

## Wichtige Werkzeuge zur Bestimmung der Verlustverteilung

Quelle:

Kreditrisikomessung – Statistische Grundlagen, Methoden und Modellierung  
Andreas Henking, Christian Bluhm, Ludwig Fahrmeir, 312 S., Springer Verlag, Berlin 2006.



## Wichtige Werkzeuge zur Bestimmung der Verlustverteilung

In diesem Kapitel beschreiben wir verschiedene prinzipielle Möglichkeiten, Verlustverteilungen zu bestimmen. Um eine plausible Verlustverteilung für ein Portfolio zu erhalten, reicht es jedoch in der Regel nicht aus, ein Verteilungsmodell zu bestimmen und mit den richtigen Parameterwerten zu befüllen. Vielmehr sind hierzu komplexere Modelle nötig, bei deren Formulierung die zuvor besprochenen Verteilungsmodelle mit ihren Eigenschaften von besonderer Bedeutung sind. Modelle dieser Art besprechen wir in Kapitel 6. Selbst wenn wir ein Portfoliomodell formuliert haben, müssen wir noch die Verlustverteilung und deren (Risiko-) Kennzahlen berechnen. Hierzu steht uns ein umfangreiches Instrumentarium zur Verfügung, dessen wichtigste Werkzeuge wir in diesem Kapitel einführend beschreiben.

### 4.1 Schätzen

Die Ausführungen in Abschnitt 2.4 zu den grundlegenden Begriffen der Statistik enthalten streng genommen nur Begriffe aus der Wahrscheinlichkeitsrechnung. Wir kennen nun Zufallsvariablen und wissen, dass Zufallsvariablen eine Verteilung haben. Dazu haben wir einige Verteilungsmodelle mit den zugehörigen Verteilungsparametern besprochen. Jedoch haben wir bisher nur vage angedeutet, wie man diese Verteilungen erhält. Und hier beginnt die Statistik erst richtig, nämlich mit dem *Schätzen* von Verteilungen oder deren Parametern. In den folgenden Unterabschnitten gehen wir auf die grundlegende Problematik des Schätzens und die Grundzüge verschiedener Schätzverfahren ein.

#### 4.1.1 Schätzproblem

Das Schätzproblem reduziert sich oft auf das Schätzen von Verteilungsparametern. Jedoch werden Verteilungen auch nicht-parametrisch durch Dichteschätzer und Histogramme geschätzt. Die Kreditrisikomodellierung geht in der Regel mit Modellannahmen einher, wie Abhängigkeit/Unabhängigkeit der Kreditnehmer und Verteilungsannahmen für die Anzahl der Ausfälle. Hiermit ist zunächst einmal ein *Mo-*

*dellrisiko* verbunden, das vor allem durch profunde Erfahrung der Modellbauer minimiert werden kann. So können z.B. beim Einsatz der Poisson- anstelle der Binomialverteilung andere Ausfallverteilungen resultieren, wie wir in den Beispielen zur Poisson-Verteilung in Abschnitt 3.2.3 exemplarisch gezeigt haben. Wir werden im Folgenden auf Modellrisiken nicht weiter eingehen. Die Modelle beinhalten aber (Verteilungs-) Parameter, wie Korrelationen und Ausfallwahrscheinlichkeiten, die geschätzt werden müssen bzw. können.

Gerade in der Kreditrisikomessung basieren Schätzungen auf Erfahrungen und Expertenwissen. Die Verteilungsparameter, z.B. die Ausfallwahrscheinlichkeit, werden oftmals von Experten- oder Entscheidungssystemen geliefert. Ratings sind ein Beispiel hierfür. Dies sind dann jedoch auch Schätzungen, deren Genauigkeit in der Praxis jedoch meist nicht ermittelt und damit bei den Portfoliomodellen nicht berücksichtigt wird.

Bei Schätzungen muss prinzipiell beachtet werden, dass sie immer mit Unsicherheit behaftet sind. Dies gilt umso mehr, wenn wir Aussagen über künftige Ereignisse treffen wollen. Die Unsicherheiten versuchen wir mit Eigenschaften der Schätzungen und mit statistischen Maßzahlen, wie der Varianz, zu quantifizieren.

