Small Data -
Darf‘s auch etwas weniger sein?

Aktuarielle Survivalanalyse für Versicherungs- und Pensionsbestände

Bernd Heistermann

Kai Kaufhold

„DAV vor Ort“, Wiesbaden, 8.3.2017

AdRes

HEISTERMANN CONSULTING

Agenda

I. Einführung

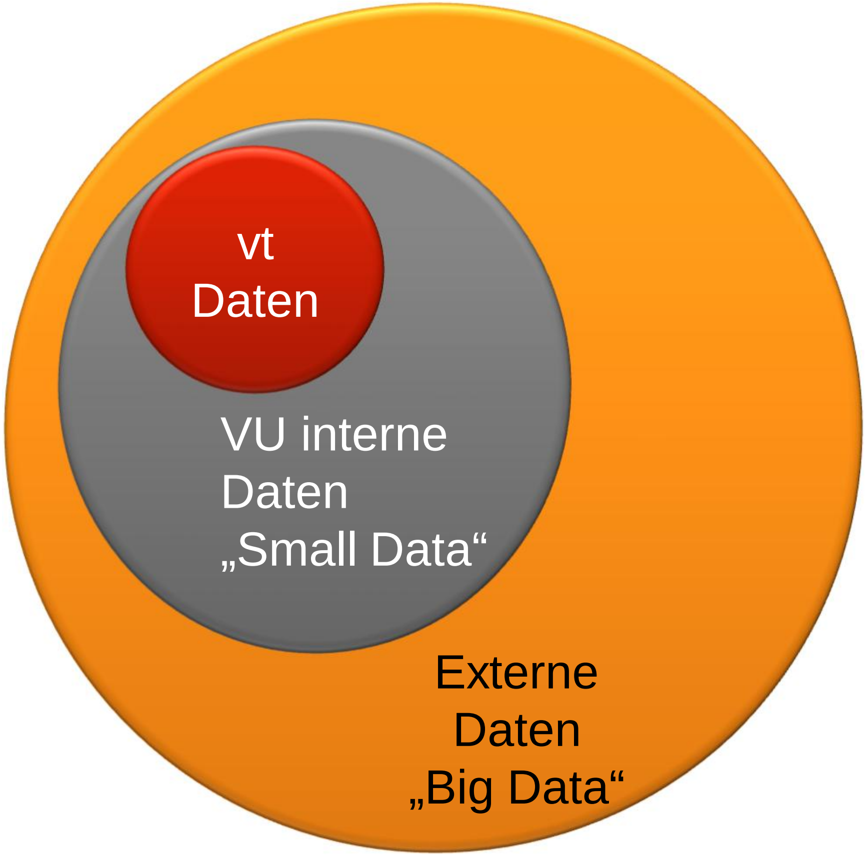
II. Probleme bei Bestandsanalysen

III. Survivalanalyse – Was ist das?

IV. Anwendungen

V. Fazit

I. Einführung - Begriffe



„Small Data“ = data in a volume and format that makes it <u>accessible</u>, <u>informative</u> and <u>actionable</u>.

„Big Data” = data characterized by 3Vs: big <u>volume</u> of data, <u>variety</u> of types of data, high process <u>velocity</u>

Quelle: Whatis.com

I. Einführung – Ziel des Vortrages

Transport einer Idee

- Parametrische Survivalanalyse als moderne Alternative zur klassischen Vorgehensweise bei der Analyse von Ausscheidewahrscheinlichkeiten

Ermutigung

- Auch kleine und mittlere Bestände können wertvolle Erkenntnisse liefern, aber man muss sie auch heben wollen.

II. Probleme bei Bestandsanalysen

Datenqualität

- Woher kommen die Daten?
- Welche (vermutlich) risikobeschreibenden Merkmale kann ich verwenden und wie ist die Erfassungsqualität?
- Kann ich nur Policen oder auch Personen auswerten?
- Sind Besonderheiten im Bestand zu erwarten?
- Wer kennt sich damit aus?

II. Probleme bei Bestandsanalysen

Datenquantität

- Sind die Bestände groß genug, um statistisch relevante Aussagen zu machen?
- Varianzvergrößerung durch Heterogenität

II. Probleme bei Bestandsanalysen

Datenpooling

- Homogenität?
- Datenqualität der Teilnehmer?
- ggT der Daten ist immer kleiner als meine eigene Datenbasis
- Vorsicht bei der Übertragung der Ergebnisse wg Basisrisiko
- +: Benchmarking
  +: geringe Kosten
- Der Poolorganisator gewinnt immer die wertvollsten Erkenntnisse

II. Probleme bei Bestandsanalysen

Organisation

- Wer sollte es machen?
- Kapazität? Priorisierung?
- Zugang zu Daten (Historie)
- Priorisierung bei IT

