

# 1 Statische Modelle

**Konvention:** Im folgenden seien alle Zufallsvariablen auf einem Wahrscheinlichkeitsraum  $(\Omega, \mathcal{A}, P)$  definiert.

Ein wichtiger Unterschied zwischen der Lebensversicherungsmathematik und der Schaden- (oder Sach-)versicherungsmathematik ergibt sich daraus, dass Lebensversicherungen i.d.R. eine deterministische Versicherungssumme vorsehen, während bei Schadenversicherungen die Schadenhöhe meist zufällig ist. Daher ist bei letzteren die Beschreibung (und die statistische Analyse) des versicherten Risikos komplexer.

Gegeben sei nun ein Portfolio aus Versicherungspolicen. Ziel soll es sein, das versicherte Risiko des gesamten Portfolios durch geeignete stochastische Modelle zu beschreiben.

Man kann dabei zwei Typen von stochastischen Modellen unterscheiden:

- Bei *statischen Modellen* wird das Zufallsverhalten der Schadenfälle in einer festen Periode (z.B. in einem Jahr) beschrieben, wobei der genaue Eintrittszeitpunkt des Schadens nicht modelliert wird. Dazu reichen meist Zufallsvariablen (Zv.) oder Zufallsvektoren aus.
- Bei *dynamischen Modellen* wird dagegen auch die zeitliche Entwicklung der Schadenfälle berücksichtigt, indem das gesamte Zufallsgeschehen durch stochastische Prozesse beschrieben wird.

In diesem Kapitel wollen wir uns auf statische Modelle beschränken.

Weiterhin unterscheidet man insbesondere bei statischen Modellen zwischen individuellen und kollektiven Modellen. (Dynamische individuelle Modell werden dagegen nur selten betrachtet.)

- Beim *individuellen Modell* wird in einem Portfolio von  $n$  versicherten Risiken für die  $i$ -te Police die zufällige Schadenhöhe  $Y_i \geq 0$  in der Versicherungsperiode betrachtet ( $1 \leq i \leq n$ ). Damit ist der Gesamtschaden des Portfolios gegeben durch

$$S_{ind} := \sum_{i=1}^n Y_i.$$

Seine Verteilung ist insbesondere durch die gemeinsame Verteilung  $P^{(Y_i)_{1 \leq i \leq n}}$  der Schäden der einzelnen Risiken festgelegt.

I.d.R. werden die  $Y_i$  als unabhängig angenommen, so dass

$$P^{(Y_i)_{1 \leq i \leq n}} = \bigotimes_{i=1}^n P^{Y_i},$$

und zur Beschreibung der Verteilung des Gesamtschadens die Kenntnis der (Rand-) Verteilungen der einzelnen Policen ausreicht. Die Modellierung dieser Verteilungen ist nun allerdings problematisch, da die Zufallsvariablen  $Y_i$  realistischerweise nicht als identisch verteilt angenommen werden dürfen, wenn das Portfolio nicht vollständig homogen ist, d.h. nur aus gleichartigen Risiken besteht.

Bei Lebensversicherungen wird das versicherte Risiko meist nur in Abhängigkeit von Alter, Geschlecht und Charakter der Versicherung (also Erlebensfallcharakter bei Rentenversicherungen u.ä., bzw. Todesfallcharakter bei Risikolebensversicherungen u.ä.) modelliert. Damit sind die homogenen Teilportfolios, die in diesen Kriterien übereinstimmen, oft groß genug, um etwa die Anwendung des Gesetzes der großen Zahlen oder des zentralen Grenzwertsatzes als gerechtfertigt erscheinen zu lassen (und so z.B. die relative Häufigkeit der Schadenfälle durch die Wahrscheinlichkeit eines Schadenfalls zu approximieren). Außerdem wird die statistische Analyse durch die Größe der homogenen Teilportfolios erheblich erleichtert.

Bei Schadenversicherungen werden dagegen i.d.R. sehr viel mehr Tarifmerkmale verwendet, die zudem oft eine größere Zahl möglicher Ausprägungen besitzen, um zwischen unterschiedlichen versicherten Risiken zu differenzieren. In solchen Fällen sind Teilportfolios mit identischer Ausprägung dieser Risikomerkmale mitunter zu klein, um Grenzwertsätze zu verwenden.

**1.1 Beispiel** In der Kfz-Haftpflichtversicherung werden stets die Tarifmerkmale Typklasse, Wohnort (in Form von Regionalklassen) und die Schadenfreiheitsklasse berücksichtigt, woraus sich bereits mehrere Tausend mögliche Kombinationen von Merkmalsausprägungen ergeben. Darüber hinaus wird oft noch unterschieden, ob der Versicherungsnehmer Beamter oder Landwirt ist oder einen sonstigen Beruf ausübt, ob eine Garage vorhanden ist, wie groß die jährlich zurückgelegte Strecke ist (wobei meist 2 oder 3 Klassen betrachtet werden) und ob nur der Versicherungsnehmer und ggf. sein Ehepartner den Wagen fahren, so dass sich die Anzahl möglicher Ausprägung auf bis zu 100.000 erhöhen kann. Bei einem einzelnen Versicherer werden daher zumindest einige der dadurch charakterisierten Teilportfolios nur wenige Policen enthalten.  $\square$

Aus den genannten Gründen spielt das individuelle Modell bei der Beschreibung des Gesamtrisikos eines Portfolios in der Risikotheorie eine untergeordnete Bedeutung. Es soll daher hier im Weiteren nicht mehr betrachtet werden.

- Bei einem *kollektiven Modell* werden die beobachteten Schadenhöhen im Portfolio nicht mehr den individuellen Policen zugeordnet, sondern z.B. einfach in der Reihenfolge ihres Auftretens registriert. Der Gesamtschaden des Portfolios wird dargestellt als

$$S_{koll} = \sum_{i=1}^N X_i,$$

wobei  $X_1, \dots, X_N > 0$  die in der Versicherungsperiode beobachteten Schadenhöhen sind und auch die Anzahl  $N$  der Schäden zufällig ist. (In diesem Skript verwenden wir stets die Konvention, dass eine leere Summe  $\sum_{i=1}^0 a_i$  gleich 0 gesetzt wird.)