#### 4.1.2 Schätzfunktionen und Schätzgenauigkeit

Für die Schätzung von Verteilungsparametern stehen uns *Schätzfunktionen* zur Verfügung, die auf der Basis von Beobachtungen, z.B. die Anzahl der Ausfälle je Ratingklasse in den vergangenen Jahren, einen Schätzwert liefern, also die eigentliche Schätzung. Es kann aber sein, dass die Beobachtungen aus einer Stichprobe stammen, zufällig keinen repräsentativen Querschnitt darstellen oder Ausreißer beinhalten. Daher rührt die Unsicherheit der Schätzungen, und Schätzfunktionen können somit als Zufallsvariablen aufgefasst werden. Die Eigenschaften dieser Zufallsvariablen nutzen wir, um die Qualität von Schätzfunktionen zu beurteilen.

Wir gehen im Folgenden davon aus, dass die Realisationen bzw. Beobachtungen  $x_1, x_2, \dots, x_n$ , die als Ergebnis einer Stichprobenziehung oder eines Zufallsexperiments vorliegen, durch die Zufallsvariablen  $X_1, X_2, \dots, X_n$  repräsentiert werden. Meist wird gefordert, dass die Realisationen unabhängige Wiederholungen einer Zufallsvariablen  $X$  sind, deren Verteilung bzw. Verteilungsparameter geschätzt werden soll. Man geht also davon aus dass die  $X_1, X_2, \dots, X_n$  unabhängig sind und alle die gleiche Verteilung haben. Diese Annahme wird uns bei den folgenden grundlegenden Überlegungen begleiten. Im Allgemeinen sind oft mehrere unbekannte Parameter zu schätzen. Wir beschränken uns aber im Folgenden auf den Fall *eines* Parameters. Für den mehrdimensionalen Fall verweisen wir auf [27].

Eine Schätzfunktion, oder kurz *Schätzer*, für einen zu schätzenden Parameter hat die allgemeine Form:

$$T = g(X_1, \dots, X_n). \quad (4.1)$$

$T$  ist also eine Funktion von Zufallsvariablen und somit selbst eine Zufallsvariable. Es kann sein, dass die Schätzfunktion  $g$  nicht explizit angegeben werden kann, was uns aber bei den folgenden Ausführungen nicht stören wird.

Werden die  $X_1, \dots, X_n$  als unabhängig angenommen, erhält man die gemeinsame Wahrscheinlichkeitsfunktion, Dichte oder Verteilungsfunktion durch Multiplikation der Wahrscheinlichkeitsfunktionen, Dichten oder Verteilungsfunktionen von  $X_1, \dots, X_n$ . Dies vereinfacht sich noch weiter, wenn die  $X_1, \dots, X_n$  auch die gleiche Verteilung haben. So bestimmt sich in diesem Fall die Verteilungsfunktion  $F_X$  von  $\mathbf{X} = (X_1, \dots, X_n)$  über:

$$F_X(x_1, \dots, x_n) = \prod_{i=1}^n F(x_i),$$

dabei bezeichnet  $F$  die Verteilungsfunktion von  $X_i, i = 1, \dots, n$ . Analog gilt für die Wahrscheinlichkeitsfunktion bzw. Dichte  $f_X$  von  $X$ :

$$f_X(x_1, \dots, x_n) = \prod_{i=1}^n f(x_i),$$

Aus der gemeinsamen Verteilung der  $X_1, \dots, X_n$  lässt sich dann prinzipiell, oft jedoch nicht in analytisch geschlossener Form die Verteilung von  $T = g(X_1, \dots, X_n)$  ableiten. Setzt man die Realisationen  $x_1, x_2, \dots, x_n$  ein, so erhält man den Schätzwert (Schätzung)

$$t = g(x_1, \dots, x_n).$$

Das bekannteste Beispiel für eine Schätzfunktion ist das arithmetische Mittel, über das Erwartungswerte geschätzt werden können.