↕

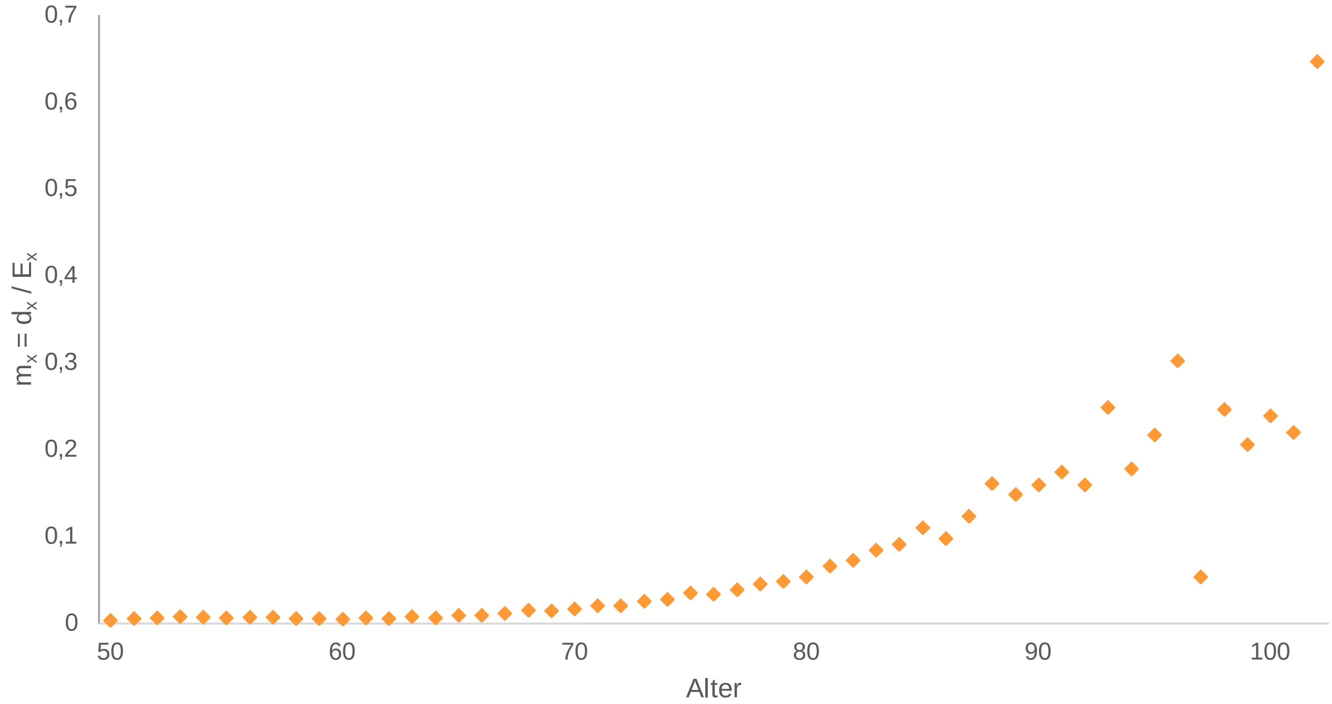
- Tools

III. Einführung Survivalanalyse

1. Graduation by mathematical formula
2. Generalised Linear Models
3. Survivalanalyse

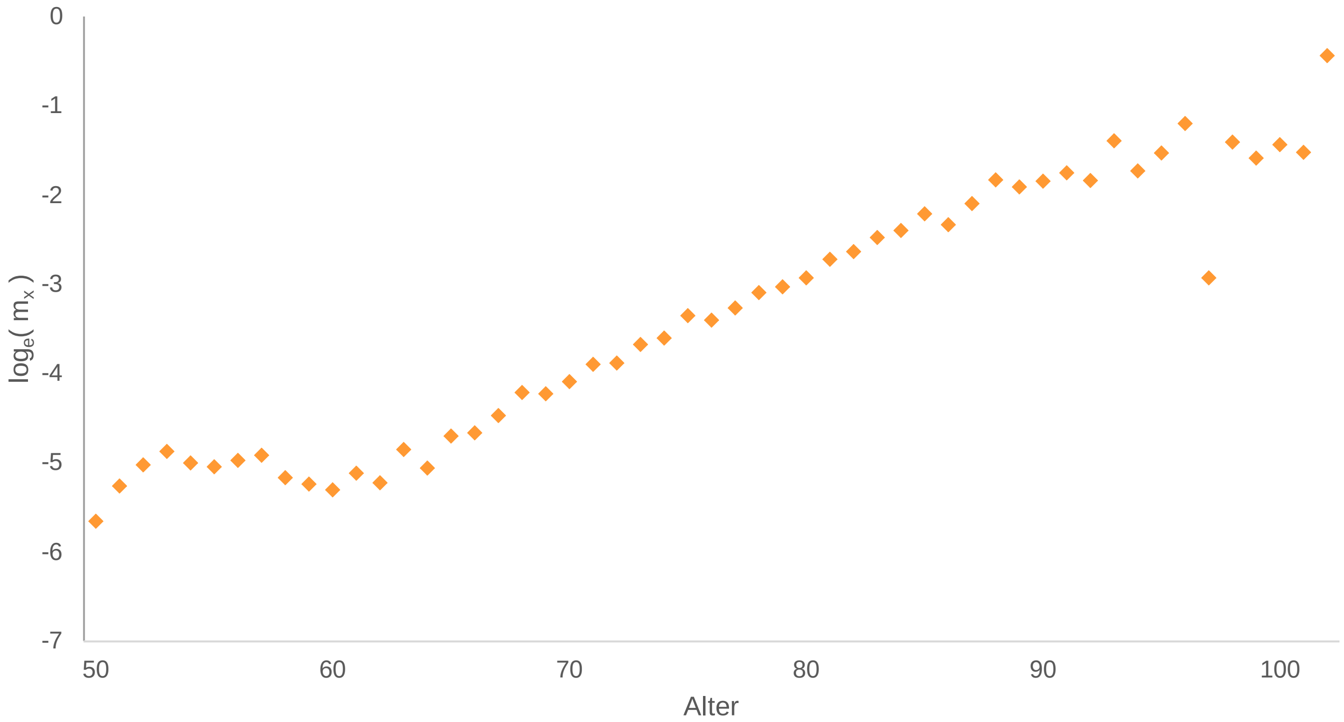
III.1 Graduation by formula

Rohe Sterbeziffern $m_x = {}^{d_x}/_{E_x} = {}^{Tote}/_{verlebte\ Zeit}$



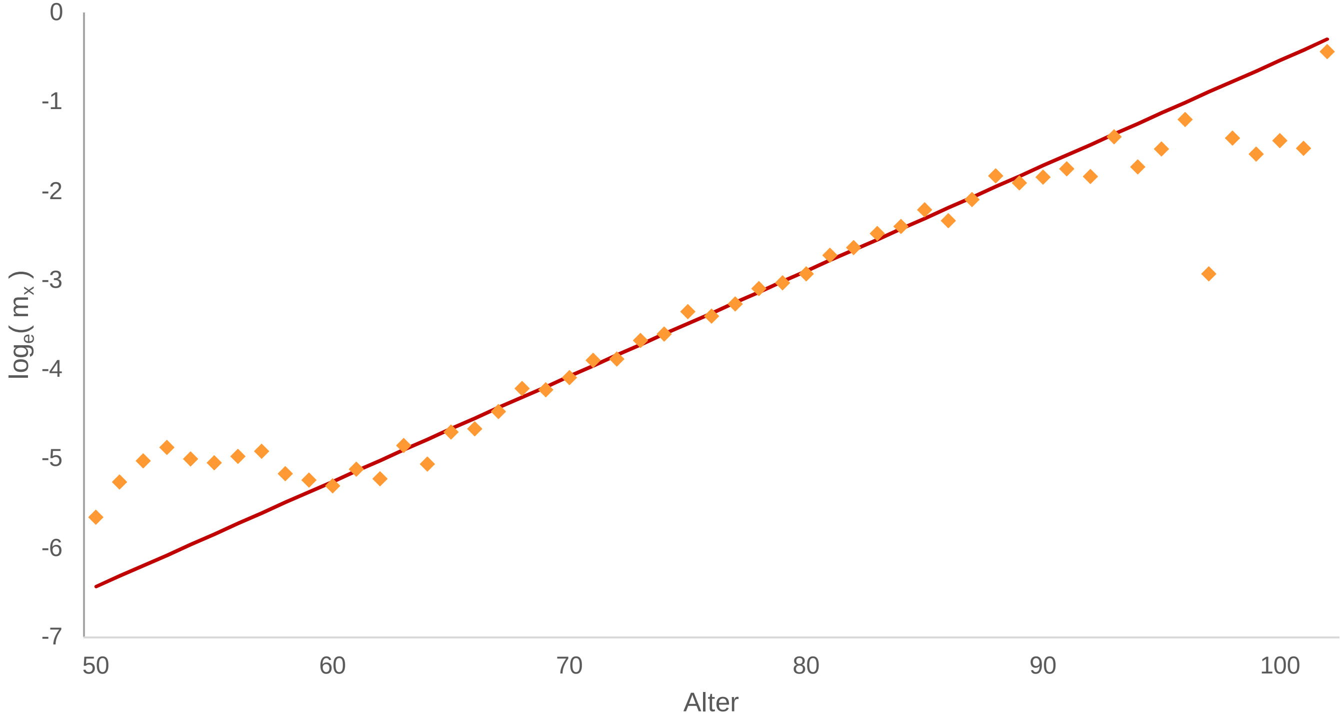
III.1 Graduation by formula

Logarithmische Skala



III.1 Graduation by formula

Gompertz (1825): $m_x = Bc^x = e^{\alpha+\beta x}$



III. Graduation by mathematical formula

Vorteile:

