t durch den Vektor $(\hat{L}, \hat{H}(1, t), \dots, \hat{H}(m, t))'$ approximiert. Die einzelnen Elemente dieses Vektors können dabei analytisch berechnet werden.

Für die Bestimmung von V_t approximieren wir die Werte von $\sigma_{ij,t}$ (dem (i, j) -ten Eintrag von V_t) mit

$$\begin{aligned} \sigma_{ij,t} &\approx \text{Cov} \left(\sum_{k_1=1}^n (D_{k_1}(i-1, t+1) \cdot \eta_{k_1}(t+1)), \sum_{k_2=1}^n (D_{k_2}(j-1, t+1) \cdot \eta_{k_2}(t+1)) \middle| \mathcal{F}_t \right) \\ &= \sum_{k_1=1}^n \sum_{k_2=1}^n D_{k_1}(i-1, t+1) \cdot D_{k_2}(j-1, t+1) \cdot \text{Cov}(\eta_{k_1}(t+1), \eta_{k_2}(t+1) | \mathcal{F}_t), \end{aligned}$$

für $i, j = 1, 2, \dots, m+1$. Dabei kann $\text{Cov}(\eta_{k_1}(t+1), \eta_{k_2}(t+1) | \mathcal{F}_t)$ analytisch bestimmt werden.

Die Approximation der Einträge $s_{iju,t}$ von S_t erfolgt mit

$$\begin{aligned}
s_{iju,t} &\approx \mathbb{E} \left[\left(\sum_{k_1=1}^n (D_{k_1}(i-1, t+1) \cdot \eta_{k_1}(t+1)) \right) \right. \\
&\quad \cdot \left(\sum_{k_2=1}^n (D_{k_2}(j-1, t+1) \cdot \eta_{k_2}(t+1)) \right) \\
&\quad \cdot \left. \left(\sum_{k_3=1}^n (D_{k_3}(u-1, t+1) \cdot \eta_{k_3}(t+1)) \right) \middle| \mathcal{F}_t \right] \\
&= \sum_{k_1=1}^n \sum_{k_2=1}^n \sum_{k_3=1}^n D_{k_1}(i-1, t+1) \cdot D_{k_2}(j-1, t+1) \cdot D_{k_3}(u-1, t+1) \\
&\quad \cdot \mathbb{E}(\eta_{k_1}(t+1)\eta_{k_2}(t+1)\eta_{k_3}(t+1) | \mathcal{F}_t),
\end{aligned}$$

für $i, j, u = 1, 2, \dots, m+1$. Dabei wird $\mathbb{E}(\eta_{k_1}(t+1)\eta_{k_2}(t+1)\eta_{k_3}(t+1) | \mathcal{F}_t)$ analytisch berechnet.

Die Approximation der Einträge $k_{ijuv,t}$ von K_t erfolgt mit

$$\begin{aligned}
k_{ijuv,t} &\approx \mathbb{E} \left[\left(\sum_{k_1=1}^n (D_{k_1}(i-1, t+1) \cdot \eta_{k_1}(t+1)) \right) \right. \\
&\quad \cdot \left(\sum_{k_2=1}^n (D_{k_2}(j-1, t+1) \cdot \eta_{k_2}(t+1)) \right) \\
&\quad \cdot \left(\sum_{k_3=1}^n (D_{k_3}(u-1, t+1) \cdot \eta_{k_3}(t+1)) \right) \\
&\quad \cdot \left. \left(\sum_{k_4=1}^n (D_{k_4}(v-1, t+1) \cdot \eta_{k_4}(t+1)) \right) \middle| \mathcal{F}_t \right] \\
&= \sum_{k_1=1}^n \sum_{k_2=1}^n \sum_{k_3=1}^n \sum_{k_4=1}^n D_{k_1}(i-1, t+1) \cdot D_{k_2}(j-1, t+1) \cdot D_{k_3}(u-1, t+1) \\
&\quad \cdot D_{k_4}(v-1, t+1) \cdot \mathbb{E}(\eta_{k_1}(t+1)\eta_{k_2}(t+1)\eta_{k_3}(t+1)\eta_{k_4}(t+1) | \mathcal{F}_t),
\end{aligned}$$

für $i, j, u, v = 1, 2, \dots, m+1$. Dabei wird $\mathbb{E}(\eta_{k_1}(t+1)\eta_{k_2}(t+1)\eta_{k_3}(t+1)\eta_{k_4}(t+1) | \mathcal{F}_t)$ analytisch berechnet.

6 Numerisches Beispiel

In diesem Abschnitt werden exemplarische Anwendungsfälle des zuvor beschriebenen Modells präsentiert. Ziel ist es, sowohl die Stärken als auch die Schwächen des Modells aufzuzeigen. Nach umfangreichen Tests wurde deutlich, dass die Leistungsfähigkeit des Modells stark von den zugrunde liegenden Daten abhängt. So treten beispielsweise Schwierigkeiten bei der Anwendung auf historische Daten aus Österreich auf, während das Modell bei der Analyse von Daten aus England und Irland sehr gute Ergebnisse liefert.

6.1 Dataset

Die für unsere Analysen verwendeten Daten stammen aus der Human Mortality Database (HMD), einer der weltweit führenden wissenschaftlichen Datenquellen zur Sterblichkeit in Industrieländern. Die HMD bietet freien Zugang zu umfassenden und qualitätsgeprüften Datensätzen, was sie zu einer idealen Grundlage für diese Untersuchung macht.

Um die Robustheit und Datenabhängigkeit des Modells zu prüfen, wurde es auf Datensätze aus verschiedenen Ländern angewendet, darunter England und Wales, Österreich, Japan, Norwegen, Irland und Kanada. So konnte das Modell unter unterschiedlichen demografischen Bedingungen getestet und bewertet werden.

Im Fokus unserer Analyse stehen vorwiegend die höheren Altersgruppen, da diese im Kontext dieser Arbeit eine besondere Relevanz aufweisen. Konkret werden Altersdaten im Bereich von $x_a = 60$ bis $x_b = 89$ betrachtet. Der untersuchte Zeitraum erstreckt sich dabei über die Jahre $t_a = 1950$ bis $t_b = 2016$. Diese Auswahl gewährleistet eine detaillierte Untersuchung der Sterblichkeitsmuster im höheren Alter über einen ausreichend langen Zeitraum.

6.2 Auswahl des Mortalitätsmodelles

Die beiden stochastischen Mortalitätsmodelle, die für die Anwendung in unserem Ansatz in Frage kommen, sind das Cairns-Blake-Dowd-Modell und das Lee-Carter-Modell. Obwohl diese Modelle nicht ineinander verschachtelt sind, werden sie anhand der vorliegenden Datensätze miteinander verglichen, um ihre jeweilige Eignung zu beurteilen.

Beide Modelle besitzen eine fundierte theoretische Grundlage und weisen spezifische Stärken auf, die ihre Berücksichtigung rechtfertigen. Dennoch fiel die Wahl auf

das Cairns-Blake-Dowd-Modell, da es im Kontext dieser Arbeit nicht nur besser interpretierbare Ergebnisse liefert, sondern sich auch in den statistischen Tests als überlegen erwiesen hat.

Für den Vergleich der Modelle wurden verschiedene Metriken zur Bewertung der Anpassungsqualität berechnet, darunter der Mean Absolute Error (MAE), der Root Mean Square Error (RMSE) und der Mean Absolute Percentage Error (MAPE). Zusätzlich wurden die Modelle anhand gängiger Informationskriterien, nämlich dem Akaike-Informationskriterium (AIC) und dem Bayes-Informationskriterium (BIC), evaluiert. Die Ergebnisse dieser Metriken bestätigten die Entscheidung zugunsten des Cairns-Blake-Dowd-Modells.