Mit der Schätzfunktion wollen wir einen (Verteilungs-) Parameter  $\theta$  schätzen. Bei allgemeinen Ausführungen ist  $\theta$  also ein Platzhalter für einen bestimmten Parameter, wie die Ausfallwahrscheinlichkeit  $p$  oder den Erwartungswert  $\mu$ . Den *Schätzwert* des Parameters bezeichnen wir mit  $\hat{\theta}$ .

Eine wichtige und wünschenswerte Eigenschaft von Schätzfunktionen ist die *Erwartungstreue*. Das bedeutet, dass der Erwartungswert von  $T$  gleich dem zu schätzenden Parameter  $\theta$  ist:

$$\mathbb{E}(T) = \theta.$$

Oft ist dies nur für  $n \rightarrow \infty$  erfüllt, d.h. es gilt  $\mathbb{E}(T) \rightarrow \theta$ . In diesem Fall sprechen wir von *asymptotischer Erwartungstreue*. Ist ein Schätzer nicht erwartungstreu, so nennen wir den Schätzer *verzerrt*. Umgekehrt nennen wir erwartungstreue Schätzer auch *unverzerrt*. Die Erwartungstreue kann man in der Regel nachweisen, ohne den wahren Wert des Parameters  $\theta$  zu kennen. Wir werden dies in Abschnitt 4.2 beispielhaft zeigen.

Die *Verzerrung* eines Schätzers, also die Differenz zwischen dem Erwartungswert des Schätzers und dem zu schätzenden Parameter, nennt man auch *Bias*, den man wie folgt bestimmt:

$$Bias_{\theta}(T) = \mathbb{E}_{\theta}(T) - \theta.$$

Die Indizierung des Erwartungswertes mit  $\theta$  zeigt an, dass für  $T$  die Verteilung mit dem Parameter  $\theta$  zugrunde gelegt wird.

Eine weitere, oftmals geforderte Eigenschaft von Schätzfunktionen ist die der *schwachen Konsistenz*. Diese liegt vor, wenn für beliebige  $\epsilon > 0$

$$\lim_{n \rightarrow \infty} \mathbb{P}(|T - \theta| < \epsilon) = 1$$

gilt.

Eine etwas stärkere Forderung ist die der *Konsistenz im quadratischen Mittel*, auch MSE-Konsistenz. Sie ist dann erfüllt ist, wenn gilt:

$$\lim_{n \rightarrow \infty} \mathbb{E}(|T - \theta|^2) = 0.$$

Oft wird nur von *Konsistenz* gesprochen, womit in der Regel die Konsistenz im quadratischen Mittel gemeint ist. Die beiden Konsistenzdefinitionen besagen, dass für sehr große Stichprobenumfänge der Schätzwert in unmittelbarer Nähe des wahren Wertes des Parameters  $\theta$  liegen sollte. Dies gewährleistet aber nicht, dass der Schätzer auch für kleine Stichprobenumfänge genau ist.

Um die Güte des Schätzers auch für endliche und insbesondere kleine  $n$  zu untersuchen, ist die Standardabweichung der Schätzfunktion, der so genannte *Standardfehler*, eine wichtige Maßzahl:

$$\sigma_T = \sqrt{\mathbb{V}(T)}.$$

Je kleiner der Standardfehler, umso weniger streut die Verteilung der Schätzfunktion, umso genauer ist der Schätzer. Zudem ist es wünschenswert, dass der Standardfehler mit anwachsendem  $n$  kleiner wird und bereits für einen geringen Stichprobenumfang  $n$  möglichst klein ist. Falls der Standardfehler eines (asymptotisch) erwartungstreuen Schätzers für  $n \rightarrow \infty$  gegen 0 geht, so ist die Schätzfunktion *konsistent im quadratischen Mittel*. Stehen für den gleichen Parameter zwei erwartungstreue Schätzfunktionen zur Verfügung, so wählen wir sicherlich diejenige mit dem kleineren Standardfehler. Anstelle  $\sigma_T$  wird häufig SE (standard error) als Abkürzung für den Standardfehler verwendet.