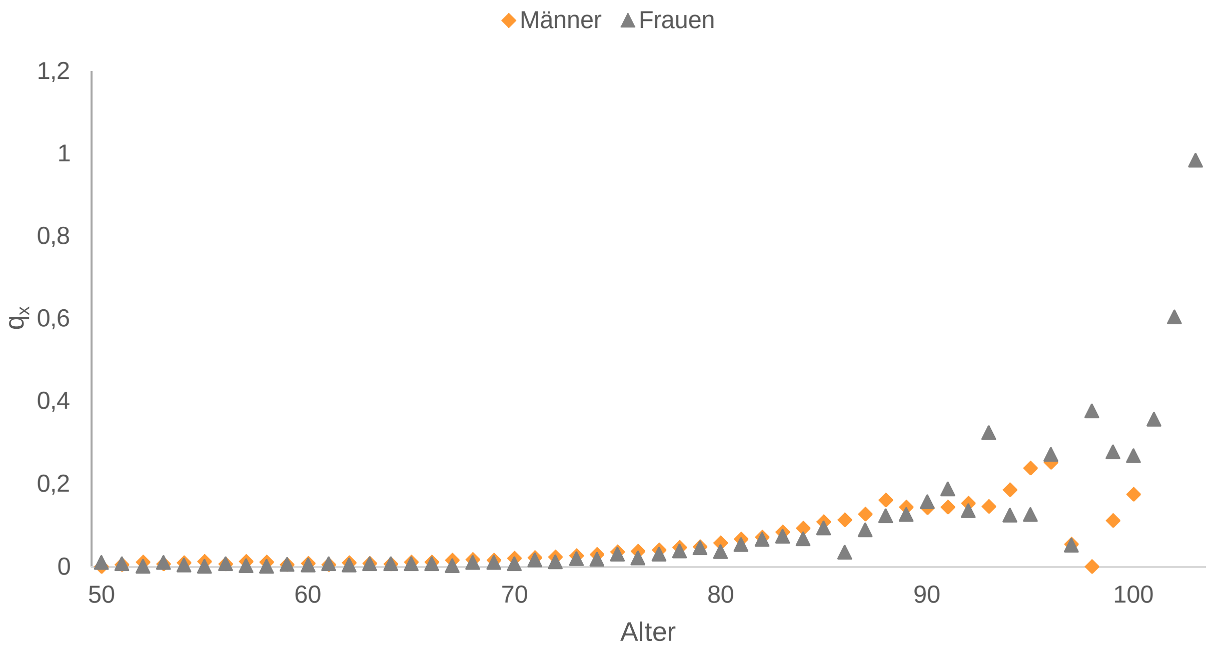
- Leichte Implementierung
- Breite Auswahl an möglichen Altersstrukturen
- Statistiken über Güte des Fit

Nachteile

- Aggregierte Daten (Altersgruppen)
- Extrapolation manchmal fraglich
- Risikogruppen nur durch Unterteilung der Daten möglich
  à Signifikanz geht verloren

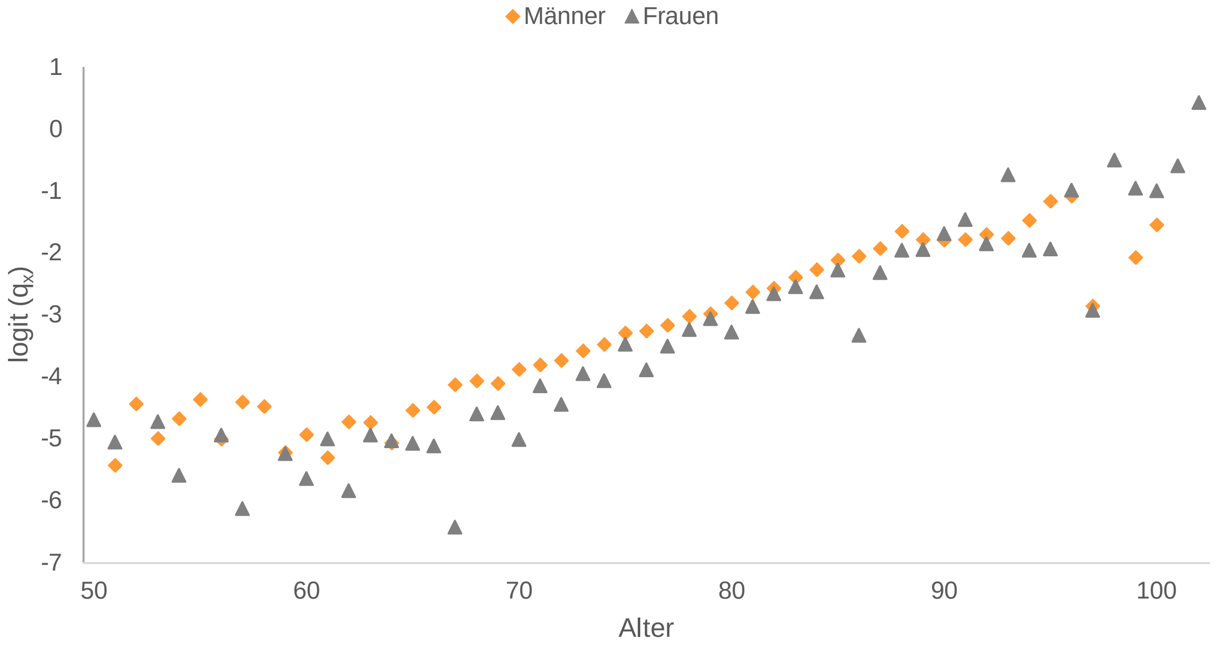
III.2 Generalised Linear Models

Sterbewahrscheinlichkeiten $q_{x-\frac{1}{2}} = 1 - e^{-m_x}$



III.2 Generalised Linear Models

Logistische Sterbew'keit: $\operatorname{logit}(q_x) = \ln\left(\frac{q_x}{1-q_x}\right)$



III.2 Generalised Linear Models

Kanonische Link-Funktion: logistische Funktion

$$q_x = \frac{e^{\alpha+\beta x}}{1 + e^{\alpha+\beta x}}$$

$$\text{logit}(q_x) = \ln\left(\frac{q_x}{1 - q_x}\right) = \alpha + \beta x$$

Zusätzliche Annahme: binomial-verteilte Todesfälle

III.2 Generalised Linear Models

Multivariate Analyse:

Simultane Parameterschätzungen für mehrere Risikofaktoren

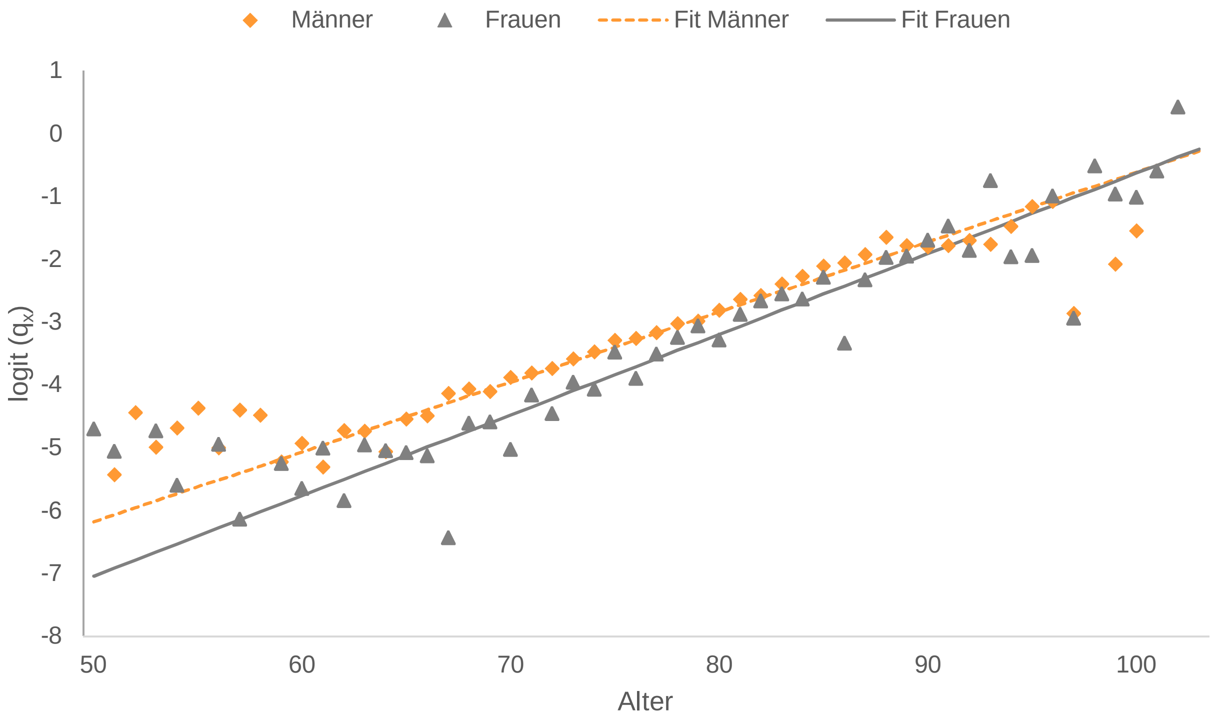
$$\alpha = \alpha_{Basis} + Z_{Frauen}\alpha_{Frauen} + \ldots$$