Zur weiteren Erläuterung der Bewertungsmethoden sollen die verwendeten Formeln im Detail beschrieben werden. Dabei sei y_i der tatsächliche Wert, \hat{y}_i der geschätzte Wert und n die Anzahl der Beobachtungen.

Der Mean Absolute Error (MAE) berechnet den durchschnittlichen absoluten Fehler zwischen den tatsächlichen und den vorhergesagten Werten und wird wie folgt definiert:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|.$$

Der Root Mean Squared Error (RMSE) berechnet die Quadratwurzel des durchschnittlichen quadratischen Fehlers. Im Vergleich zum MAE betont der RMSE größere Fehler stärker, da die Fehler quadriert werden:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}.$$

Der Mean Absolute Percentage Error (MAPE) misst den durchschnittlichen prozentualen Fehler und bietet eine relative Bewertung der Modellgenauigkeit im Verhältnis zu den tatsächlichen Werten:

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \cdot 100.$$

Zusätzlich wird der adjusted R-Squared betrachtet, der den Anteil der Varianz beschreibt, den das Modell erklärt, wobei er die Anzahl der im Modell verwendeten Parameter berücksichtigt. Ein höherer Wert deutet auf eine bessere Modellanpassung hin. Der adjusted R-Squared wird definiert als:

$$\text{Adjusted } R^2 = 1 - \left(\frac{(1 - R^2) \cdot (n - 1)}{n - k - 1} \right),$$

wobei n die Anzahl der Beobachtungen, k die Anzahl der unabhängigen Variablen

im Modell und R^2 der Bestimmtheitsmaß ist:

$$R^2 = 1 - \frac{\text{SSR}}{\text{SST}},$$

mit SSR (Sum of Squared Residuals) als Summe der quadrierten Abweichungen zwischen Modellvorhersagen und tatsächlichen Werten und SST (Total Sum of Squares) als Summe der quadrierten Abweichungen der tatsächlichen Werte vom Durchschnittswert der tatsächlichen Werte.

Zuletzt werden die Informationskriterien Akaike-Informationskriterium (AIC) und Bayesian-Informationskriterium (BIC) beschrieben, die zur Bewertung der Modellgüte unter Berücksichtigung der Modellkomplexität verwendet werden. Hierbei sind die absoluten Werte von AIC und BIC für sich genommen nicht aussagekräftig; entscheidend ist der Vergleich zwischen Modellen, wobei ein niedrigerer Wert auf ein besseres Modell hinweist:

$$\text{AIC} = 2k - 2\log(\tilde{L}), \quad \text{BIC} = k \cdot \log(n) - 2\log(\tilde{L}),$$

wobei \tilde{L} die maximale Likelihood des Modells darstellt und k die Anzahl der geschätzten Parameter im Modell ist. (vgl. Hastie et al., 2000)

Die oben genannten statistischen Bewertungsmethoden wurden auf verschiedenen Datensätzen angewandt, um die Güte und Leistungsfähigkeit der Modelle zu vergleichen. Zur Illustration werden die Ergebnisse anhand der Daten aus Österreich präsentiert:

Die oben genannten statistischen Methoden wurden für diverse Datensätze berechnet, als Beispiel werden jene von den Daten aus Österreich illustriert.

Statistische Methode	Lee-Carter-Modell	Cairns-Blake-Dowd-Modell
MAPE	32,3137	4,3164
RMSE	0,0373	0,0050
MAE	0,0252	0,0031
adjusted R^2	0,7253	0,9938
AIC	- 7738,20	- 15621,16
BIC	- 7721,38	- 15609,95

Tabelle 3: Ergebnisse verschiedener statistischer Methoden anhand der Daten aus Österreich

Die Tabelle zeigt, dass das Cairns-Blake-Dowd-Modell in allen betrachteten statistischen Kriterien signifikant bessere Ergebnisse liefert als das Lee-Carter-Modell. Insbesondere beim MAPE, RMSE und MAE weist das Cairns-Blake-Dowd-Modell erheblich geringere Fehler auf, was auf eine genauere Anpassung an die beobachteten Daten hindeutet. Auch die Informationskriterien (AIC und BIC) sowie das angepas-

te Bestimmtheitsmaß (R^2) sprechen klar für das Cairns-Blake-Dowd-Modell, da es sowohl eine höhere Modellgüte als auch eine bessere Erklärung der Daten liefert.

6.3 Parameter

Bei der Anwendung des Cairns-Blake-Dowd-Modells müssen zwei zeitabhängige Parameter, $\kappa_t^{(1)}$ und $\kappa_t^{(2)}$, geschätzt werden. Die Details zu diesen Parametern wurden in Kapitel 2 ausführlich beschrieben. In diesem Abschnitt werden die Schätzungen anhand der österreichischen Daten illustriert. Für $\kappa_t^{(1)}$ ergibt sich die folgende Abbildung:

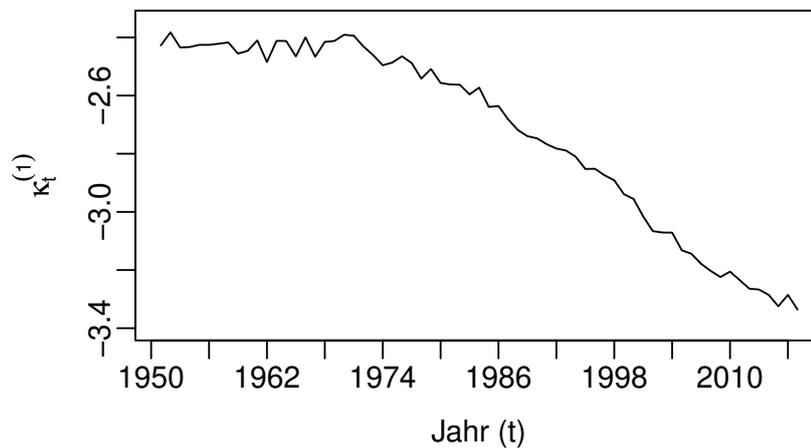


Abbildung 6.1: Schätzung von $\kappa_t^{(1)}$ basierend auf den österreichischen Daten

Für die Schätzungen von $\kappa_t^{(2)}$ erhalten wir die folgende Abbildung:

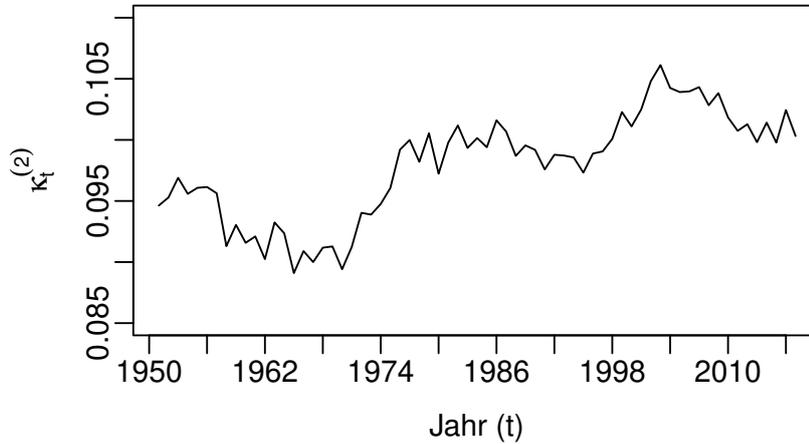


Abbildung 6.2: Schätzung von $\kappa_t^{(2)}$ basierend auf den österreichischen Daten

Zur Prognose der zukünftigen Werte von $\kappa_t^{(1)}$ und $\kappa_t^{(2)}$ wird ein bivariater Random Walk mit Drift verwendet. Die Driftterme werden dabei aus den Schätzungen der Parameter berechnet und mit $c^{(i)}$ für $i = 1, 2$ bezeichnet. Für die österreichischen Daten ergeben sich folgende Werte:

$$c^{(1)} = -0,0154 \text{ und } c^{(2)} = 0,00018.$$

Diese Driftterme beschreiben die langfristigen Trends in den zeitlichen Entwicklungen der Parameter und sind ein essenzieller Bestandteil für die Modellierung der zukünftigen Mortalität.