Der Bias und der Standardfehler wird durch die *mittlere quadratische Abweichung* (*mean squared error*, MSE) simultan gemessen:

$$\begin{aligned} MSE &= \mathbb{E}([T - \theta]^2) \\ &= \text{Var}(T) + \text{Bias}(T)^2. \end{aligned}$$

Man sieht leicht, warum die Konsistenz im quadratischen Mittel auch MSE-Konsistenz heißt, und dass diese für asymptotisch erwartungstreue Schätzer erfüllt ist, deren Standardfehler gegen null geht.

### 4.1.3 Schätzverfahren

Wir erläutern kurz verschiedene Konstruktionsprinzipien für Schätzfunktionen.

**Maximum Likelihood-Schätzer:** Unter der *Likelihoodfunktion*  $L(\theta)$  versteht man die Wahrscheinlichkeitsfunktion oder Dichte der Stichprobe  $\mathbf{X} = (X_1, \dots, X_n)$  in Abhängigkeit des zu schätzenden Parameters  $\theta$ :

$$L(\theta) = f_X(x_1, \dots, x_n | \theta).$$

Das Maximum Likelihood-Prinzip besagt, dass zu den Realisationen  $x_1, \dots, x_n$  derjenige Parameter  $\hat{\theta}_{ML}$  als Schätzung gewählt wird, für den die Likelihood maximal ist:

$$L(\hat{\theta}_{ML}) = \max_{\theta} L(\theta).$$

Einen derart konstruierter Schätzer heißt *Maximum Likelihood-Schätzer*, der als eine Schätzfunktion  $T$  aufgefasst werden kann. Üblich ist auch die Kurzschreibweise *ML-Schätzer*. Man berechnet den Maximum Likelihood-Schätzer entweder auf analytischem Wege oder mit Hilfe numerischer Verfahren.

**Kleinste-Quadrate-Schätzer:** Die *Kleinste-Quadrate-Schätzer* (KQ-Schätzer) sind jene Schätzer  $\hat{\theta}_{KQ}$ , welche die quadrierten Abweichungen zu den Realisationen minimieren. Zur Konstruktion des Schätzers wird  $\hat{\theta}_{KQ}$  so bestimmt, dass

$$\sum_{i=1}^n (x_i - \hat{\theta}_{KQ})^2 \rightarrow \min$$

erfüllt ist. Damit erhalten größere Abweichungen von  $\hat{\theta}_{KQ}$  und insbesondere Ausreißer ein überproportional höheres Gewicht bei der Ermittlung des Schätzers. Der KQ-Schätzer findet insbesondere in der *Regressionsanalyse* Anwendung. Auch wenn das ML-Prinzip einen anderen Zugang wählt, so stimmen ML- und KQ-Schätzer oftmals überein.

#### 4.1.4 Konfidenzintervalle

Um Informationen für die Genauigkeit einer Schätzung mitzuliefern, werden oftmals Konfidenzintervalle gebildet. Dies sind Intervalle um den Schätzwert, die mit einer vorgegebenen (kleinen) Wahrscheinlichkeit  $\alpha$  den gesuchten Parameter *nicht* enthalten. Die Wahrscheinlichkeit  $1 - \alpha$  heißt *Überdeckungswahrscheinlichkeit* oder *Konfidenzniveau*.

Bei der Konstruktion von Konfidenzintervallen bildet man aus den Stichprobenvariablen eine Untergrenze  $G_u = g_u(X_1, \dots, X_n)$  und eine Obergrenze  $G_o = g_o(X_1, \dots, X_n)$ , so dass gilt:

$$\mathbb{P}(G_u \leq \theta \leq G_o) = 1 - \alpha.$$

Es ist zu beachten, dass Zufallsvariablen in obiger Wahrscheinlichkeit in der Unter- und Obergrenze und nicht in  $\theta$  stecken. Setzt man in  $G_u$  und  $G_o$  die Realisationen

$x_1, \dots, x_n$  ein, so ergeben sich die realisierten Grenzen  $g_u = g_u(x_1, \dots, x_n)$  und  $g_o = g_o(x_1, \dots, x_n)$  und  $[g_u, g_o]$  bildet das Konfidenzintervall.