$$\beta = \beta_{Basis} + Z_{Frauen}\beta_{Frauen} + \ldots$$

Indikatorfunktion $Z_{Frauen} = \begin{cases} 1 & Frauen \\ 0 & sonst \end{cases}$

III.2 Generalised Linear Models

GLM-Fit: $\alpha_{Frauen} = -1{,}71; \beta_{Frauen} = 0{,}017$



III.2 Generalised Linear Models

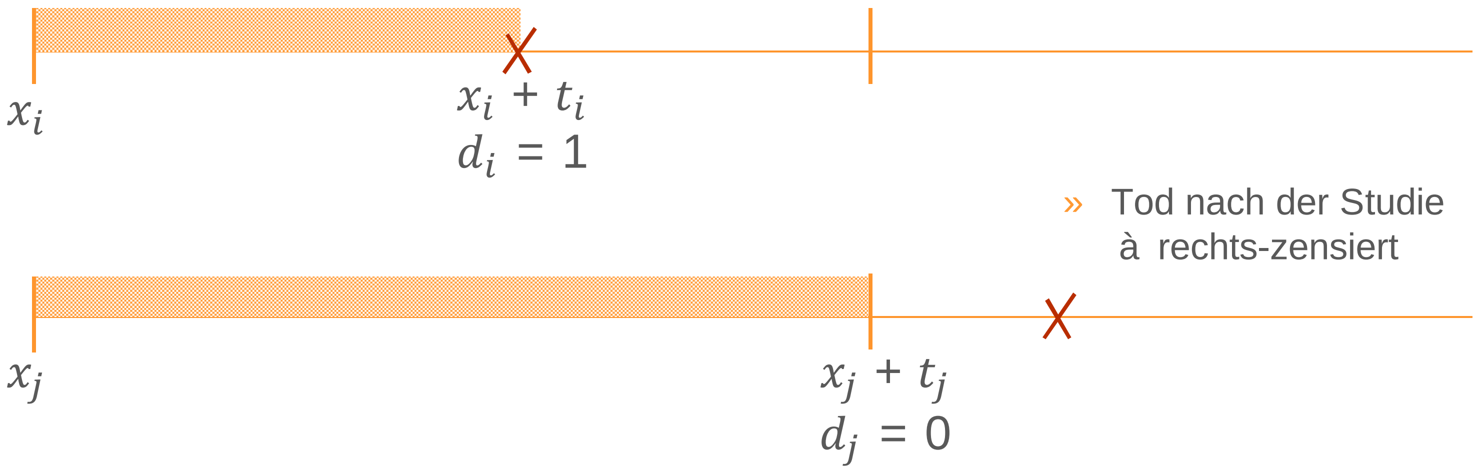
- Vorteile:
  - Multivariate Analyse nach Risikogruppen
  - Einfache Implementierung
  - Statistische Aussagekraft kann gemessen werden.
- Einschränkungen:
  - Aggregierte Daten
  - Eingeschränkte Wahl der Altersstruktur der Sterbegesetze

III.3 Survival-Modelle

- Beobachte jedes Leben i im Bestand N.
- Eintrittalter x_i, beobachtete Zeit t_i, Status d_i

» Tod innerhalb der Studie

» Tod nach der Studie
à rechts-zensiert



III.3 Survival-Modelle

Einfaches Beispiel: Makeham (1859)

Sterbeintensität:

$$\mu_x = A + Bc^x = e^{\epsilon} + e^{\alpha+\beta x}$$

Überlebenswahrscheinlichkeit:

$$S_x = {}_tp_x = \exp\left(-\int_0^t \mu_{x+s} ds\right) = e^{-H_x(t)}$$

III.3 Survival-Modelle

Verteilung für Restlebensdauer T des Lebens i:

$$f_T(t_i) = P(T = t_i) = {}_{t_i}p_{x_i}\mu_{x_i+t_i}^{d_i}$$

Likelihoodfunktion L:

$$L = \prod_i {}_{t_i}p_{x_i}\mu_{x_i+t_i}^{d_i}$$

Log-Likelihood ll:

$$ll = \ln(L) = -\sum_i H_{x_i}(t_i) + \sum_i d_i \ln(\mu_{x_i+t_i})$$

III.3 Survival-Modelle

Verteilung für Restlebensdauer T des Lebens i:

$$f_T(t_i) = P(T = t_i) = {}_{t_i}p_{x_i}\mu_{x_i+t_i}^{d_i}$$

Likelihoodfunktion L:

$$L = \prod_i {}_{t_i}p_{x_i}\mu_{x_i+t_i}^{d_i}$$

Log-Likelihood ll:

$$ll = \ln(L) = -\sum_i H_{x_i}(t_i) + \sum_i d_i \ln\left(\mu_{x_i+t_i}\right)$$

III.3 Survival-Modelle

Verteilung für Restlebensdauer T des Lebens i:

$$f_T(t_i) = P(T = t_i) = {}_{t_i}p_{x_i}\mu_{x_i+t_i}^{d_i}$$

Likelihoodfunktion L:

$$L = \prod_i {}_{t_i}p_{x_i}\mu_{x_i+t_i}^{d_i}$$

Log-Likelihood ll:

$$ll = \ln(L) = -\sum_i H_{x_i}(t_i) + \sum_i d_i \ln(\mu_{x_i+t_i})$$

III.3 Survival-Modelle

Verteilung für Restlebensdauer T des Lebens i:

$$f_T(t_i) = P(T = t_i) = {}_{t_i}p_{x_i}\mu_{x_i+t_i}^{d_i}$$

Likelihoodfunktion L:

$$L = \prod_i {}_{t_i}p_{x_i}\mu_{x_i+t_i}^{d_i}$$

Log-Likelihood ll:

$$ll = \ln(L) = -\sum_i H_{x_i}(t_i) + \sum_i d_i \ln(\mu_{x_i+t_i})$$

III.3 Survival-Modelle

Noch ein Beispiel: Makeham-Beard (1959):

$$\mu_{x,y} = \frac{e^{\epsilon} + e^{\alpha+\beta x+\delta(y-2000)}}{1 + e^{\alpha+\varrho+\beta x+\delta(y-2000)}}$$

- Stetige Funktion in Alter x und Zeit y.
- Parameter α, β, ϵ und ϱ beschreiben Sterblichkeit in verschiedenen Altersbereichen,
- δ einen zeitlichen Trend.