6.4 Überprüfung der Verteilungsannahmen

Im weiteren Verlauf analysieren wir den Innovationsvektor $\vec{\eta}(t)$, welcher ein zentraler Bestandteil des Modells ist. Hier könnten erste Schwierigkeiten auftreten, da angenommen wird, dass dieser Vektor einer Skew-t-Verteilung folgt. Mit den in Kapitel 3.2 beschriebenen Formeln lassen sich die Werte von $\vec{\eta}(t) = (\eta_1(t), \eta_2(t))'$ berechnen. Die relevanten Notationen sind:

$$\vec{\alpha}(t) = \begin{pmatrix} \kappa_t^{(1)} \\ \kappa_t^{(2)} \end{pmatrix}, \quad \vec{c} = \begin{pmatrix} c_t^{(1)} \\ c_t^{(2)} \end{pmatrix}, \quad \vec{\eta}(t) = \begin{pmatrix} \eta_1(t) \\ \eta_2(t) \end{pmatrix}, \quad A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Die zugrunde liegende Übergangsgleichung ist:

$$\vec{\alpha}(t) = \vec{c} + A\vec{\alpha}(t-1) + \vec{\eta}(t),$$

Um die historischen Innovationsvektoren $\vec{\eta}(t)$ zu berechnen, lösen wir die Gleichung nach $\vec{\eta}(t)$ auf:

$$\vec{\eta}(t) = \vec{\alpha}(t) - \vec{c} - A\vec{\alpha}(t-1).$$

Auf Basis unserer Schätzungen erhalten wir zwei Vektoren der Länge 66 für $\eta_1(t)$ und $\eta_2(t)$, deren Verteilung im Folgenden untersucht wird.

Viele Modelle gehen davon aus, dass $\vec{\eta}(t)$ einer Normalverteilung folgt. Daher überprüfen wir zunächst, ob eine Normalverteilung plausibel ist. Darüber hinaus wird geprüft, ob eine Skew-t-Verteilung in Betracht gezogen werden kann.

Für die Überprüfung der Normalverteilung jedes einzelnen Vektorelements verwenden wir den Jarque-Bera-Test. Der Jarque-Bera-Test ist ein statistischer Test zur Überprüfung der Normalverteilungsannahme von Residuen oder anderen Daten. Er wurde von Carlos M. Jarque und Anil K. Bera entwickelt und nutzt das Prinzip des Lagrange-Multiplikators (LM). Die Grundidee des Jarque-Bera-Test, dass normalverteilte Daten eine Schiefe von 0 und eine Kurtosis von 3 haben. Der Jarque-Bera-Test prüft, ob die Verteilung der Daten signifikant von diesen Werten abweicht, indem er Schiefe und Kurtosis gemeinsam betrachtet.

Die Teststatistik ist definiert als:

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

mit

- n : Stichprobengröße,
- S : Schiefe und
- K : Kurtosis.

Die Teststatistik ist asymptotisch Chi-Quadrat-verteilt mit zwei Freiheitsgraden (χ_2^2) und die Nullhypothese beschreibt ein Vorliegen einer Normalverteilung. (vgl. Jarque & Bera, 1980)

Die Berechnung erfolgt mit der Funktion `jarque.bera.test()` aus dem R-Package `tseries`. Der Test liefert einen p-Wert, welcher mit einem Signifikanzniveau von 0,05 verglichen wird, um die Hypothese der Normalverteilung zu akzeptieren oder abzulehnen.

Um die gemeinsame Verteilung der Innovationsvektoren zu überprüfen, führen wir weiters einen multivariaten Test durch. Dafür verwenden wir den Henze-Zirkler-Test,

welcher mit der Funktion `mvn()` aus dem R-Package `MVN` implementiert ist. Dieser Test erlaubt die Untersuchung, ob die Multivarianzannahme aufrechterhalten werden kann. Er wurde von Henze und Zirkler (1990) entwickelt und die Grundidee ist die Prüfung, ob eine gegebene Stichprobe mit einer multivariaten Normalverteilung konsistent ist. (vgl. Henze & Zirkler, 1990)

Die Teststatistik ist gegeben mit:

$$HZ = n \left(\frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n \phi(\|X_i - X_j\|) - \frac{2}{n} \sum_{i=1}^n \mathbb{E}[\phi(\|X_i - Z\|)] \mathbb{E}[\phi(\|Z_1 - Z_2\|)] \right),$$

wobei:

- X_i, X_j die Beobachtungen der Stichprobe ($n \times p$ -Matrix) sind,
- Z, Z_1, Z_2 multivariat normalverteilte Zufallsvektoren sind,
- $\phi(r) = e^{-\gamma r^2}$ eine Funktion der Distanz r ist, mit $\gamma > 0$,
- n die Stichprobengröße ist.

Die Nullhypothese beschreibt ein Vorliegen einer multivariaten Normalverteilung und der Test ist asymptotisch Chi-Quadrat-verteilt, das bedeutet, die Teststatistik wird mit einem kritischen Wert der Chi-Quadrat-Verteilung verglichen. (vgl. Korkmaz et al., 2014)

Die Ergebnisse der durchgeführten Tests zur Überprüfung der Normalverteilung zeigen eine starke Abhängigkeit von den zugrunde liegenden Daten. Bei den Daten aus Österreich ist die Annahme einer Normalverteilung beispielsweise plausibel.

Für den univariaten Jarque-Bera-Test ergeben sich folgende p-Werte:

- Für $\eta_1(t)$: $p = 0,3050$
- Für $\eta_2(t)$: $p = 0,4643$

Da beide p-Werte über dem Signifikanzniveau von 0,05 liegen, kann die Nullhypothese, dass die Innovationsvektoren $\eta_1(t)$ und $\eta_2(t)$ normalverteilt sind, nicht abgelehnt werden.

Auch der multivariate Henze-Zirkler-Test bestätigt dieses Ergebnis mit einem p-Wert von $p = 0,7396$. Somit ist auch die Hypothese, dass die Innovationsvektoren gemeinsam einer Normalverteilung folgen, nicht abzulehnen.

Ähnliche Ergebnisse lassen sich bei den Daten aus Norwegen und Kanada beobachten. In diesen Fällen liefert der Jarque-Bera-Test ebenfalls p-Werte, die auf eine Normalverteilung hinweisen, und auch der Henze-Zirkler-Test unterstützt diese Annahme.

Für andere Datensätze zeigt sich, dass die Annahme einer Normalverteilung nicht haltbar ist. Ein Beispiel hierfür sind die Daten aus England und Wales.

Die Ergebnisse des Jarque-Bera-Tests liefern folgende p-Werte:

- Für $\eta_1(t)$: $p = 0,0063$
- Für $\eta_2(t)$: $p = 0,000007188$

Da beide p-Werte deutlich unter dem Signifikanzniveau von 0,05 liegen, ist die Nullhypothese, dass die Innovationsvektoren $\eta_1(t)$ und $\eta_2(t)$ normalverteilt sind, abzulehnen.

Auch der multivariate Henze-Zirkler-Test unterstützt dieses Ergebnis, da der p-Wert von $p = 0,0275$ ebenfalls unterhalb der Schwelle von 0,05 liegt. Folglich muss die Annahme, dass die Innovationsvektoren einer Normalverteilung folgen, für die Daten aus England und Wales verworfen werden.

Ähnliche Resultate lassen sich bei den Daten aus Irland und Japan beobachten. In diesen Fällen sind die p-Werte der Tests ebenfalls so gering, dass die Normalverteilung nicht angenommen werden kann.