Formal muss man daher eine ganze Folge von  $(0, \infty)$ -wertigen Zufallsvariablen  $X_i$ ,  $i \in \mathbb{N}$ , betrachten sowie eine  $\mathbb{N}_0$ -wertige Zufallsvariable  $N$ . Meist werden dabei  $(X_i)_{i \in \mathbb{N}}$  und  $N$  als unabhängig angenommen. Diese Annahme ist in manchen Fällen eine deutliche Vereinfachung der Realität. Z.B. ist es nicht unüblich, dass bei gewissen Wetterbedingungen eine relativ hohe Zahl relativ leichter Kfz-Unfälle eintritt;

dies könnte dazu führen, dass ein großer Wert von  $N$  tendenziell mit niedrigen Schadenhöhen  $X_i$  einhergeht, was der obigen Unabhängigkeitsannahme widersprechen würde.

Da nun die Schadenhöhen nicht mehr den individuellen Policen (mit ihren unterschiedlichen Risikoprofilen) zugeordnet werden, können die Zufallsvariablen  $X_i$  als identisch verteilt angenommen werden, wenn sich das versicherte Risiko über die betrachtete Versicherungsperiode nicht ändert. (Wenn man die beobachteten Schäden zusätzlich noch zufällig permutiert, kann auf diese einschränkende Annahme auch noch verzichtet werden.) Des weiteren werden die  $X_i$  meist auch als unabhängig angenommen (was wiederum eine gewisse vereinfachende Abbildung der Realität darstellt).

Die oben skizzierten Annahmen führen zu der folgenden Definition.

**1.2 Definition** *Beim Standardmodell der kollektiven Risikotheorie wird der Gesamtschaden eines Portfolios beschrieben durch die Zufallsvariable*

$$S_{koll} = \sum_{i=1}^N X_i,$$

wobei die  $(0, \infty)$ -wertigen Zufallsvariablen  $X_i$ ,  $i \in \mathbb{N}$ , unabhängig und identisch verteilt (i.i.d.) und gemeinsam von der  $\mathbb{N}_0$ -wertigen Zufallsvariable  $N$  unabhängig sind, d.h.

$$P^{((X_i)_{i \in \mathbb{N}}, N)} = (P^{X_1})^{\mathbb{N}} \otimes P^N. \quad \square$$

Es ist offensichtlich, dass eine solche kollektive Sichtweise nur geeignet ist, das Gesamtrisiko des Portfolios zu beschreiben, nicht aber, um eine angemessene Prämie für einzelne versicherte Risiken zu bestimmen. Auf diesen Aspekt der Schadenversicherungsmathematik werden wir in späteren Kapiteln eingehen.

## Zusammenhang zwischen individuellem und kollektivem Modell

Wenn man vereinfachend annimmt, dass jede Police in einer Periode maximal einen Schaden produziert, so erhält man anschaulich gesehen ein kollektives Modell aus einem individuellen, indem man zunächst zufällig eine Police aus dem betrachteten Portfolio heraus greift, wobei die Ziehungswahrscheinlichkeit für jede Police proportional zur Wahrscheinlichkeit sei, dass die Police einen (echten) Schaden (mit strikt positiver Schadenhöhe) produziert. In einem zweiten Schritt wird dann die zufällige Schadenhöhe gemäß der bedingten Verteilung der individuellen Schadenhöhe dieser Police ermittelt, gegeben dass die Schadenhöhe positiv ist. Bezeichnen wie oben  $Y_i$ ,  $1 \leq i \leq n$ , die Zv. der Schadenhöhen im individuellen Modell, so lässt sich dieses zweistufige Zufallsexperiment formal wie folgt beschreiben:

- (i) Sei  $I$  eine  $\{1, \dots, n\}$ -wertige Zufallsvariable mit

$$P\{I = i\} = \frac{P\{Y_i > 0\}}{\sum_{j=1}^n P\{Y_j > 0\}}.$$

- (ii) Ferner sei für alle  $1 \leq i \leq n$  eine von  $I$  unabhängige  $(0, \infty)$ -wertige Zv.  $Z_i$  gegeben mit  $P^{Z_i} = P^{Y_i | Y_i > 0}$ , d.h.  $P\{Z_i \in B\} = P(Y_i \in B \mid Y_i > 0) = P\{Y_i \in B\} / P\{Y_i > 0\}$  für alle  $B \in \mathcal{B}(0, \infty)$ .

Man erhält dann die Schadenzahl im kollektiven Modell als  $N = \sum_{i=1}^n 1_{\{Y_i > 0\}}$  und die Verteilung der Schadenhöhen als  $P^{X_1} = P^{Z_1}$ .

Die 2. Stufe des Zufallsexperiments lässt sich formal durch sog. Markov-Kerne beschreiben (s. Maßtheorie, Definition 4.1).

**1.3 Definition** Seien  $(\Omega_1, \mathcal{A}_1)$  und  $(\Omega_2, \mathcal{A}_2)$  zwei Messräume. Eine Abbildung  $K : \Omega_1 \times \mathcal{A}_2 \rightarrow [0, \infty]$  heißt **Übergangsmaß** von  $(\Omega_1, \mathcal{A}_1)$  nach  $(\Omega_2, \mathcal{A}_2)$ , wenn

(i)  $K(\cdot, A_2) : \omega_1 \mapsto K(\omega_1, A_2)$  für jedes feste  $A_2 \in \mathcal{A}_2$  eine  $\mathcal{A}_1, \mathcal{B}([0, \infty])$ -messbare Abbildung ist und

(ii)  $K(\omega_1, \cdot) : \mathcal{A}_2 \mapsto [0, \infty]$  für jedes feste  $\omega_1 \in \Omega_1$  ein Maß ist.