III.3 Survival-Modelle

Zwei Parameter verletzen die Linearität:

$$\mu_{x,y} = \frac{e^{\epsilon} + e^{\alpha+\beta x+\delta(y-2000)}}{1 + e^{\alpha+\varrho+\beta x+\delta(y-2000)}}$$

- Parameter ϵ beschreibt altersunabhängige Übersterblichkeit,
- ϱ das asymptotische Verhalten bei hohen Altern.

$$\mu_x \xrightarrow[x\to\infty]{} e^{-\varrho}$$

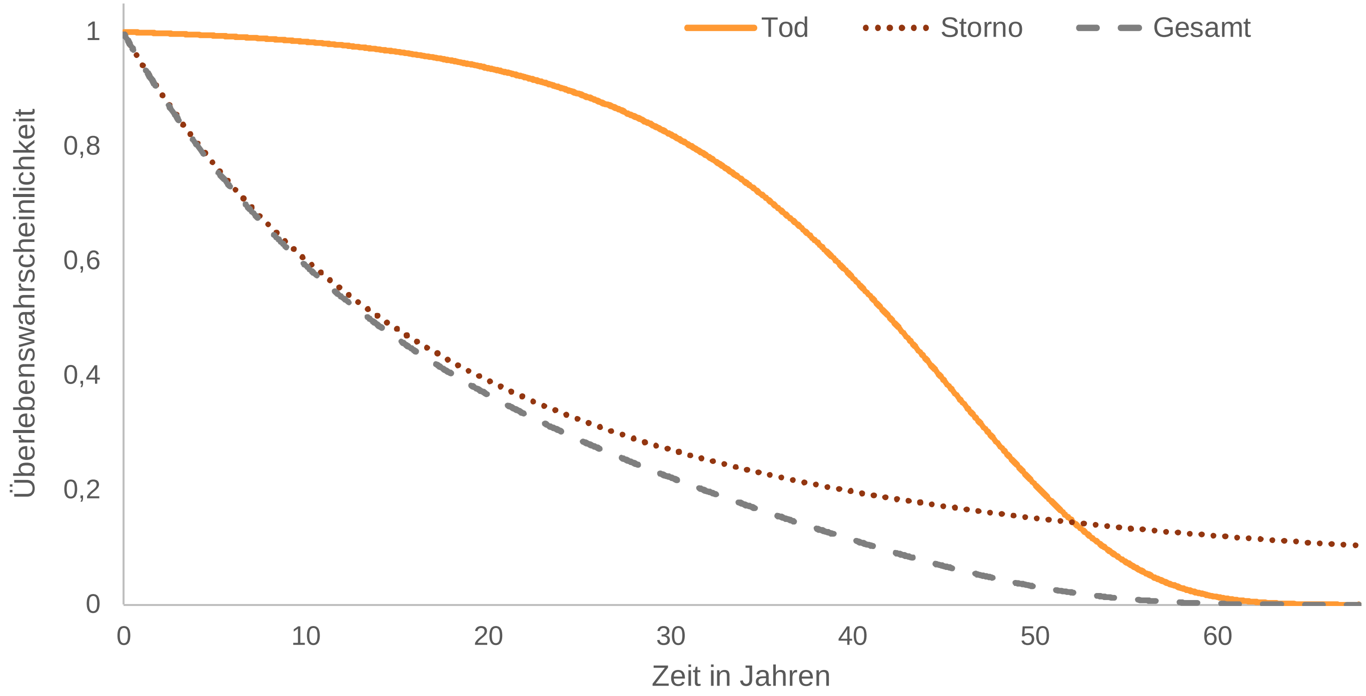
IV. Anwendung – Reservierung

Sicherheitsmargen explizit bestimmen

- Monte Carlo-Simulation des Bestandes
- Leistungen und Storno modellieren und vorhersagen
- Schätzrisiko und Zufallsrisiko messbar