Diese Ergebnisse verdeutlichen, dass die Annahme einer Normalverteilung stark von den verwendeten Datensätzen abhängt. Für Datensätze wie die aus England und Wales, Irland oder Japan scheint eine alternative Verteilung, wie etwa die Skew-t-Verteilung, besser geeignet zu sein. Dies betont die Bedeutung, die Verteilungsannahmen individuell an die jeweiligen Daten anzupassen, um realistische und verlässliche Ergebnisse zu erzielen.

Bei den Datensätzen, die die Hypothese der Normalverteilung nicht verwerfen, treten erhebliche Probleme bei der Schätzung der Parameter der Skew-t-Verteilung auf, insbesondere hinsichtlich der Freiheitsgrade. Die Schätzung der Parameter erfolgt mit der Funktion `fit_skew_t` aus dem R-Package `sn`. Zum Vergleich werden die geschätzten Parameter für einen Datensatz präsentiert, bei dem die Normalverteilungsannahme verworfen wird, sowie für einen Datensatz, der diese Annahme nicht verwirft. Hierfür ziehen wir die Daten aus England und Wales sowie aus Österreich heran.

Parameter	England und Wales	Österreich
ξ	$\begin{pmatrix} -0,0024883115 \\ 0,0009727239 \end{pmatrix}$	$\begin{pmatrix} 1,149097 \times 10^{-3} \\ -9,943992 \times 10^{-5} \end{pmatrix}$
Ω	$\begin{pmatrix} 7,321501 \times 10^{-4} & 1,684717 \times 10^{-5} \\ 1,684717 \times 10^{-5} & 2,447596 \times 10^{-6} \end{pmatrix}$	$\begin{pmatrix} 1,033583 \times 10^{-3} & 1,768814 \times 10^{-5} \\ 1,768814 \times 10^{-5} & 2,799492 \times 10^{-6} \end{pmatrix}$
ζ	$\begin{pmatrix} 1,051722 \\ -1,997987 \end{pmatrix}$	$\begin{pmatrix} 1,933445e - 02 \\ -2,507496e - 06 \end{pmatrix}$
ν	5,391	12883,33

Tabelle 4: Vergleich der Schätzungen der Parameter der Skew-t-Verteilung

Beim Vergleich der Ergebnisse zeigen sich deutliche Unterschiede zwischen den beiden Datensätzen. Besonders auffällig sind die Werte des Schiefeparameters ζ sowie der Freiheitsgrade ν . Die Daten aus England und Wales weisen eine klare Schiefe auf, was mit den Annahmen unseres Modells übereinstimmt. Im Gegensatz dazu liegt der Schiefeparameter ζ für die Daten aus Österreich nahe bei Null, was eine starke Annäherung an die Normalverteilung nahelegt.

Auch die geschätzten Freiheitsgrade unterstützen diese Beobachtung. Für die Daten aus Österreich nehmen die Freiheitsgrade mit $\nu = 12.883,33$ einen sehr hohen Wert an, wodurch sich die Skew-t-Verteilung an die Skew-Normalverteilung annähert. Diese hohen Freiheitsgrade und die fehlende Schiefe deuten darauf hin, dass die Annahme einer Normalverteilung für die Daten aus Österreich gerechtfertigt ist. Infolgedessen eignen sich diese Daten weniger für unser Modell, welches explizit von einer Abweichung von der Normalverteilung ausgeht.

Im Gegensatz dazu sind die Freiheitsgrade für die Daten aus England und Wales mit $\nu = 5,391$ vergleichsweise niedrig, was die Annahme einer Skew-t-Verteilung unterstützt. Die ausgeprägte Schiefe in diesen Daten unterstreicht zusätzlich die Relevanz unseres Modellansatzes.

Die Analyse zeigt, dass die Wahl der zugrunde liegenden Daten für die Modellierung von entscheidender Bedeutung ist. Aufgrund der unterschiedlichen Charakteristika der Datensätze lassen sich bestimmte Verteilungsannahmen nicht universell treffen. Für den weiteren Verlauf dieser Arbeit konzentrieren wir uns daher auf die Daten aus England und Wales, Japan und Irland. Diese Datensätze erfüllen die Voraussetzungen unseres Modells, da sie eine deutliche Abweichung von der Normalverteilung und eine erkennbare Schiefe aufweisen.

6.5 Alternative Verteilungsannahme: Skew-Normalverteilung

In diesem Abschnitt wird eine alternative Verteilungsannahme zur Skew-t-Verteilung betrachtet, um auch jene Datensätze einbeziehen zu können, die die ursprünglichen Annahmen verletzen. Die Skew-Normalverteilung wird als mögliche Alternative her-

angezogen, da sie für Daten geeignet sein könnte, die die Annahme der Normalverteilung nicht vollständig erfüllen, aber dennoch keine extreme Abweichung davon zeigen.

Die Skew-Normalverteilung ist eine Erweiterung der Normalverteilung, die es ermöglicht, asymmetrische Daten zu modellieren. Diese Verteilung wurde erstmals von Azzalini (1985) eingeführt und hat sich als nützlich erwiesen, wenn Daten zwar „normalähnlich“ sind, aber eine gewisse Schiefe aufweisen.

Die multivariate Skew-Normalverteilung wird durch drei Parameter beschrieben:

- ξ : Ein Vektor, der die Mittelwerte der Verteilung angibt.
- Ω : Eine positiv semidefinite Kovarianzmatrix, die die Streuung und die Korrelationen beschreibt.
- α : Ein Vektor, der die Schiefe der Verteilung repräsentiert.

Die multivariate Dichtefunktion ist definiert als:

$$f(x; \xi, \Omega, \alpha) = 2 \cdot \phi_p(x; \xi, \Omega) \cdot \Phi(\alpha^\top \Omega^{-1/2}(x - \xi)),$$

wobei:

- $\phi_p(x; \xi, \Omega)$ die Dichtefunktion der p -dimensionalen multivariaten Normalverteilung mit Erwartungswert $\xi \in \mathbb{R}^p$ und Kovarianzmatrix $\Omega \in \mathbb{R}^{p \times p}$ ist.
- $\Phi(z)$ die Verteilungsfunktion der Standardnormalverteilung bezeichnet, angewendet auf $z \in \mathbb{R}$.
- $\Omega^{-1/2}$ die Matrixwurzel der inversen Kovarianzmatrix darstellt.

Die Skew-Normalverteilung stellt eine einfache Erweiterung der Normalverteilung dar, da sie durch den Parameter α die Möglichkeit bietet, asymmetrische Daten zu modellieren. Für den Spezialfall $\alpha = 0$ entspricht sie der Normalverteilung. Ihre Flexibilität, sowohl symmetrische als auch asymmetrische Verteilungen zu modellieren, sowie ihre mathematische Handhabbarkeit machen sie zu einer attraktiven Alternative für Datensätze, die nicht den Annahmen der Skew-t-Verteilung genügen (vgl. Azzalini & Capitanio, 1999).

Zur Anpassung der Daten an die Skew-Normalverteilung wurde die R-Funktion `msn.mple()` aus dem Paket `sn` verwendet. Diese Funktion liefert die zugrunde liegenden Parameter ξ , Ω und α mittels Maximum-Likelihood-Schätzung. Die folgenden Tabellen zeigen die ermittelten Parameterschätzungen für jene Datensätze (Österreich, Kanada, Norwegen), die die Voraussetzungen der Skew-t-Verteilung nicht erfüllten:

Parameter	Wert	
Österreich:	ξ	$\begin{pmatrix} 0.009189624 \\ 0.001749563 \end{pmatrix}$
	Ω	$\begin{pmatrix} 1.090138 \times 10^{-3} & 3.158322 \times 10^{-5} \\ 3.158322 \times 10^{-5} & 6.188544 \times 10^{-6} \end{pmatrix}$
	α	$\begin{pmatrix} 0.2059668 \\ -2.3992292 \end{pmatrix}$

Tabelle 5: Parameterschätzungen der Skew-Normalverteilung für die Daten aus Österreich.