Ist  $K(\omega_1, \cdot)$  für alle  $\omega_1 \in \Omega_1$  sogar ein Wahrscheinlichkeitsmaß (W.maß), so heißt  $K$  **Übergangsw.maß** oder **Markov-Kern**.  $\square$

Ein Markov-Kern kann man also als Beschreibung des Zufallsverhaltens der 2. Stufe des Zufallsexperiments in Abhängigkeit des Ausgangs  $\omega_1$  der 1. Stufe auffassen. Ist zusätzlich ein W.maß  $Q$  auf  $(\Omega_1, \mathcal{A}_1)$  gegeben, das die 1. Stufe des Zufallsexperiments beschreibt, so erhält man daraus eine stochastische Beschreibung des Gesamtexperiments:

**1.4 Satz und Definition** Seien ein W.maß  $Q$  auf  $(\Omega_1, \mathcal{A}_1)$  und ein Markov-Kern  $K$  von  $(\Omega_1, \mathcal{A}_1)$  nach  $(\Omega_2, \mathcal{A}_2)$  gegeben. Dann definiert

$$Q \otimes K(A) := \int_{\Omega_1} \int_{\Omega_2} 1_A(\omega_1, \omega_2) K(\omega_1, d\omega_2) Q(d\omega_1), \quad A \in \mathcal{A}_1 \otimes \mathcal{A}_2,$$

ein W.maß  $Q \otimes K$  auf  $(\Omega_1 \times \Omega_2, \mathcal{A}_1 \otimes \mathcal{A}_2)$ , die **Koppelung** von  $Q$  und  $K$ .  $\square$

BEWEIS. s. Maßtheorie, Satz und Definition 4.4.  $\square$

**1.5 Bemerkung** Implizit besagt dieses Resultat auch, dass die Abbildung

$$\omega_1 \mapsto \int_{\Omega_2} 1_A(\omega_1, \omega_2) K(\omega_1, d\omega_2)$$

$\mathcal{A}_1, \mathcal{B}[0, 1]$ -messbar ist für alle  $A \in \mathcal{A}_1 \otimes \mathcal{A}_2$ .  $\square$

Wenn nicht der Ausgang des Gesamtexperiments von Interesse ist, sondern nur das Ergebnis der 2. Stufe des Experiments, so reicht es, das W.maß  $\mathcal{A}_2 \ni A_2 \mapsto Q \otimes K(\Omega_1 \times A_2)$  auf  $(\Omega_2, \mathcal{A}_2)$  zu betrachten. Dieses lässt sich als Mischung der W.masse  $K(\omega_1, \cdot)$  interpretieren, wobei diese gewissermaßen mit den Eintrittsw.  $Q\{\omega_1\}$  gewichtet werden. (Dies ist so allerdings nur bei einem diskreten W.raum  $(\Omega_1, \mathcal{A}_1, Q)$  formal richtig.)

**1.6 Korollar und Definition** (i) Seien  $Q_i, 1 \leq i \leq n$ , W.maße auf einem Messraum  $(\mathcal{X}, \mathcal{C})$  und  $p_i \geq 0, 1 \leq i \leq n$ , mit  $\sum_{i=1}^n p_i = 1$ , so definiert

$$Q(C) = \sum_{i=1}^n p_i Q_i(C), \quad C \in \mathcal{C},$$

ein W.maß auf  $(\mathcal{X}, \mathcal{C})$ , die **(endliche) Mischung der  $Q_i$  mit den Gewichten  $p_i$** .

(ii) Seien  $(\Theta, \mathcal{S})$  ein Messraum und  $Q_\vartheta, \vartheta \in \Theta$ , W.maße auf  $(\mathcal{X}, \mathcal{C})$  so, dass  $K : (\vartheta, C) \mapsto Q_\vartheta(C)$  ein Markov-Kern von  $(\Theta, \mathcal{S})$  nach  $(\mathcal{X}, \mathcal{C})$  ist, und sei  $Q_\Theta$  ein W.maß auf  $(\Theta, \mathcal{S})$ , so definiert

$$Q(C) = \int_{\Theta} Q_\vartheta(C) Q_\Theta(d\vartheta) = Q_\Theta \otimes K(\Theta \times C), \quad C \in \mathcal{C},$$

ein W.maß auf  $(\mathcal{X}, \mathcal{C})$ , die **Mischung der  $Q_\vartheta$  mit Mischungsmaß  $Q_\Theta$  oder kurz Mischung der  $Q_\vartheta$  unter  $Q_\Theta$** . Bezeichnet  $pr_2 : \Theta \times \mathcal{X} \rightarrow \mathcal{X}, (\vartheta, x) \mapsto x$  die Projektion auf die 2. Koordinate, so ist  $Q$  gerade die Verteilung von  $pr_2$  unter der Koppelung  $Q_\Theta \otimes K$ .  $\square$

**1.7 Bemerkung** Satz 1.6 (i) ist ein Spezialfall von Satz 1.6 (ii) mit  $(\Theta, \mathcal{S}) = (\{1, \dots, n\}, 2^{\{1, \dots, n\}})$  und  $Q_\Theta = \sum_{i=1}^n p_i \varepsilon_i$ . Hierbei bezeichnet wie üblich  $\varepsilon_i$  das Dirac-Maß mit Masse 1 im Punkt  $i$ , d.h.  $\varepsilon_i(A) = 1_A(i)$  für alle  $A \subset \{1, \dots, n\}$ .  $\square$

BEWEIS. Gemäß Satz 1.4 gilt:

$$\begin{aligned} (Q_\Theta \otimes K)^{pr_2}(C) &= Q_\Theta \otimes K(pr_2^{-1}(C)) \\ &= Q_\Theta \otimes K(\Theta \times C) \\ &= \int_{\Theta} \int_{\mathcal{X}} \underbrace{1_{\Theta \times C}(\vartheta, x)}_{=1_C(x)} K(\vartheta, dx) Q_\Theta(d\vartheta) \\ &= \int_{\Theta} K(\vartheta, C) Q_\Theta(d\vartheta) \\ &= \int_{\Theta} Q_\vartheta(C) Q_\Theta(d\vartheta) \\ &= Q(C) \end{aligned} \quad \square$$

In der oben genannten Situation erhält man also für die Verteilung der Schadenhöhen im kollektiven Modell

$$\begin{aligned} P\{X_1 \in B\} &= P\{Z_I \in B\} \\ &= \sum_{i=1}^n P\{I = i, Z_i \in B\} \\ &= \sum_{i=1}^n \frac{P\{Y_i > 0\}}{\sum_{j=1}^n P\{Y_j > 0\}} P(Y_i \in B \mid Y_i > 0) \quad \text{für alle } B \in \mathcal{B}(0, \infty), \end{aligned}$$

d.h.  $P^{X_1}$  ist die endliche Mischung der W.mäße  $Q_i := P^{Y_i|Y_i>0}$  mit den Gewichten  $p_i = P\{Y_i > 0\} / \sum_{j=1}^n P\{Y_j > 0\}$ .