IV. Anwendung – Reservierung

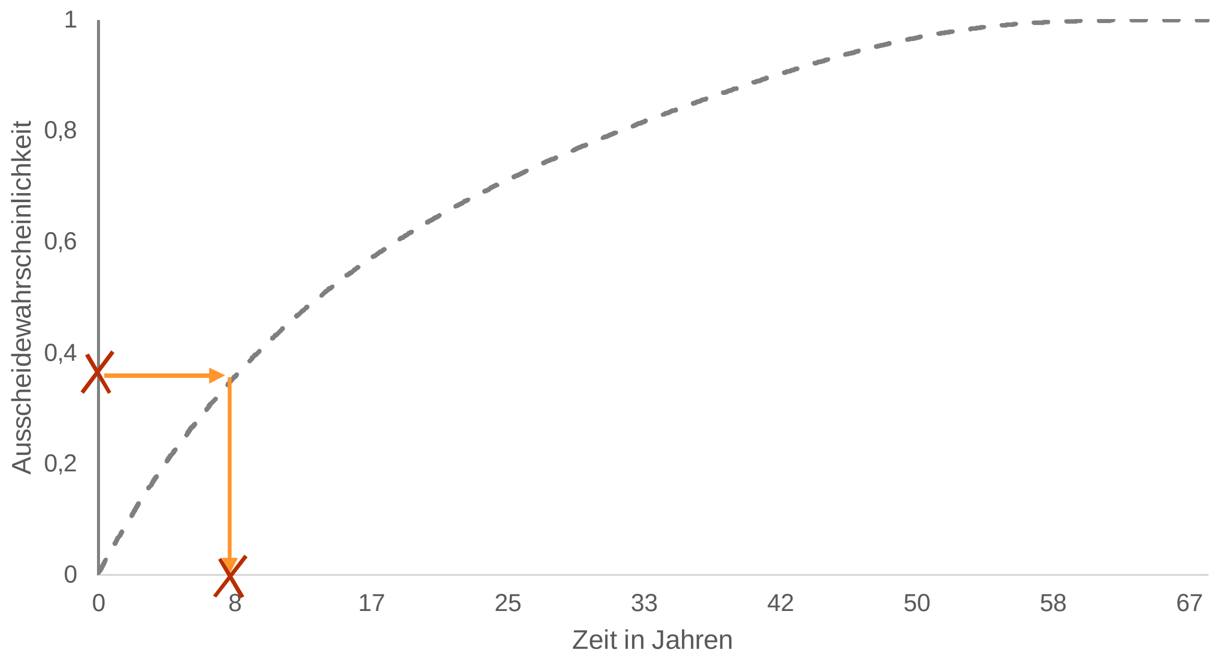
Multidekrement-Modell für Verweildauer



IV. Anwendung – Reservierung

Stochastisches Modell für Ausscheiden

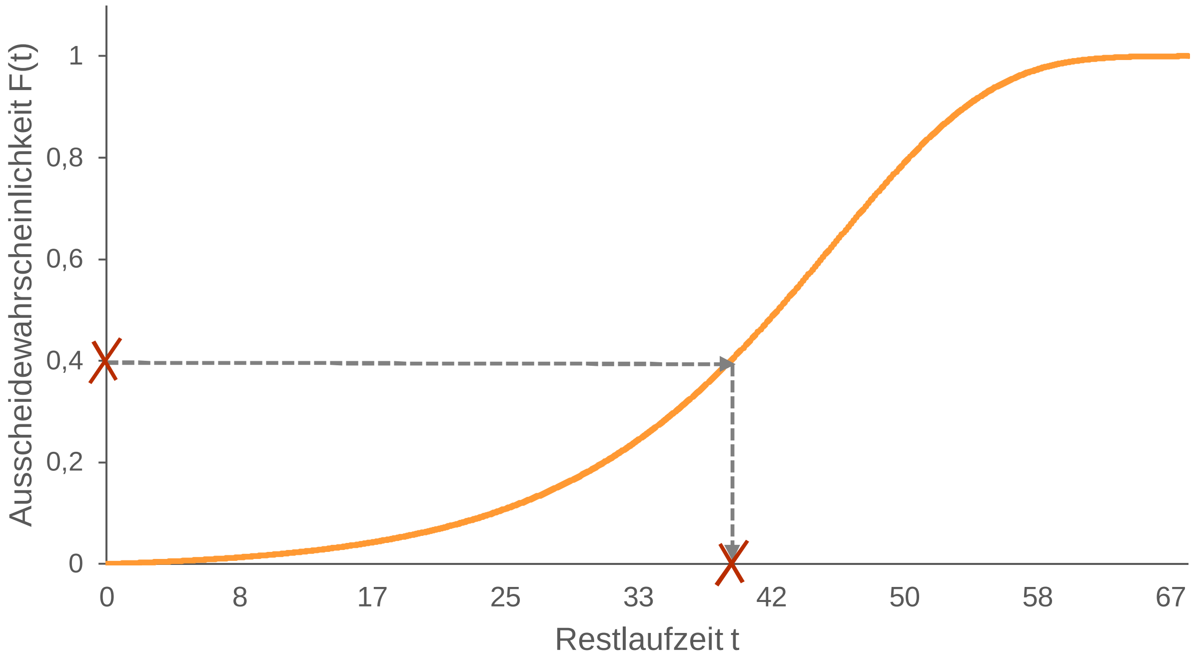
Ziehe gleich-
verteilte
Zufallszahl



IV. Anwendung – Reservierung

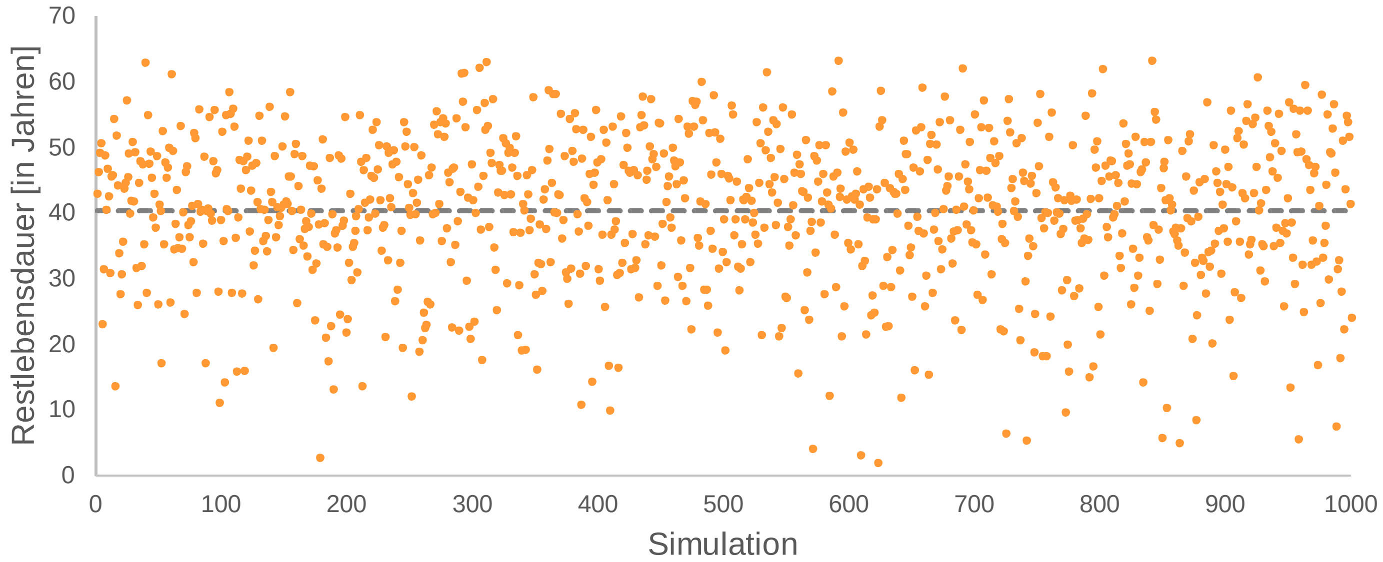
Stochastisches Modell für Todesfälle

Ziehe gleich-verteilte Zufallszahl



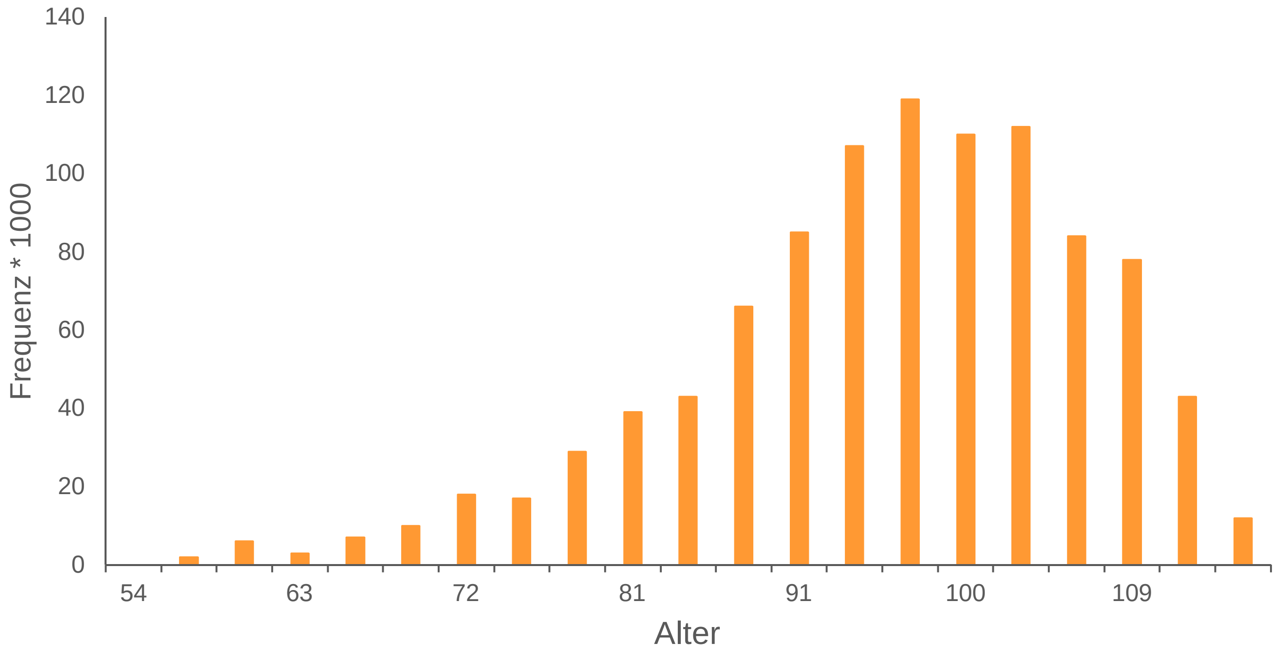
IV. Anwendung – Reservierung

Simulierte Restlebensdauern



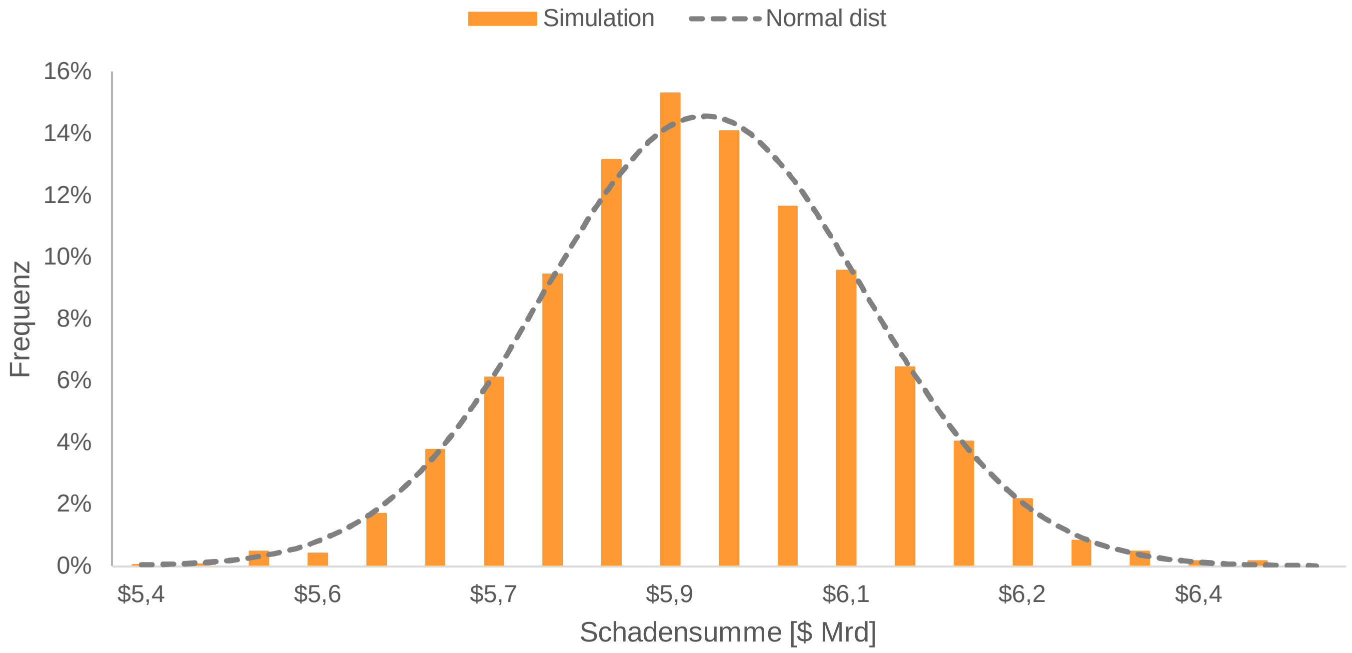
IV. Anwendung – Reservierung

Simulierte Restlebensdauern



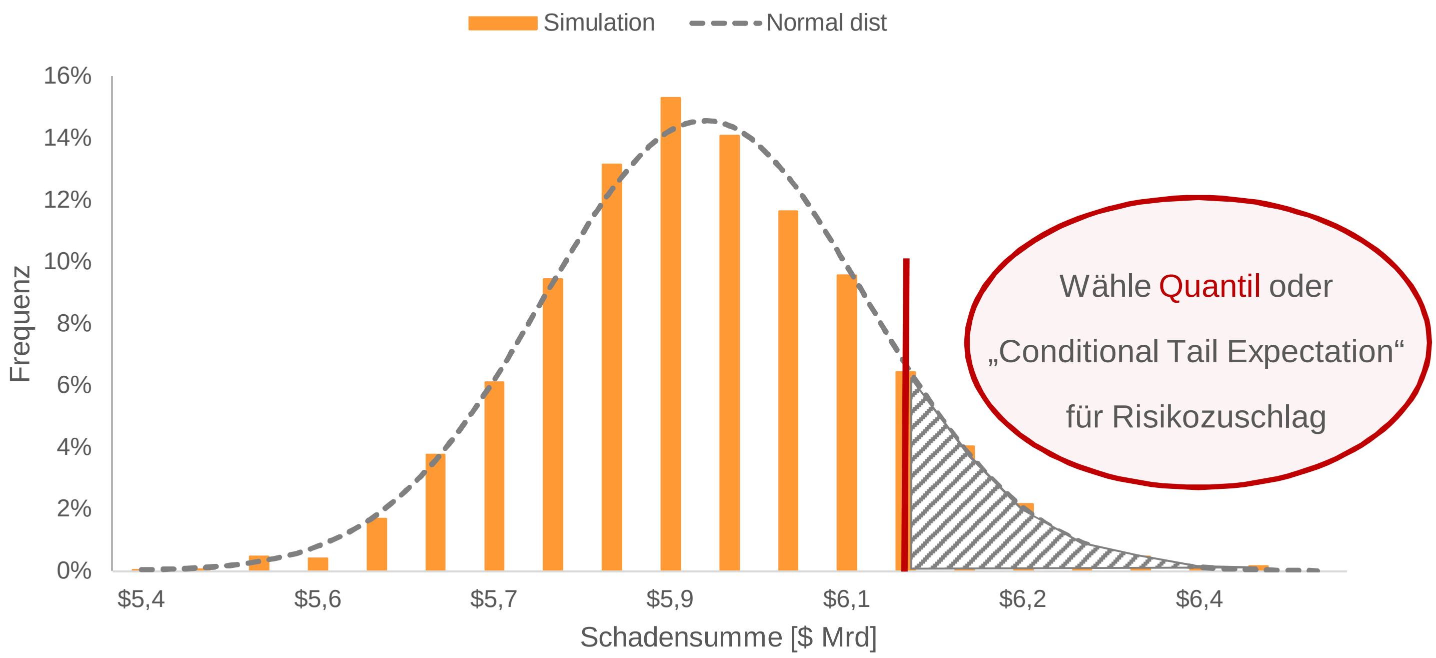
IV. Anwendung – Reservierung

Simulationsergebnisse



IV. Anwendung – Reservierung

Simulationsergebnisse



IV. Anwendung - SII Internes Modell

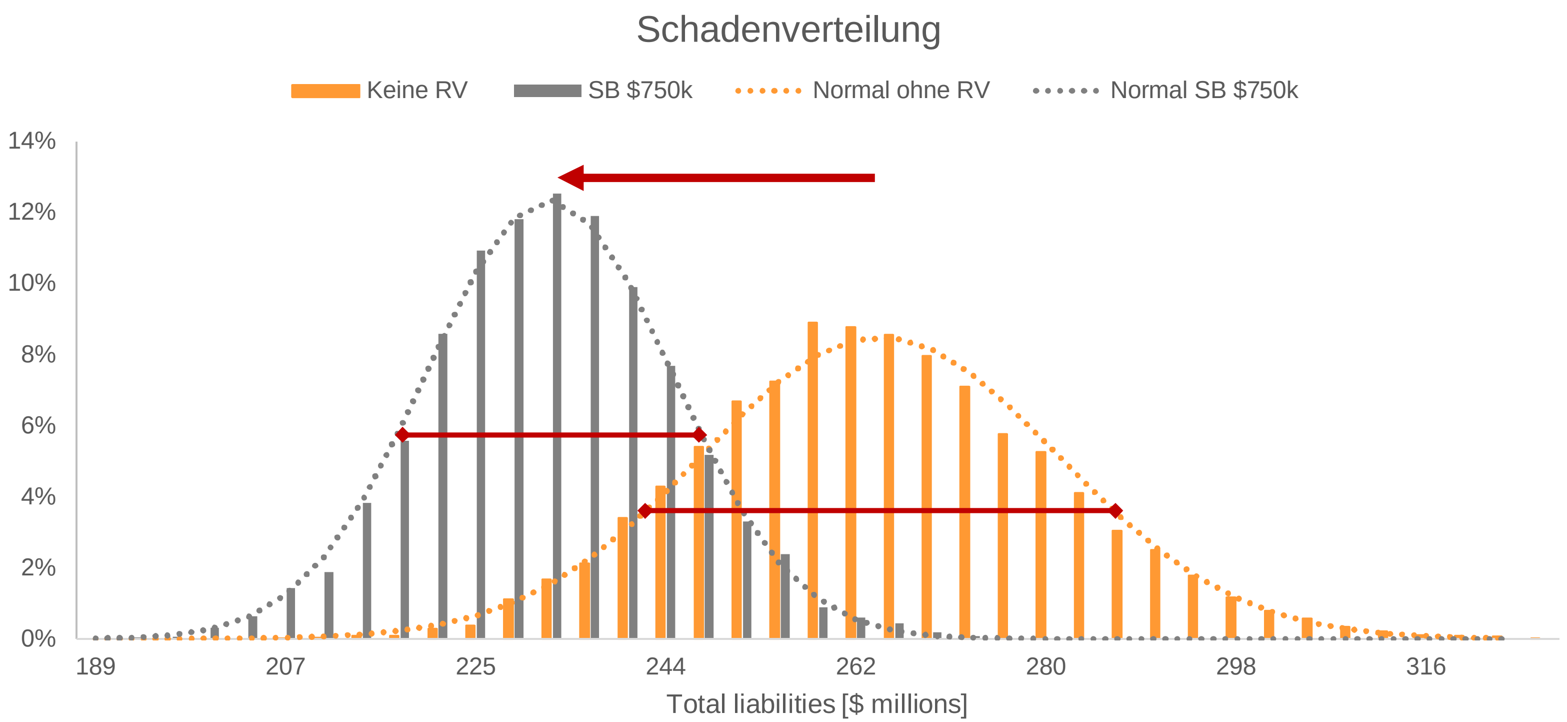
Methode für Solvency II zur Kalibrierung des SCR

- Sterblichkeit / Langlebigkeit
- Storno
- Morbidität / BU

Auch Auswirkungen von Rückversicherung auf die biometrischen Ergebnisse und deren Schwankung.

IV. Anwendung – Rückversicherung

Unterschiedliche Selbstbehalte



AdRes HEISTERMANN CONSULTING

IV. Anwendungen – Risikomanagement in der Biometrie

Überprüfung von Underwriting-Richtlinien

- Sind Zuschläge und ihre Höhe sachgerecht?
- Z.B. Modellierung eines Survival-Models mit Parameter „Zuschlag ja/nein“