Typischerweise benötigt man zur Konstruktion von Konfidenzintervallen einen Schätzer  $\hat{\theta}$  und die (approximative) Verteilung dieses Schätzers, insbesondere die  $\alpha/2$  und  $(1 - \alpha/2)$  Quantile dieser Verteilung.

#### 4.1.5 Empirische Verteilungsfunktion

Ein besondere Bedeutung kommt der empirischen Verteilungsfunktion zu, die direkt aus den Realisationen  $x_1, \dots, x_n$  gebildet wird und die Verteilungsfunktion „auf“ den Daten ist. Damit hat die empirische Verteilungsfunktion alle Eigenschaften einer diskreten Verteilungsfunktion, wie in Abschnitt 2.4.3 beschrieben. Besondere Bedeutung erlangt die empirische Verteilungsfunktion bei der Verarbeitung der Ergebnisse einer Monte-Carlo-Simulation, die in Abschnitt 4.6 beschrieben ist.

Seien  $x_{(1)}, \dots, x_{(n)}$  die geordneten Realisationen. Dabei ist die Umklammerung der Indices eine übliche Schreibweise für geordnete Werte. Es gilt somit:

$$x_{(i)} \leq x_{(j)} \quad \text{für} \quad 1 \leq i < j \leq n.$$

Die absoluten kumulierten Häufigkeiten sind:

$$H(x) = \max_j \{j : x_{(j)} \leq x\} = \text{Anzahl der Werte } x_i \text{ mit } x_i \leq x$$

Die *empirische Verteilungsfunktion*  $F_n(x)$ , als Analogon zur diskreten Verteilungsfunktion, bildet sich aus den relativen kumulierten Häufigkeiten:

$$F_n(x) = \frac{H(x)}{n} = \text{Anteil der Werte } x_i \text{ mit } x_i \leq x.$$

Da die empirische Verteilungsfunktion alle Eigenschaften einer diskreten Verteilungsfunktion besitzt, entspricht auch ihre grafische Darstellung der einer diskreten Verteilungsfunktion (vgl. Abbildung 2.10 auf Seite 53).

## 4.2 Schätzung von Risikokennzahlen

### 4.2.1 Ausfallwahrscheinlichkeit

Die Ausfallwahrscheinlichkeit, so wie Wahrscheinlichkeiten allgemein, wird normalerweise über die relative Häufigkeit geschätzt. Seien dazu

$$\mathbf{1}_{D_i}, \quad i = 1, \dots, n$$

die *unabhängigen* Ausfallindikatoren der  $n$  Kreditnehmer eines Portfolios. Die Ausfallindikatoren sind Bernoulli-verteilt (vgl. Abschnitt 3.2.1):

$$\mathbf{1}_{D_i} \sim \mathcal{B}(1, p), \quad i = 1, \dots, n.$$

Dann ist

$$\hat{P} = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{D_i}$$

ein Schätzer für die Ausfallwahrscheinlichkeit  $p$ , der nichts anders als die relative Häufigkeit, also der Anteil, der zufälligen Ausfälle ist. Man nennt  $\hat{P}$  auch *Ausfallquote*. Setzen wir die beobachteten Ausfälle  $\mathbf{1}_{d_i}$  ein, so erhalten wir den Schätzwert

$$\hat{p} = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{d_i}$$

Um die Eigenschaften des Schätzers  $\hat{P}$  genauer zu untersuchen, verwenden wir, dass die Summe unabhängiger Bernoulli-verteilter Zufallsvariablen binomialverteilt ist (vgl. Formel (3.22)):

$$\sum_{i=1}^n \mathbf{1}_{D_i} \sim \mathcal{B}(n, p).$$

Somit kennen wir die Momente von  $n\hat{P} = \sum_{i=1}^n \mathbf{1}_{D_i}$ :

$$\mathbb{E} \left( \sum_{i=1}^n \mathbf{1}_{D_i} \right) = np$$

$$\mathbb{V} \left( \sum_{i=1}^n \mathbf{1}_{D_i} \right) = np(1-p)$$

Also ist (vgl. Gleichung (3.4)):

$$\mathbb{E}(\hat{P}) = \frac{1}{n} \mathbb{E} \left( \sum_{i=1}^n \mathbf{1}_{D_i} \right) = \frac{1}{n} np = p$$

und daher der Schätzer  $\hat{P}$  erwartungstreu.