Im folgenden werden wir stets das Standardmodell der kollektiven Risikothorie voraussetzen, wenn nicht explizit ein anderes Modell angegeben wird.

Als Grundlage für die Tarifierung werden zunächst realistische Verteilungsannahmen für die Schadenhöhen und die Schadenzahl benötigt.

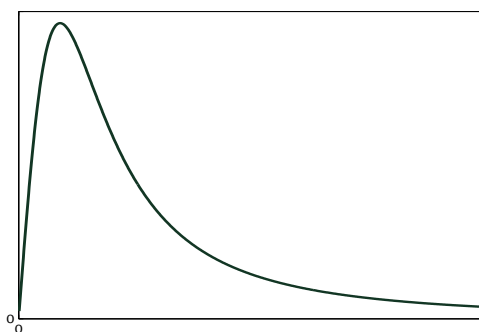
## Modelle für die Schadenhöhenverteilung

Typischerweise hat die empirische Schadenhöhenverteilung die folgenden Charakteristika:

- Sehr kleine Schäden treten relativ selten auf. (Dies lässt sich z.B. dadurch erklären, dass Versicherungsnehmer solche Schäden dem Versicherer gar nicht melden, oder dass bei fälligen Reparaturarbeiten minimale Fixkosten anfallen.)
- Mitttelgroße Schäden treten besonders oft auf.
- Sehr große Schäden sind wiederum selten.

Oft scheint dabei die Eintrittswahrscheinlichkeit in gleich großen Klassen von Schadenhöhen zunächst anzusteigen und ab einem Punkt wieder abzufallen. Möchte man die Schadenhöhen durch eine Verteilung modellieren, die eine (Lebesgue-)Dichte besitzt, so sind daher insbesondere unimodale Dichten von besonderem Interesse, also Dichten mit nur einer lokalen Maximalstelle (Modus), die zudem meist als strikt positiv angenommen wird.

Skizze:

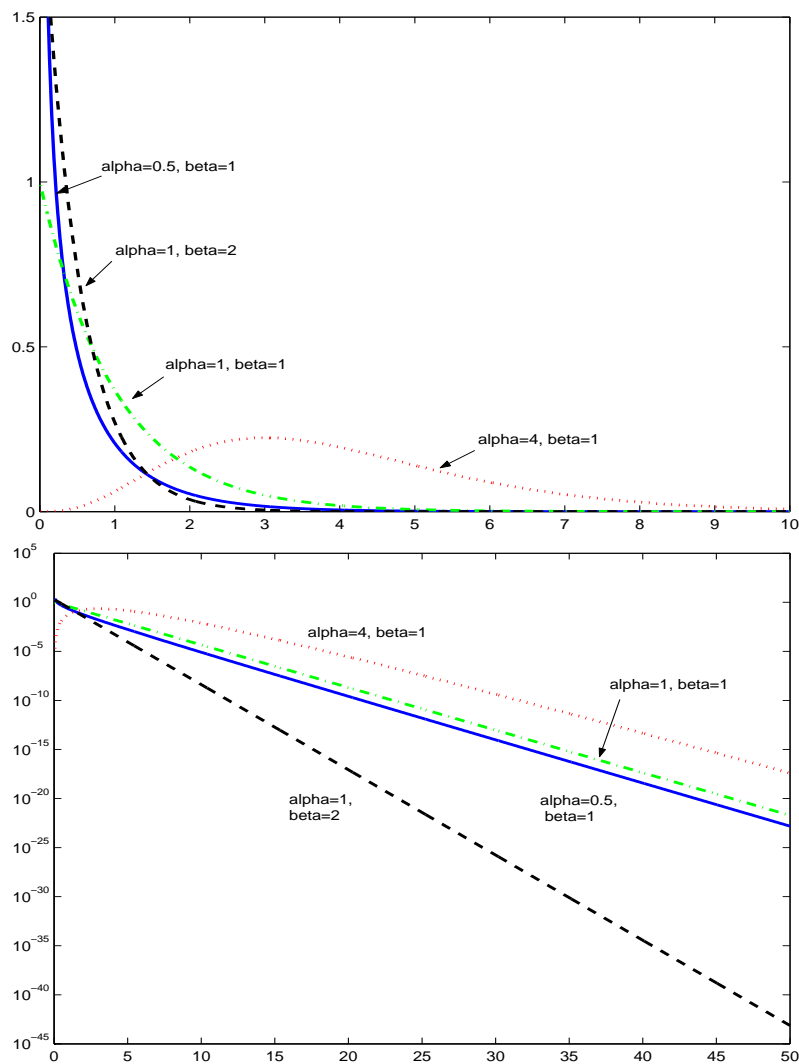


Es folgt eine (nicht erschöpfende) Liste mit in der Praxis oft für die Modellierung von Schadenhöhen verwendeten Verteilungen mit (Lebesgue-)Dichten  $f(x)$ ,  $x > 0$ .

- *Gammaverteilung*  $\Gamma_{\alpha,\beta}$ ,  $\alpha, \beta > 0$ :  
Die Dichte

$$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, \quad x > 0,$$

ist im Fall  $\alpha \in (0, 1]$  monoton fallend mit  $\lim_{x \downarrow 0} f(x) = \infty$ , falls  $\alpha < 1$ , und sie ist unimodal mit Modus  $(\alpha - 1)/\beta$  im Fall  $\alpha > 1$ . Der Parameter  $1/\beta$  ist ein Skalenparameter, der die Verteilungsfunktion (Vf.)  $\Gamma_{\alpha,1}$  parallel zur  $x$ -Achse staucht ( $\beta > 1$ ) bzw. streckt ( $\beta < 1$ ). Der Parameter  $\alpha$  bestimmt dagegen die Form der Verteilungsfunktion bzw. der Dichte. Im Fall  $\alpha = 1$  erhält man speziell eine Exponentialverteilung.