Parameter	Wert	
Norwegen:	ξ	$\begin{pmatrix} 0.0286773524 \\ 0.0005039792 \end{pmatrix}$
	Ω	$\begin{pmatrix} 1.830284 \times 10^{-3} & 2.786836 \times 10^{-5} \\ 2.786836 \times 10^{-5} & 3.368710 \times 10^{-6} \end{pmatrix}$
	α	$\begin{pmatrix} -1.8168954 \\ -0.3282486 \end{pmatrix}$

Tabelle 6: Parameterschätzungen der Skew-Normalverteilung für die Daten aus Norwegen.

Parameter	Wert	
Kanada:	ξ	$\begin{pmatrix} -0.007288754 \\ 0.001160604 \end{pmatrix}$
	Ω	$\begin{pmatrix} 2.923677 \times 10^{-4} & -4.947766 \times 10^{-6} \\ -4.947766 \times 10^{-6} & 3.536533 \times 10^{-6} \end{pmatrix}$
	α	$\begin{pmatrix} 1.679594 \\ -3.489364 \end{pmatrix}$

Tabelle 7: Parameterschätzungen der Skew-Normalverteilung für die Daten aus Kanada.

Die Anpassung der Daten an die Skew-Normalverteilung zeigt insgesamt plausible Ergebnisse. Besonders auffällig ist, dass der Schiefeparameter α in allen Fällen signifikant von null abweicht. Dies deutet darauf hin, dass die Skew-Normalverteilung die in den Daten enthaltene Schiefe adäquat abbildet.

Im weiteren Verlauf dieser Arbeit wird wieder auf die Skew-t-Verteilung zurückgegriffen, um die ursprünglichen Fragestellungen weiter zu untersuchen. Dennoch zeigt dieser Abschnitt, dass die Anpassung an alternative Verteilungen wie die Skew-Normalverteilung eine sinnvolle Ergänzung darstellen kann, insbesondere wenn die ursprünglichen Annahmen nicht vollständig erfüllt werden.

6.6 Generelle Annahmen zur Modellierung

Im Folgenden treffen wir die Annahme, dass die zu sichernde Verbindlichkeit aus einem Portfolio von Leibrenten besteht, mit einer Laufzeit von $T = 25$ Jahren. Um die Vergleichbarkeit der Ergebnisse zwischen den verschiedenen Ländern nicht zu beeinträchtigen, betrachten wir ausschließlich Rentenverpflichtungen, die an die männliche Bevölkerung der jeweiligen Länder ausgegeben werden.

Die betrachtete Bevölkerung besteht aus Personen, die zum Zeitpunkt des Verkaufs der Leibrenten am Ende des Jahres $t_0 = 2016$ ein Alter von $x = 60$ Jahren erreicht haben. Die Absicherung wird ebenfalls zu diesem Zeitpunkt, also am Ende des Jahres $t_0 = 2016$, eingerichtet.

Jede Leibrente zahlt dem Rentenempfänger am Ende eines jeden Jahres einen konstanten Betrag von 1 USD. Die Rentenzahlungen enden entweder mit dem Ableben des Empfängers oder nach Erreichen der Laufzeit $T = 25$ Jahre, je nachdem, welches Ereignis zuerst eintritt.

Um die Modellierung zu vereinfachen, nehmen wir an, dass das betrachtete Rentenportfolio groß genug ist, sodass das mikroökonomische Langlebkeitsrisiko (d.h. die Unsicherheit durch individuelle Sterblichkeit einzelner Rentenempfänger) vernachlässigbar ist. Damit bleibt das makroökonomische Langlebkeitsrisiko, also die Unsicherheit in Bezug auf allgemeine Sterblichkeitstrends der Bevölkerung, als zentraler Risikofaktor bestehen.

Zur Absicherung des Langlebkeitsrisikos werden in unserem Modell $m = 4$ q-Forwards verwendet. Diese Instrumente wurden alle zum Zeitpunkt $t_0 = 2016$ emittiert und sind an die Sterblichkeit der männlichen Bevölkerung des jeweiligen Datensatzes gekoppelt. Die Details der vier q-Forwards sind in Tabelle 6 zusammengefasst:

j	x_j	T_j	$v(x_j)$
1	60	10	2,25 %
2	65	10	2,92 %
3	70	10	3,66 %
4	75	10	4,28 %

Tabelle 8: Details der vier q-Forwards in unserem Portfolio

Die Tabelle zeigt für jedes der vier q-Forwards:

- das Referenzalter x_j ,
- die Restlaufzeit T_j (in Jahren),

- und die Volatilität $v(x_j)$ der jährlichen Änderungen in den geglätteten Sterbewahrscheinlichkeiten.

Die Werte $v(x_j)$ werden dabei auf Basis geglätteter Sterbewahrscheinlichkeiten berechnet. Dies ist notwendig, da die Preisbildung von q-Forwards kein kleines Stichprobenrisiko berücksichtigen sollte, das durch Diversifikation eliminiert werden kann. Die Forward-Sterblichkeitsrate $q^f(x_j, t_0 + T_j)$ für den j -ten q-Forward kann unter Verwendung der geschätzten Volatilität $v(x_j)$, eines Marktrisikopreises $\ell = 0,25$ und der geglätteten Sterbewahrscheinlichkeit $\hat{q}(x_j, t_0 + T_j)$ mithilfe der folgenden Gleichung berechnet werden:

$$q^f(x_j, t_0 + T_j) = (1 - T_j \cdot \ell \cdot v(x_j)) \cdot \hat{q}(x_j, t_0 + T_j).$$

Dabei gilt, dass ℓ der Marktrisikopreis ist, der die Risikoprämie widerspiegelt, die von Marktteilnehmern für die Übernahme des Langlebigkeitsrisikos verlangt wird. Obwohl der Marktrisikopreis ℓ in der Realität zeitlich variabel sein könnte, da sich die Einschätzungen der Marktteilnehmer hinsichtlich des Langlebigkeitsrisikos im Laufe der Zeit ändern, gehen wir in unseren numerischen Beispielen davon aus, dass ℓ nicht zeitabhängig ist. Diese Annahme ist gerechtfertigt, da der derzeit noch wenig entwickelte Markt für Langlebigkeitsrisiken keine ausreichende Grundlage bietet, um ein realistisch zeitlich variierendes Modell zu erstellen.

Für die Absicherung setzen wir den Budgetparameter C aus unserem Polynomial Goal Programming-Modell auf 0,5 %. Das bedeutet, dass der Absichernde für die zum Zeitpunkt t formulierte Absicherungsstrategie ein Budget in Höhe von $0,005 \cdot \mathbb{E}[V_L(t+1)|\mathcal{F}_t]$ zur Verfügung hat, wobei $\mathbb{E}[V_L(t+1)|\mathcal{F}_t]$ den erwarteten Wert der Verbindlichkeit zum Zeitpunkt $t+1$, gegeben die verfügbare Information \mathcal{F}_t , darstellt. Zusätzlich setzen wir den Diskontierungszinssatz r auf 1 %.

6.7 Berechnung der Momente

In diesem Abschnitt betrachten wir die Approximation der analytischen Formeln für die Momente M_t, V_t, S_t, K_t anhand des Datensatzes für England und Wales. Dafür berechnen wir zuerst die Momente von $\vec{\eta}(t+1) = (\eta_1(t+1), \eta_2(t+1))'$, gegeben \mathcal{F}_t . Die Herleitung basiert auf den in Abschnitt 5.3 abgeleiteten Ergebnissen.