Nutzung der Eingabedaten von Risikoprüfungsexpertensystemen

- Differenzierung der Inzidenzen nach Vorerkrankungen (BU)
- Voraussetzung: die Daten aus der Risikoprüfung behalten!

IV. Anwendungen – Risikomanagement in der Biometrie

Veränderungen über die Zeit

- Beispiel Sterbekassen:
  - Antiselektion steigt in den letzten Jahren

  ⇨ Veränderung der Zeichnungsrichtlinien

IV. Anwendungen – Risikomanagement in der Biometrie

Firmen- und Gruppengeschäft

- Der Wettbewerb in diesem Segment wird in den nächsten Jahren noch härter:
  - Margen in der Biometrie werden zunehmend weitergegeben und müssen deshalb genauer bewertet werden

IV. Anwendungen – Pricing, Marketing

Wer seine zielgruppenspezifische Inzidenzen kennt, kann das Marketing verfeinern

- Differenzierung im Preis oder in der Überschussbeteiligung
- Oder man nutzt diese Kenntnisse zur Selektion der Risiken und Vertriebssteuerung
- In Nichtleben machen wir das schon lange, warum eigentlich nur dort?

IV. Anwendungen – Pricing, Marketing

Zielgruppensteuerung

- Beispiel Rentenversicherung:
  - Sind „wohlhabende Kunden“ profitable Risiken?

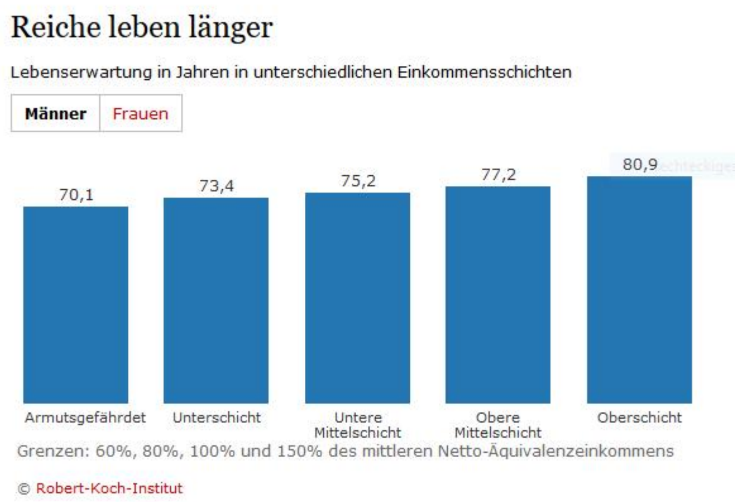
FAZ.net vom 4.3.17

Soziale Ungleichheit

Reiche leben länger

Die Lebenserwartung steigt mit dem Einkommen - vor allem bei Männern. Der Grund ist aber nicht allein das Geld.