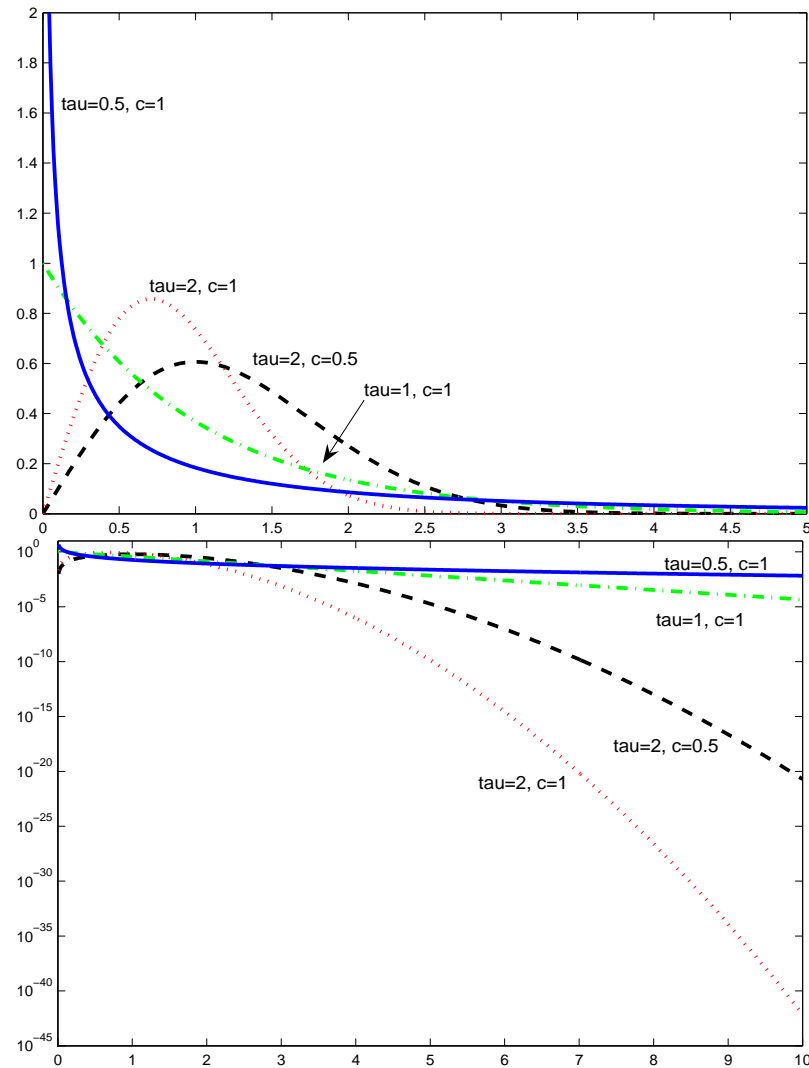
Dichte der Gammaverteilung für verschiedene Parameterkonstellationen mit linear (oben) bzw. logarithmisch (unten) skaliertem  $y$ -Achse.

- Weibull-Verteilung  $W_{c,\tau}$ ,  $c, \tau > 0$ :

Die Dichte

$$f(x) = c\tau x^{\tau-1} e^{-cx^\tau}, \quad x > 0,$$

ist im Fall  $\tau \in (0, 1]$  monoton fallend mit  $\lim_{x \downarrow 0} f(x) = \infty$ , falls  $\tau < 1$ , und sie ist unimodal mit Modus  $((\tau - 1)/(c\tau))^{1/\tau}$  im Fall  $\tau > 1$ . Der Parameter  $\tau$  bestimmt die Form der Dichte, während  $c^{-1/\tau}$  ein Skalenparameter ist. Im Fall  $\tau = 1$  erhält man speziell eine Exponentialverteilung, im Fall  $\tau > 1$  ( $\tau < 1$ ) fällt die Dichte für  $x \rightarrow \infty$  schließlich schneller (langsamer) als jede Exponentialfunktion.



Dichte der Weibull-Verteilung für verschiedene Parameterkonstellationen mit linear (oben) bzw. logarithmisch (unten) skaliertem  $y$ -Achse.