Für das erste Moment gilt:

$$\mathbb{E}[\vec{\eta}(t+1)|\mathcal{F}_t] = \vec{0}.$$

Die Varianz-Kovarianz-Matrix des zweiten Moments ist gegeben durch:

$$\mathbb{E}[\vec{\eta}(t+1)\vec{\eta}(t+1)'|\mathcal{F}_t] = \begin{bmatrix} \mathbb{E}[\eta_1^2(t+1)] & \mathbb{E}[\eta_1(t+1)\eta_2(t+1)] \\ \mathbb{E}[\eta_2(t+1)\eta_1(t+1)] & \mathbb{E}[\eta_2^2(t+1)] \end{bmatrix}.$$

Für die Daten aus England und Wales erhalten wir:

$$\begin{bmatrix} 1,15444 \times 10^{-3} & 3,017550 \times 10^{-5} \\ 3,017550 \times 10^{-5} & 2,683895 \times 10^{-6} \end{bmatrix}.$$

Für das dritte Moment sind die berechneten Werte gegeben durch:

$$\mathbb{E}[\vec{\eta}(t+1)\vec{\eta}(t+1)'\vec{\eta}(t+1)|\mathcal{F}_t] = \begin{bmatrix} 4,491473 \times 10^{-6} & -4,596939 \times 10^{-7} & -4,596939 \times 10^{-7} & -2,174154 \times 10^{-8} \\ -4,596939 \times 10^{-7} & -2,174154 \times 10^{-8} & -2,174154 \times 10^{-8} & -4,599545 \times 10^{-9} \end{bmatrix}.$$

Das vierte Moment umfasst mehrere kombinierte Terme. Die Ergebnisse sind in den folgenden Matrizen dargestellt:

Erste Teilmatrix:

$$\begin{bmatrix} 9,775893 \times 10^{-6} & 2,520186 \times 10^{-7} & 2,478152 \times 10^{-7} & 2,203836 \times 10^{-8} \\ 2,520186 \times 10^{-7} & 1,957070 \times 10^{-9} & 2,203836 \times 10^{-8} & 5,897801 \times 10^{-10} \end{bmatrix}.$$

Zweite Teilmatrix:

$$\begin{bmatrix} 2,478152 \times 10^{-7} & 1,274576 \times 10^{-8} & 3,325230 \times 10^{-9} & 6,353141 \times 10^{-10} \\ 1,274576 \times 10^{-8} & 5,897801 \times 10^{-10} & 6,353141 \times 10^{-10} & 6,660878 \times 10^{-11} \end{bmatrix}.$$

Des Weiteren wollen wir das Cairns-Blake-Dowd-Modell in einen Zustandsraum überführen, wie es bereits in Abschnitt 3.2 beschrieben wurde. Dabei betrachten wir zunächst die Beobachtungsgleichung, die sich durch folgende Darstellung ergibt:

$$\vec{y}(t) = \begin{pmatrix} \ln\left(\frac{q(x_a,t)}{1-q(x_a,t)}\right) \\ \ln\left(\frac{q(x_a+1,t)}{1-q(x_a+1,t)}\right) \\ \vdots \\ \ln\left(\frac{q(x_b,t)}{1-q(x_b,t)}\right) \end{pmatrix}, \quad \vec{d} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & (x_a - \bar{x}) \\ 1 & (x_a + 1 - \bar{x}) \\ \vdots & \vdots \\ 1 & (x_b - \bar{x}) \end{pmatrix}.$$

Hierbei ist $\vec{y}(t)$ der Vektor der beobachteten logit-transformierten Sterbewahrscheinlichkeiten für die Altersgruppe $[x_a, x_b]$, während der Vektor \vec{d} eine konstante Null-Vorgabe enthält. Die Matrix B repräsentiert die lineare Struktur des Modells, wobei \bar{x} den Durchschnitt der betrachteten Altersgruppe darstellt.

Die zugehörigen Übergangsgleichungen sind bereits in Abschnitt 6.4 detailliert be-

schrieben und werden hier nicht erneut ausgeführt. Dort wurde gezeigt, wie sich die Zustandsvariablen dynamisch über die Zeit entwickeln.

Unser Ziel ist es, die Werte für die Momente M_t, V_t, S_t und K_t zu berechnen. Dazu betrachten wir zunächst die partiellen Ableitungen

$$D_k(0, t_0 + 1) = \frac{\partial L}{\partial \eta_k(t_0 + 1)} \quad \text{und} \quad D_k(j, t_0 + 1) = \frac{\partial H(j, t_0)}{\partial \eta_k(t_0 + 1)}, \quad j = 1, \dots, m.$$

Die benötigten Formeln sind bereits in Abschnitt 5.2 hergeleitet worden. Zur Prognose der zukünftigen Mortalitätswahrscheinlichkeiten im Cairns-Blake-Dowd-Modell verwenden wir einen bivariaten Random Walk mit Drift \vec{c} , um die Entwicklung der Parameter $\kappa_t^{(1)}$ und $\kappa_t^{(2)}$ zu schätzen. Die resultierenden Werte für $D_k(0, t_0 + 1)$ und $D_k(j, t_0 + 1)$ für die Daten aus England und Wales sind in der folgenden Tabelle zusammengefasst:

k	$D_k(0, t_0 + 1)$	$D_k(1, t_0 + 1)$	$D_k(2, t_0 + 1)$	$D_k(3, t_0 + 1)$	$D_k(4, t_0 + 1)$
1	-3,34416	-0,00497167	-0,00835200	-0,0139453	-0,0230487
2	12,62698	0,07208915	0,07934401	0,0627540	-0,0115244

Tabelle 9: Werte von $D_k(0, t_0 + 1)$ und $D_k(j, t_0 + 1)$

Im nächsten Schritt approximieren wir $V_L(t_0 + 1)$ und $V_H(t_0 + 1, j, t_0)$ für $j = 1, 2, 3, 4$. Die hierfür verwendeten Formeln sind in Abschnitt 5.4 beschrieben.

1. Berechnung von \hat{L} :

Mithilfe der Ergebnisse aus Abschnitt 3.4 und basierend auf den prognostizierten Überlebenswahrscheinlichkeiten aus dem Cairns-Blake-Dowd-Modell ergibt sich:

$$\hat{L} = 21,4327.$$

2. Berechnung von $V_L(t_0 + 1)$:

Unter Verwendung der eben berechneten Werte $D_k(0, t_0 + 1)$, $k = 1, 2$ und der Prognosen für die Parameter $\kappa_t^{(1)}$ und $\kappa_t^{(2)}$, die mithilfe des Driftparameters \vec{c} geschätzt wurden, erhalten wir:

$$V_L(t_0 + 1) = 21.1405.$$

3. Berechnung von $V_H(t_0 + 1, j, t_0)$:

Analog zum Vorgehen bei $V_L(t_0 + 1)$ berechnen wir die Werte für $V_H(t_0 + 1, j, t_0)$,

wobei wir zunächst

$$\hat{H} = \begin{pmatrix} -0,000327015 \\ -0,000716130 \\ -0,001510040 \\ -0,002955557 \end{pmatrix}$$

berechnen und schließlich

$$V_H(t_0 + 1, j, t_0) = \begin{pmatrix} -0,000862273 \\ -0,001536334 \\ -0,002747654 \\ -0,004783101 \end{pmatrix}$$

erhalten.

6.8 Optimierungsproblem

Nun wenden wir uns der Berechnung der Werte von M_t, V_t, S_t und K_t zu, um das Teiloptimierungsproblem aus Abschnitt 4.1 weiter zu bearbeiten. Mit diesen Berechnungen werden wir schließlich die Werte von $\mathcal{M}^{(i)}$, für $i = 1, 2, 3, 4$, ermitteln.