Für die Varianz von  $\hat{P}$  gilt mit Gleichung (3.8)

$$\mathbb{V}(\hat{P}) = \frac{1}{n^2} \mathbb{V} \left( \sum_{i=1}^n \mathbf{1}_{D_i} \right) = \frac{1}{n^2} np(1-p) = \frac{p(1-p)}{n}.$$

Wir sehen, dass für  $n \rightarrow \infty$  die Varianz von  $\hat{P}$  verschwindet. Deshalb und wegen der Erwartungstreue ist  $\hat{P}$  ein konsistenter Schätzer.

Lässt man die Unabhängigkeitsannahme der Ausfallindikatoren  $\mathbf{1}_{D_i}$  fallen und geht in einem einfachen Portfoliomodell (vgl. Ein-Faktor-Modell in Abschnitt 6.1) von positiver Assetkorrelation aus, so ist  $\hat{P}$  immer noch erwartungstreu aber nicht mehr zwingend konsistent (vgl. Höse, Huschens[48]).



### 4.2.2 Erwartungswert / Erwarteter Verlust

Seien  $L_1, \dots, L_n$  unabhängige Zufallsvariablen mit der gleichen Verteilung wie der Portfolioverlust  $L_{PF}$ . Dann ist

$$\bar{L}_{PF} = \frac{1}{n} \sum_{i=1}^n L_i$$

ein Schätzer für den erwarteten Verlust bzw. für den Erwartungswert von  $L_{PF}$ .

Der Erwartungswert wird oft als *Mittelwert* bezeichnet, was aber streng genommen falsch ist. Der Begriff wird deswegen verwendet, da der üblichste Schätzer für den Erwartungswert das *arithmetische Mittel* von Beobachtungen einer Zufallsvariablen ist. Seien  $l_1, \dots, l_n$  Realisationen von  $L_1, \dots, L_n$ . So ist das arithmetische Mittel  $\bar{l}_{PF}$  der Verluste der Schätzwert für den erwarteten Verlust  $\mathbb{E}(L_{PF})$ :

$$\bar{l}_{PF} = \frac{1}{n} \sum_{i=1}^n l_i. \quad (4.2)$$

Die Zufallsvariable  $L_{PF}$ , damit auch die Zufallsvariablen  $L_1, \dots, L_n$ , haben den Erwartungswert  $\mu_{PF}$  und die Varianz  $\sigma_{PF}^2$ . Dann ist wegen der Linearität des Erwartungswertes (vgl. Formel (3.5))

$$\mathbb{E}(\bar{L}_{PF}) = \frac{1}{n} \sum_{i=1}^n \mathbb{E}(L_i) = \frac{1}{n} \sum_{i=1}^n \mu_{PF} = \mu_{PF}$$

und somit ist  $\bar{L}_{PF}$  erwartungstreu.

Die Varianz von  $\bar{L}_{PF}$  ist mit Gleichung (3.10):

$$\mathbb{V}(\bar{L}_{PF}) = \frac{1}{n^2} \sum_{i=1}^n \mathbb{V}(L_i) = \frac{1}{n^2} \sum_{i=1}^n \sigma_{PF}^2 = \frac{1}{n} \sigma_{PF}^2.$$

Somit ergibt sich für den Standardfehler von  $\bar{L}_{PF}$ :

$$\sigma_{\bar{L}_{PF}} = \frac{1}{\sqrt{n}} \sigma_{PF}. \quad (4.3)$$

Wir sehen, dass  $\bar{L}_{PF}$  konsistent ist, da der Schätzer erwartungstreu und der Standardfehler für  $n \rightarrow \infty$  gegen 0 geht.