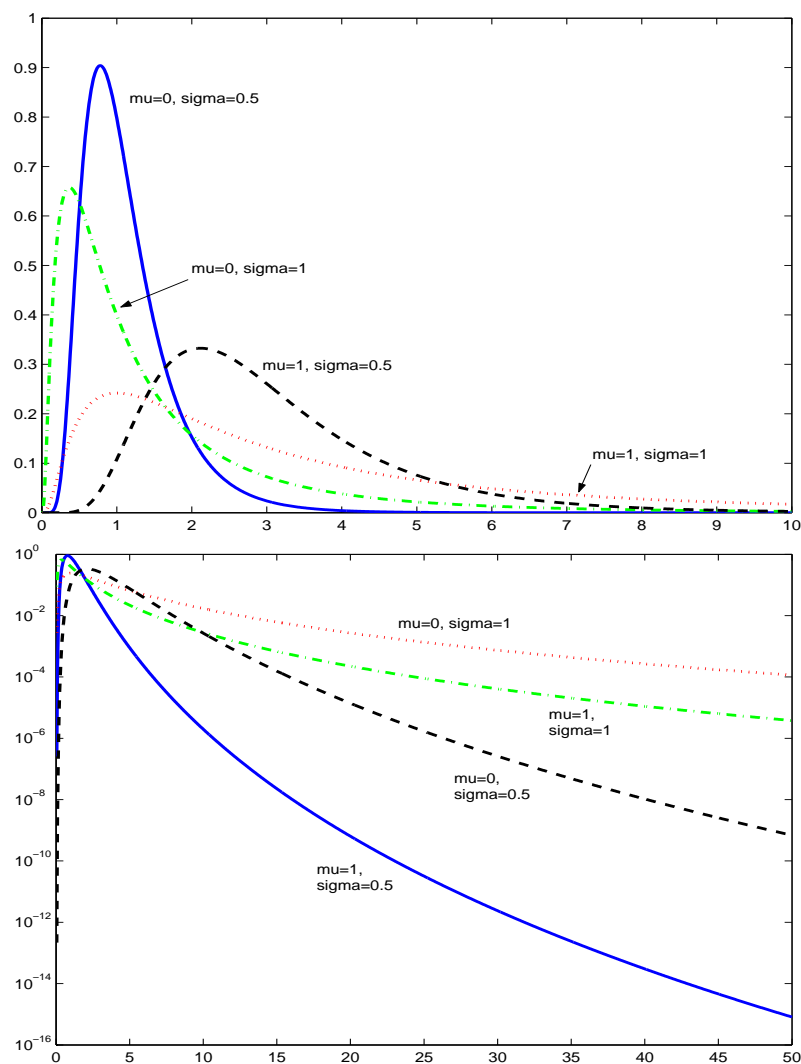


- *Log-Normalverteilung*  $LN_{\mu,\sigma^2}$ ,  $\mu \in \mathbb{R}$ ,  $\sigma > 0$ :

Ist  $X$  gemäß  $LN_{\mu,\sigma^2}$  verteilt, so ist  $\log X$  gerade  $\mathcal{N}_{(\mu,\sigma^2)}$ -verteilt. Die Dichte

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma x} \exp\left(-\frac{(\log x - \mu)^2}{2\sigma^2}\right), \quad x > 0,$$

ist stets unimodal mit Modus  $e^\mu$ . Sie fällt für  $x \rightarrow \infty$  langsamer ab als jede Exponentialfunktion, aber schneller als jede Potenz von  $x$  mit negativem Exponenten.

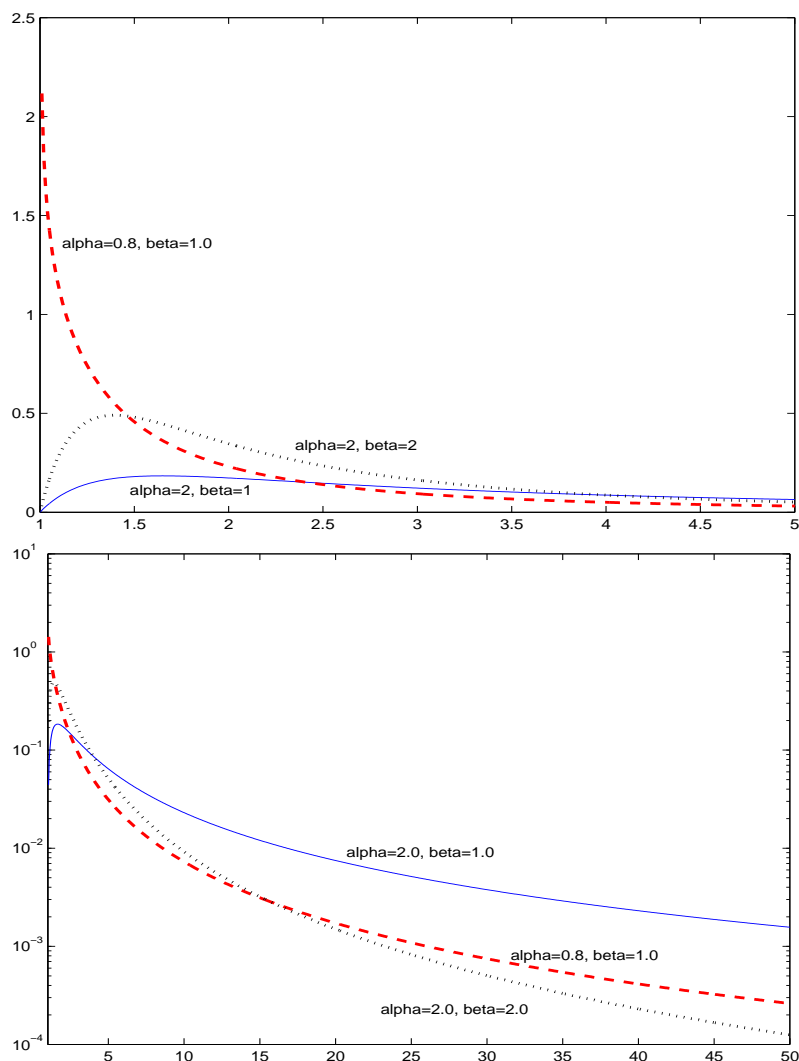


Dichte der Log-Normalverteilung für verschiedene Parameterkonstellationen mit linear (oben) bzw. logarithmisch (unten) skaliertem  $y$ -Achse.

- *Log-Gammaverteilung*  $LG_{\alpha,\beta}$ ,  $\alpha, \beta > 0$ :  
Ist  $X$  gemäß  $LG_{\alpha,\beta}$  verteilt, so ist  $\log X$  gerade  $\Gamma_{\alpha,\beta}$ -verteilt. Die Dichte

$$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} (\log x)^{\alpha-1} x^{-(\beta+1)} 1_{(1,\infty)}(x)$$

ist auf  $(1, \infty)$  monoton fallend, falls  $\alpha \leq 1$ , mit  $\lim_{x \downarrow 1} f(x) = \infty$  im Fall  $\alpha < 1$ , und sie ist unimodal mit Modus  $\exp((\alpha - 1)/(\beta + 1))$ , falls  $\alpha > 1$ . Die Dichte fällt “im Wesentlichen” wie die Potenz  $x^{-(\beta+1)}$  ab.



Dichte der Log-Gammaverteilung für verschiedene Parameterkonstellationen mit linear (oben) bzw. logarithmisch (unten) skaliertem  $y$ -Achse.