Zur Berechnung von M_t, V_t, S_t und K_t greifen wir auf die in Kapitel 5.3 hergeleiteten Formeln zurück. Die Berechnung erfordert insbesondere die Terme $D_k(j, t_0 + 1)$ sowie die Momente von $\vec{\eta}(t + 1)$. Für den Datensatz England und Wales werden die Ergebnisse für M_t und V_t in dieser Arbeit präsentiert. Die Resultate für S_t und K_t werden aufgrund ihrer Größe nicht explizit angegeben, da es sich hierbei um einen Tensor dritten bzw. vierten Grades mit Dimensionen $5 \times 5 \times 5$ bzw. $5 \times 5 \times 5 \times 5$ handelt.

Das berechnete Ergebnis für M_t lautet:

$$M_t = \begin{bmatrix} 21,4327 & -0,000327 & -0,000716 & -0,001510 & -0,002956 \end{bmatrix}.$$

Das berechnete Ergebnis für V_t lautet:

$V_t =$

$$\begin{pmatrix} 1,0790 \times 10^{-2} & 1,2467 \times 10^{-5} & 2,3743 \times 10^{-5} & 4,4318 \times 10^{-5} & 8,0972 \times 10^{-5} \\ 1,2467 \times 10^{-5} & 2,0852 \times 10^{-8} & 3,3215 \times 10^{-8} & 5,2430 \times 10^{-8} & 8,1648 \times 10^{-8} \\ 2,3743 \times 10^{-5} & 3,3215 \times 10^{-8} & 5,7431 \times 10^{-8} & 9,8618 \times 10^{-8} & 1,6749 \times 10^{-7} \\ 4,4318 \times 10^{-5} & 5,2430 \times 10^{-8} & 9,8618 \times 10^{-8} & 1,8226 \times 10^{-7} & 3,3032 \times 10^{-7} \\ 8,0972 \times 10^{-5} & 8,1648 \times 10^{-8} & 1,6749 \times 10^{-7} & 3,3032 \times 10^{-7} & 6,2967 \times 10^{-7} \end{pmatrix}.$$

Im nächsten Schritt berechnen wir die Werte von $M^{(i)}$ für $i = 1, 2, 3, 4$, indem wir das Teiloptimierungsproblem aus Abschnitt 4.2 lösen. Für die numerische Berechnung der Lösung wurde die R-Funktion `nloptr` aus dem gleichnamigen Package verwendet. Die resultierenden Werte sind wie folgt:

$$M^{(i)} = \begin{cases} 21.4327, & i = 1 \\ 0.0106, & i = 2 \\ -0.4327, & i = 3 \\ 7.3134, & i = 4 \end{cases} .$$

Damit sind alle Vorbereitungen für die Lösung des Optimierungsproblems abgeschlossen.

6.9 Ergebnisse

Schließlich berechnen wir die Momente von $V_P(t_0 + 1, \vec{N}_{t_0})$. Die Ergebnisse des Optimierungsproblems sind in der folgenden Tabelle dargestellt. Es wurden sieben verschiedene Risikopräferenzen berücksichtigt, die sich durch unterschiedliche Gewichtungen in der Optimierung ergeben. Die Tabelle zeigt die vier Einträge des Vektors der Nominalbeträge \vec{N}_0 zum Startzeitpunkt sowie die daraus resultierenden Werte für die berechneten Momente.

	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)
λ_1	1	1	1	1	1	1	1
λ_2	1	1	1	1	1	1	1
λ_3	0	0,5	1	2	0	0	0
λ_4	0	0	0	0	0,5	1	2
Mittelwert	21,435037	21,435045	21,435066	21,435142	21,435031	21,435031	21,435038
Var ($\times 10^{-2}$)	1,065508	1,065508	1,065508	1,065508	1,065508	1,065508	1,065508
Schiefe	-0,300628	-0,300703	-0,300898	-0,301566	-0,300578	-0,300583	-0,300636
Kurtosis	7,465808	7,465885	7,466086	7,466774	7,465756	7,465751	7,465716

Tabelle 10: Ergebnisse des Optimierungsproblems für verschiedene Risikopräferenzen.

Zunächst vergleichen wir die Risikopräferenz (i) mit den Präferenzen (ii) bis (iv). Risikopräferenz (i) repräsentiert den klassischen Mean-Variance-Ansatz, bei dem weder die Schiefe noch die Kurtosis berücksichtigt werden ($\lambda_3 = \lambda_4 = 0$). Im Gegensatz dazu beziehen die Präferenzen (ii) bis (iv) die Schiefe durch eine Gewichtung von λ_3 in die Optimierung mit ein.

Wie erwartet, nimmt der Wert der Schiefe ab, je mehr Gewicht auf λ_3 gelegt wird. Dies zeigt, dass eine stärkere Berücksichtigung der Schiefe zu einer gezielten Re-

duktion dieses Moments führt. Allerdings verschlechtern sich bei einer stärkeren Gewichtung der Schiefe zwangsläufig die anderen Momente, insbesondere die Kurtosis.

Ein ähnliches Verhalten lässt sich auch bei der Gewichtung der Kurtosis beobachten, wie in den Präferenzen (v) bis (vii) sichtbar wird. Eine zunehmende Gewichtung von λ_4 reduziert den Wert der Kurtosis, beeinflusst jedoch auch die Schiefe leicht. Dies verdeutlicht, dass die Momente nicht unabhängig voneinander sind und eine Veränderung der Gewichtung eines Moments auch Auswirkungen auf die anderen Momente hat.

In der Arbeit von Siu et al. (2023), in der ein ähnliches Modell entwickelt wurde, werden zusätzliche Analysen zu Gewinnen und Verlusten durchgeführt. Die Ergebnisse stimmen mit den in dieser Arbeit dargestellten Resultaten überein: Ein Ansatz, der die ersten vier Momente berücksichtigt, liefert bessere Ergebnisse, solange kein Gewicht übermäßig stark belastet wird.

Darüber hinaus vergleichen Siu et al. das Modell mit dem Value at Risk (VaR) und dem Tail-Value at Risk (TVaR) bei einem Konfidenzniveau von 99,5 %. Sie zeigen, dass das Modell weiterhin valide bleibt, solange das Konfidenzniveau hoch bleibt.

Für spezifische Hedging-Strategien, die auf ein bestimmtes Ziel wie die Minimierung des Value at Risk abzielen, erzielen spezialisierte Ansätze zwar bessere Ergebnisse in Bezug auf dieses Ziel. Dennoch zeigt sich, dass das in dieser Arbeit vorgestellte Modell ein überlegenes Gesamtergebnis liefert, indem es alle Momente ausgewogen berücksichtigt und somit eine robuste Lösung bietet.

7 Conclusio

In dieser Arbeit wurde ein Modell entwickelt, das sich mit der Modellierung von Langlebigkeit und der Absicherung des damit verbundenen Risikos befasst. Dabei wurde ein innovativer Ansatz verfolgt, bei dem anstelle der üblichen Annahme von Normalverteilungen eine asymmetrische Verteilung verwendet wurde. Um die praktische Anwendbarkeit zu erhöhen, wurde zudem eine Zustandsraumdarstellung implementiert.

Die Ergebnisse sind gemischt: Während das Modell in bestimmten Fällen vielversprechende Resultate liefert, zeigt sich, dass es stark datenabhängig ist. Dies bedeutet, dass das Modell keine allgemeine Eignung besitzt, sondern die zugrunde liegenden Daten zunächst auf ihre Verteilungsstruktur hin überprüft werden müssen.