Wenn  $\mu_{PF}$  geschätzt werden muss, dann ist in der Regel auch  $\sigma_{PF}^2$  unbekannt. Daher setzt man zur Abschätzung des Standardfehlers die geschätzte Varianz bzw. Standardabweichung in (4.3) ein. Schätzfunktionen hierzu besprechen wir weiter unten. Es scheint kaum möglich zu sein, genügend Realisationen des Portfolioverlustes  $L_{PF}$  zu beobachten (Zumindest wenn man fordert, dass  $n$  nicht zu klein ist, damit gewisse oder vorgegebene Genauigkeitsanforderungen an den Schätzer  $\bar{L}_{PF}$  erfüllt sind). Jedoch erzeugt man beispielsweise in Monte-Carlo-Simulationen beliebig viele Realisationen von  $L_{PF}$ , wo dann der Schätzer seinen berechtigten Einsatz findet.

### 4.2.3 Varianz / Unerwarteter Verlust

Die Varianz  $\sigma_{PF}^2$  von  $L_{PF}$  schätzt man aus  $L_1, \dots, L_n$  über

$$S^2 := \hat{\sigma}_{PF}^2 = \frac{1}{n-1} \sum_{i=1}^n (L_i - \bar{L}_{PF})^2.$$

Mit ein paar Umformungen zeigt man, dass  $\hat{\sigma}_{PF}^2$  erwartungstreu ist, also

$$\mathbb{E}(\hat{\sigma}_{PF}^2) = \sigma_{PF}^2.$$

Dividiert man bei dem Schätzer  $\hat{\sigma}_{PF}^2$  durch  $n$  anstatt durch  $(n-1)$ , so ist der resultierende Schätzer verzerrt:

$$\mathbb{E}\left(\frac{1}{n} \sum_{i=1}^n (L_i - \bar{L}_{PF})^2\right) = \mathbb{E}\left(\frac{n-1}{n} \hat{\sigma}_{PF}^2\right) = \frac{n-1}{n} \sigma_{PF}^2.$$

Der Verzerrungsfaktor  $\frac{n-1}{n}$  geht jedoch für  $n \rightarrow \infty$  gegen 0, so dass  $\frac{n-1}{n} \hat{\sigma}_{PF}^2$  *asymptotisch erwartungstreu* ist.

Ein Schätzer für die Standardabweichung von  $L_{PF}$  und somit für den unerwarteten Verlust ist die Wurzel aus dem Varianzschätzer:

$$\hat{\sigma}_{PF} = \sqrt{\hat{\sigma}_{PF}^2}.$$

Jedoch ist dieser Schätzer im Allgemeinen nicht erwartungstreu, was darin begründet ist, dass die Wurzelfunktion nicht linear ist.

Sind  $L_1, \dots, L_n$  unabhängig und identisch  $\mathcal{N}(\mu, \sigma^2)$ -verteilt, so ist die Varianz von  $\hat{\sigma}_{PF}^2$ :

$$\mathbb{V}(\hat{\sigma}_{PF}^2) = \frac{2}{n-1} \sigma^4.$$

Der Standardfehler ist wie gehabt die Wurzel daraus. Somit ist auch  $\hat{\sigma}_{PF}^2$  ein konsistenter Schätzer.

Man kann auch voraussetzen, dass der Erwartungswert  $\mu_{PF}$  von  $L_{PF}$  bekannt ist. Der Varianzschätzer ist dann

$$\hat{\sigma}_{PF,\mu}^2 = \frac{1}{n} \sum_{i=1}^n (L_i - \mu_{PF})^2.$$

Dieser Schätzer ist ebenfalls erwartungstreu und konsistent. Er hat mit  $\sqrt{\frac{2}{n} \sigma^4}$  einen geringeren Standardfehler als  $\hat{\sigma}_{PF}^2$ , da durch den bekannten Erwartungswert zusätzliche Informationen in die Schätzung einfließen.