- *Pareto-Verteilung*  $P_{\alpha,\mu,\sigma}$ ,  $\alpha, \sigma > 0$ ,  $\mu \geq -\sigma$ :

Die Dichte

$$f(x) = \frac{\alpha}{\sigma} \left( \frac{x - \mu}{\sigma} \right)^{-(\alpha+1)} 1_{(\mu+\sigma, \infty)}(x)$$

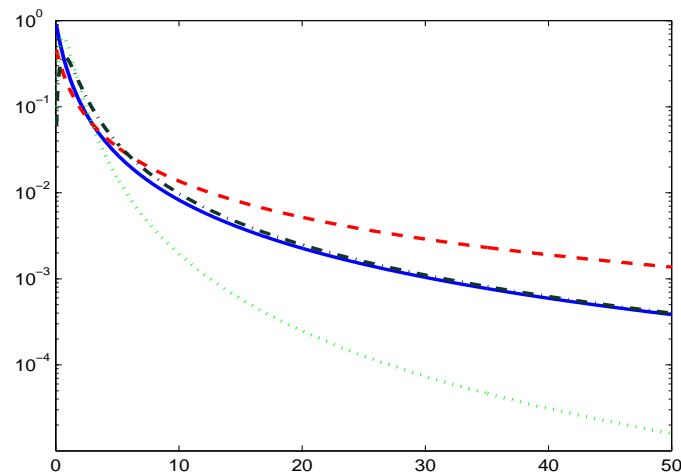
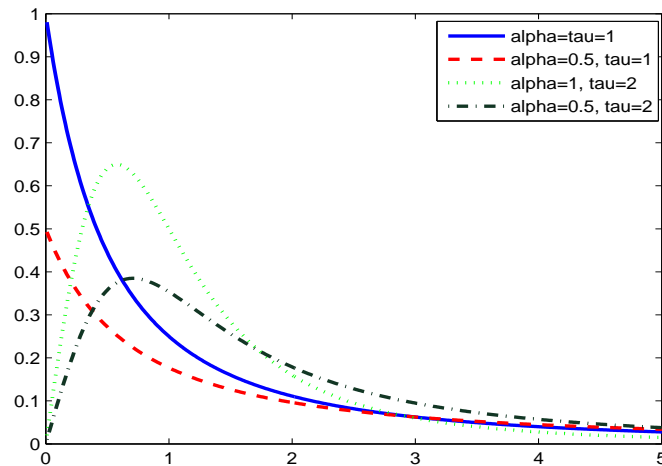
ist auf  $(\mu + \sigma, \infty)$  monoton fallend. Der Parameter  $\alpha$  bestimmt die Geschwindigkeit, mit der  $f(x)$  für  $x \rightarrow \infty$  gegen 0 konvergiert,  $\mu$  ist ein Lokationsparameter, um den die Vf.  $P_{\alpha,\mu,\sigma}$  bzw. die zugehörige Dichte gegenüber der Vf. bzw. Dichte zum Parameter  $\mu = 0$  parallel zur  $x$ -Achse verschoben wird,  $\sigma$  ist ein Skalenparameter, der die Stauchung  $\sigma < 1$  bzw. Streckung  $\sigma > 1$  der Vf. quantifiziert.

- *Burr-Verteilung*  $Burr_{\alpha,\tau,\sigma}$ ,  $\alpha, \tau, \sigma > 0$ :

Die Dichte

$$f(x) = \frac{\alpha\tau}{\sigma} \left( \frac{x}{\sigma} \right)^{\tau-1} \left( 1 + \left( \frac{x}{\sigma} \right)^\tau \right)^{-(\alpha+1)}, \quad x > 0,$$

ist monoton fallend, falls  $\tau \geq 1$  und unimodal mit Modus  $\sigma((\tau - 1)/(\alpha\tau + 1))^{1/\tau}$  sonst. Die Parameter  $\alpha$  und  $\tau$  sind Formparameter, wobei die Dichte für  $x \rightarrow \infty$  wie  $x^{-\alpha\tau-1}$  abfällt;  $\sigma$  ist wiederum ein Skalenparameter.



Dichte der Burr-Verteilung für verschiedene Parameterkonstellationen mit linear (oben) bzw. logarithmisch (unten) skaliertem  $y$ -Achse; es ist stets  $\sigma = 1$ .

Wird eine dieser (oder anderer) Verteilungsfamilien als angemessen angesehen, so sind die unbekannten Parameter noch durch geeignete statistische Verfahren (z.B. Maximum Likelihood-Schätzer oder Momenten-Schätzer) zu schätzen. Außerdem sollten die Modellvoraussetzungen *stets* wenigstens durch geeignete graphische Methoden zur Modellvalidierung wie QQ-Plots oder Plots der empirischen *mean excess* Funktion (besser noch durch formale statistische Tests) auf Plausibilität überprüft werden, da nicht gerechtfertigte Modellannahmen zu einer massiven Fehleinschätzung des versicherten Risikos führen können. Die Bedeutung dieses in der Praxis leider oft nicht hinreichend gewürdigten Analyseschritts kann kaum überschätzt werden!

## Modelle für die Schadenzahlverteilung

Zum Abschluss dieses Kapitels sollen einige Modelle für die zufällige Schadenanzahl  $N$  im Portfolio diskutiert werden. Die folgenden Verteilungen sind in diesem Zusammenhang besonders populär:

- *Binomialverteilung:*

$$P\{N = k\} = \mathcal{B}_{(n,p)}\{k\} = \binom{n}{k} p^k (1-p)^{n-k}, \quad 0 \leq k \leq n,$$

für gewisse  $n \in \mathbb{N}$  und  $p \in (0, 1)$ .

- *Poisson-Verteilung:*

$$P\{N = k\} = \mathcal{P}_\lambda\{k\} = e^{-\lambda} \frac{\lambda^k}{k!}, \quad k \in \mathbb{N}_0,$$

für ein  $\lambda > 0$ .

- *negative Binomialverteilung:*

$$P\{N = k\} = \mathcal{B}_{(\beta,p)}^-\{k\} = \binom{\beta + k - 1}{k} p^\beta (1-p)^k, \quad k \in \mathbb{N}_0,$$

für gewisse  $\beta \in (0, \infty)$  und  $p \in (0, 1)$ , wobei

$$\binom{\beta + k - 1}{k} := \frac{(\beta + k - 1) \cdot (\beta + k - 2) \cdots \beta}{k!}.$$

**1.8 Bemerkung** (i) Die Binomialverteilung lässt sich natürlich insbesondere durch den oben skizzierten Zusammenhang zwischen dem individuellen und dem kollektiven Modell motivieren, wenn Schäden der individuellen Policen unabhängig voneinander mit der gleichen W.  $p$  eintreten. Da die Annahme einer identischen Schadeneintrittsw. aber sehr restriktiv ist, wird diese Verteilungsannahme von den drei genannten am seltensten verwendet.