04.03.2016, von PATRICK BERNAU



IV. Anwendungen – Pricing, Marketing

Zielgruppensteuerung

- Beispiel BU:
  - Ist die Berufsgruppe 1 profitabler als Berufsgruppe 4?
  - Die Antwort ist in hohem Maße von Ihrem Zugang zu den Zielgruppen und Zeichnungsrichtlinien abhängig.
  - Diese Frage können Sie daher nur in Ihrem eigenen Bestand beantworten!

IV. Anwendungen – Bewertung von Beständen

- Run Off der Klassikbestände wird ein Thema
  - Bei Verkauf von Beständen wird eine Bewertung notwendig
    - Ultimate Inzidenzraten bestimmen
  - Reichen die Reserven (BU- und Rentenbestände)?
- Run Off Dienstleister
  - Veränderungen bei den Ausscheiderechnungsgrundlagen (insbesondere Storno) beobachten
  - RV-Politik

V. Fazit

- Bestandsdaten werden besser, wenn man sie benutzt.
- Daten auf Einzelrisikobasis auswerten
- Bestandsspezifische Analysen schlagen Poollösungen.
- Auch Small Data ist wertvoll

Literaturhinweise

- D. Dickson, M. Hardy, H. Waters (2012): Actuarial Mathematics for Life Contingent Risks (Chapter 2: Survival Models)
- S.J. Richards (2016) Mis-estimation risk: measurement and impact, *British Actuarial Journal.*
- K. Kaufhold und W. Lennartz (2016) Optimizing Risk Retention, *SOA Research Paper.*

Kontakt

Bernd Heistermann

Tel: 02241 205280

bh@HeistermannConsulting.de

Kai Kaufhold

Tel: 02233 9464700

Kai.Kaufhold@AdReServices.com