Ein Aspekt dieser Arbeit war die Untersuchung alternativer Verteilungsannahmen, insbesondere der Skew-Normalverteilung, als Ergänzung zur Skew-t-Verteilung. Die Ergebnisse zeigen, dass die Skew-Normalverteilung eine sinnvolle Alternative darstellt, wenn die Daten die Voraussetzungen der Skew-t-Verteilung nicht erfüllen, insbesondere bei einer geringeren Abweichung von der Symmetrie. Durch die Anpassung der Skew-Normalverteilung konnte die Schiefe der Daten zuverlässig abgebildet werden, was auf die Flexibilität dieser Verteilung hinweist. Dennoch bleibt die Skew-t-Verteilung aufgrund ihrer robusteren Eigenschaften gegenüber Ausreißern und ihrer breiteren Anwendbarkeit in der Modellierung von Langlebigkeitsrisiken die bevorzugte Wahl. Die Analyse verdeutlicht jedoch, dass die Wahl der Verteilungsannahme entscheidend von den Eigenschaften der zugrunde liegenden Daten abhängt und in jedem spezifischen Anwendungsfall sorgfältig geprüft werden sollte.

Bei der Umsetzung des Optimierungsproblems wurde deutlich, dass für Datensätze, die die Modellannahmen erfüllen, erwartungsgemäße Ergebnisse erzielt werden können. Jedoch ist eine sorgfältige Gewichtung der Momente essenziell. Eine übermäßige Gewichtung eines einzelnen Moments kann zu unbefriedigenden Resultaten führen. Dies legt nahe, dass weitere Einschränkungen und Anpassungen notwendig sind, um die Robustheit des Modells zu verbessern.

Ein weiterer wichtiger Punkt betrifft die Hedgestrategie: Das in dieser Arbeit betrachtete Modell basiert auf einem statischen Hedge. Eine mögliche Erweiterung wäre die Entwicklung und Integration dynamischer Hedges, um flexiblere und anpassungsfähigere Ansätze zu ermöglichen.

Zusammenfassend zeigt sich, dass das vorgeschlagene Modell unter den richtigen Annahmen und einer ausgewogenen Gewichtung der Momente Verbesserungen in der Absicherung von Langlebigkeitsrisiken bieten kann. Dennoch bestehen klare Potenziale für zukünftige Forschung und Weiterentwicklung, um die Vielseitigkeit und Anwendbarkeit des Modells zu steigern.

Literatur

- Azzalini, A. and Capitanio, A. (1999), *Statistical applications of the multivariate skew normal distribution*. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 61: 579-602.
<https://doi.org/10.1111/1467-9868.00194>
- Azzalini, A. and Capitanio, A. (2003), *Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution*. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 65: 367-389.
<https://doi.org/10.1111/1467-9868.00391>
- Bauer, Daniel, Börger, Matthias and Ruß, Jochen, (2010), *On the pricing of longevity-linked securities*, Insurance: Mathematics and Economics, 46, issue 1, p. 139-149,
<https://EconPapers.repec.org/RePEc:eee:insuma:v:46:y:2010:i:1:p:139-149>.
- Bayraktar, Erhan & Milevsky, Moshe & Promislow, S. & Young, Virginia. (2009). *Valuation of Mortality Risk via the Instantaneous Sharpe Ratio: Applications to Life Annuities*. Journal of Economic Dynamics and Control. 33. 676-691.
10.1016/j.jedc.2008.09.004.
- Biffis, Enrico & Blake, David P., *Mortality-Linked Securities and Derivatives* (2009). Preprint. <https://ssrn.com/abstract=1340409>
- Booth, Heather & Tickle, Leonie. (2008). *Mortality Modelling and Forecasting: a Review of Methods*. Annals of Actuarial Science. 3. 3-43.
10.1017/S1748499500000440.
- Coughlan, Guy & Epstein, David & Sinha, Amit & Honig, Paul. (2007). *q-Forwards: Derivatives for Transferring Longevity and Mortality Risks*. Pension Advisory Group, JPMorgan Chase Bank NA,
<https://www.researchgate.net/publication/256109844>
- Deckro, Richard & Hebert, John. (1988). *Polynomial goal programming: A procedure for modeling preference trade-offs*. Journal of Operations Management. 7. 149-164. 10.1016/0272-6963(81)90010-3.
- Fung, M.C., & Peters, G.W., & Shevchenko, P.V., *A unified approach to mortality modelling using state-space framework: characterisation, identification, estimation and forecasting*. Annals of Actuarial Science. 2017;11(2):343-389.
doi:10.1017/S1748499517000069
- Hastie, Trevor & Tibshirani, Rob & Friedman, Jerome. (2000). *The Elements of Statistical Learning*. Springer-Verlag, New York. 10.1007/978-0-387-21606-5.
- Henze, N., & Zirkler, B. (1990). *A class of invariant consistent tests for*

multivariate normality. Communications in Statistics - Theory and Methods, 19(10), 3595–3617. <https://doi.org/10.1080/03610929008830400>

- Hunt, Andrew & Blake, David. (2020). *On the Structure and Classification of Mortality Models*. North American Actuarial Journal. 25. 1-20. 10.1080/10920277.2019.1649156.
- Jarque, C. & Bera, Anil. (1980). *Efficient Test for Normality, Heteroscedasticity and Serial Independence of Regression Residuals*. Economics Letters. 6. 3. 255-259. [https://doi.org/10.1016/0165-1765\(80\)90024-5](https://doi.org/10.1016/0165-1765(80)90024-5).
- Kainhofer, R. F., & Predota, M., & Schmock, U. (2006). *The New Austrian Annuity Valuation Table AVÖ 2005R*. Mitteilungen Der Aktuarvereinigung Österreichs, 13, 55–135. <http://hdl.handle.net/20.500.12708/172051>
- Kirkwood, Thomas. (2015). *Deciphering death: A commentary on Gompertz (1825) 'On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies'*. Philosophical transactions of the Royal Society of London. Series B, Biological sciences. 370. 10.1098/rstb.2014.0379.
- Korkmaz, Selcuk & Goksuluk, Dincer & Zararsiz, Gokmen. (2014). *MVN: An R Package for Assessing Multivariate Normality*. The R Journal. 6. 151-162. 10.32614/RJ-2014-031.
- Li, S. H., & Chan, W. S. (2007). *The Lee-Carter Model for Forecasting Mortality, Revisited*. North American Actuarial Journal, 11(1), 68–89. <https://doi.org/10.1080/10920277.2007.10597438>
- Lin, Yijia & Cox, Samuel. (2005). *Securitization of Mortality Risks in Life Annuities*. Journal of Risk & Insurance. 72. 227-252. 10.1111/j.1539-6975.2005.00122.x.
- Liu, Yanxin & Li, Johnny Siu-Hang. (2018). *A strategy for hedging risks associated with period and cohort effects using q-forwards*. Insurance Mathematics & Economics, 78:267-285. doi: 10.1016/J.INSMATHECO.2017.09.007
- Peiro, Amado. (1999). *Skewness in financial returns*. Journal of Banking & Finance, Elsevier, vol. 23(6), pages 847-862, June. <https://EconPapers.repec.org/RePEc:eee:jbfina:v:23:y:1999:i:6:p:847-862>
- Renshaw, A.E. & Haberman, Steven. (2006). *A Cohort-Based Extension to the Lee-Carter Model for Mortality Reduction Factors*. Insurance: Mathematics and Economics. 38. 556-570. 10.1016/j.insmatheco.2005.12.001.
- Safitri, Y. & Mardiyati, Sri & Malik, Maulana. (2019). *The Cairns-Blake-Dowd model to forecast Indonesian mortality rates*. AIP Conference Proceedings.

2168. 020039. 10.1063/1.5132466.

Siu, Johnny & Liu, Yanxin & Chan, Wai. (2023). *Hedging Longevity Risk under Non-Gaussian State-Space Stochastic Mortality Models: A Mean-Variance-Skewness-Kurtosis Approach*. Insurance: Mathematics and Economics. 113. 96-121. 10.1016/j.insmatheco.2023.08.001.