(ii) Die Poisson-Verteilung ist insbesondere geeignet, die Schadenzahl zu modellieren, wenn die W. für den Eintritt eines Schadens für jede Police klein ist.

Gilt z.B. in der unter (i) geschilderten Situation bei steigendem Portfolioumfang  $np \rightarrow \lambda$ , so gilt nach dem Poissonschen Grenzwertsatz für die Gesamtzahl der Schäden

$$P\{N = k\} \rightarrow \mathcal{P}_\lambda\{k\} \quad \forall k \in \mathbb{N}_0.$$

Allgemeiner gilt sogar: Sind  $Z_{n,i}$ ,  $1 \leq i \leq n$ , für alle  $n \in \mathbb{N}$  unabhängige  $\{0, 1\}$ -wertige Zv., so gilt für  $n \rightarrow \infty$

$$P\left\{\sum_{i=1}^n Z_{n,i} = k\right\} \rightarrow \mathcal{P}_\lambda\{k\} \quad \forall k \in \mathbb{N}_0,$$

falls für die  $p_{n,i} := P\{Z_{n,i} = 1\}$  gilt:  $\max_{1 \leq i \leq n} p_{n,i} \rightarrow 0$  und  $\sum_{i=1}^n p_{n,i} \rightarrow \lambda$ ; in dem Fall ist  $\lambda$  dann die erwartete Schadenzahl. Liegt also konkret ein großes Portfolio vor, bei dem für jede Police unabhängig voneinander mit geringer W. ein Schaden eintritt, so sollte die Poisson-Verteilung die zufällige Schadenzahl im Portfolio gut beschreiben.

- (iii) Das Poisson-Modell ist allerdings insofern sehr einschränkend, als bei einer Poisson-verteilten Schadenzahl der Erwartungswert und die Varianz identisch sind. Tatsächlich beobachtet man aber oft, dass die Varianz deutlich größer als die erwartete Schadenzahl zu sein scheint.

Eine solche erhöhte Varianz (*Überdispersion*) kann man z.B. dadurch erhalten, dass man statt einer festen Poisson-Verteilung eine Mischung von verschiedenen Poisson-Verteilungen über den Parameter  $\lambda$  betrachtet, d.h.

$$P\{N = k\} = \int_0^\infty \mathcal{P}_\lambda\{k\} Q(d\lambda),$$

wobei  $Q$  eine geeignete Mischungsverteilung ist. Dieser Modellansatz lässt sich wie folgt anschaulich motivieren: Die Argumentation aus (ii) legt nahe, dass man die Schadenzahlverteilung gut durch eine Poisson-Verteilung approximieren kann, wobei der Parameter  $\lambda$  aber von gewissen Rahmenbedingungen abhängen kann, die auf alle Policen zugleich wirken; z.B. kann ein Winter mit viel Schnee und Glatteis zu vermehrten Verkehrsunfällen führen oder ein stürmisches Winterhalbjahr zu einer ungewöhnlich hohen Zahl von Schadenfällen in der Wohngebäudeversicherung. Modelliert man diese Rahmenbedingungen ebenfalls als zufällig, so führt dies gerade dazu, dass man statt einer festen Poisson-Verteilung eine Mischung von Poisson-Verteilungen betrachtet, wobei die Verteilung  $Q$  die Eintrittsw. der unterschiedlichen Szenarien von Rahmenbedingungen beschreibt.

Wählt man speziell  $Q = \Gamma_{\alpha,\beta}$  als Mischungsverteilung, so führt dieser Ansatz gerade auf die negative Binomialverteilung  $\mathcal{B}_{(\alpha,\beta/(\beta+1))}^-$  (s. Übungen).  $\square$

Prinzipiell lassen sich die oben genannten Schätzverfahren auch verwenden, um die unbekannten Parameter der Verteilung von  $N$  zu bestimmen.<sup>1</sup> Zu diesem Zweck werden allerdings die beobachteten Schadenzahlen aus mehreren Versicherungsperioden benötigt,

<sup>1</sup>Beim Maximum Likelihood-Verfahren ist dazu beruhend auf den beobachteten Schadenzahlen  $n_1, \dots, n_k$  der Vorjahre die Wahrscheinlichkeit  $\prod_{i=1}^k P\{N = n_i\}$  zu maximieren.

was insbesondere deshalb problematisch ist, weil sich die Größe des Versicherungsportfolios mit der Zeit ändert und daher lang zurückliegende Versicherungsperioden kaum Rückschlüsse auf die in der aktuellen Periode zu erwartenden Schadenzahl erlauben. Diesen Effekt kann man beim Poisson-Modell dadurch (teilweise) berücksichtigen, dass man den Parameter  $\lambda$  als proportional zu dem Portfolioumfang  $L$  (also i.d.R. zu der Anzahl von Policen) modelliert, d.h. in  $k$  Versicherungsperioden mit Portfolioumfängen  $L_1, \dots, L_k$  werden die beobachteten Schadenzahlen  $n_1, \dots, n_k$  als Realisierungen von unabhängigen  $\mathcal{P}_{L_i, \vartheta}$ -verteilten Zufallsvariablen aufgefasst, wobei nur  $\vartheta$  unbekannt ist. Bei diesem Ansatz ist es realistischer,  $\vartheta$  über einen längeren Zeitraum als konstant anzunehmen, als die erwartete Schadenzahl  $\lambda$  selber als konstant anzusetzen. Der Parameter  $\vartheta$  kann dann geschätzt werden durch den Maximum-Likelihood-Schätzer

$$\hat{\vartheta}_k := \frac{\sum_{i=1}^k n_i}{\sum_{i=1}^k L_i},$$

woraus sich ein Schätzer für die erwartete aktuelle Schadenzahl durch Multiplikation mit dem aktuellen Portfolioumfang ergibt. (Hierbei wird implizit vorausgesetzt, dass wenigstens ein  $n_i > 0$  ist, d.h. in der Vergangenheit wenigstens ein Schaden aufgetreten ist; ist dies nicht der Fall, so besitzt die Likelihoodfunktion keine Maximalstelle.)