### 4.3 Momentenmethode

Bei der Momentenmethode werden die Momente einer interessierenden Variablen, die ohne komplexes Modell und ohne Verteilungsannahmen zur Verfügung stehen, ermittelt. Im Kontext des Portfolioverlustes sind dies typischerweise Erwartungswert und Varianz. Zur Bestimmung der Varianz muss eventuell vereinfachend von der Unabhängigkeit der Kreditnehmer ausgegangen werden. Diese Momente werden dann als Parameter einer Modellverteilung eingesetzt. Unter Umständen ist hierzu noch eine Transformation der Momente in die Verteilungsparameter notwendig, wie wir dies am Beispiel der Betaverteilung vorführen. Die Anwendung der Momentenmethode setzt allerdings voraus, dass die gesuchte Verteilung ausreichend gut durch die Momente charakterisiert werden kann. Inwieweit dies erfüllt ist, lässt sich in der Praxis nicht immer befriedigend abschätzen. Somit geht mit der Momentenmethode ein starkes *Modellrisiko* einher.

Die Momentenmethode hat z.B. im Zusammenhang mit der Betaverteilung durchaus praktische Relevanz. Nach der Anpassung der Momente liegt die Verteilung in geschlossener Form vor und Berechnungen sind in Sekundenschnelle möglich. Wegen des hohen Modellrisikos eignet sich die Betaverteilung allerdings nicht, um die Verlustrisiken möglichst exakt abzuschätzen. Jedoch wird die Betaverteilung oft als Vergleichs- oder Referenzverteilung (z.B. in [82]) verwendet. Vor allem bei der Betrachtung extremer Verluste, wie den Value at Risk, eignet sich die Momentenmethode nicht sonderlich gut, da bei Verwendung eines Lageparameters (Erwartungswert) und eines Streuungsparameters (Varianz) die Verteilung in erster Linie in der „Mitte“ angenähert wird und nicht an den Enden.

Wir erklären die Momentenmethode exemplarisch anhand von Portfolio II.

#### Portfolio II

Wir wollen die Verlustverteilung von Portfolio II durch eine Betaverteilung modellieren. Dazu ist der Übergang von diskreten zu stetigen Verlusten notwendig, da es sich bei der Betaverteilung um ein stetiges Verteilungsmodell handelt. Aufgrund der geringen Anzahl von Kreditnehmern in Portfolio I würde sich dieses keines Falls eignen, um den Übergang von einer diskreten zu einer stetigen Verlustverteilung vorzunehmen. Selbst die 80 Kreditnehmer in Portfolio II scheinen noch nicht die Näherung einer diskreten Verteilung durch die Betaverteilung zu rechtfertigen. Dennoch führen wir den Übergang von der Binomialverteilung zur Betaverteilung anhand des Portfolios II vor. Nach unserer Erfahrung wird in der Praxis auf die Güte einer solchen Näherung oftmals wenig Wert gelegt. Umso wichtiger erscheint es uns, diese Näherung durchzuführen und eine erste Diskussion der Güte dieser Näherung anzuregen.

Wir wollen die binomialverteilte Zufallsvariable

$X$  : Anzahl der Ausfälle in Portfolio II

durch eine betaverteilte Zufallsvariable annähern. Hierbei legen wir wieder die mittlere 5-jährige Ausfallwahrscheinlichkeit zugrunde und nehmen dem Beispiel zu Portfolio II auf Seite 83 folgend  $X \sim \mathcal{B}(80; 0, 0195)$  an.

Da die Beta-Verteilung, so wie wir sie in Abschnitt 3.2.7 eingeführt haben, auf das Einheitsintervall beschränkt ist, betrachten wir die Zufallsvariable

$Y$  : Prozentualer Anteil der Ausfälle in Portfolio II.

Wir sehen leicht, dass der Zusammenhang

$$Y = \frac{1}{n}X$$

gilt, wobei  $n = 80$  die Anzahl der Kreditnehmer in Portfolio II ist.

Über den Erwartungswert und die Varianz von  $X$  bestimmen wir nun Erwartungswert und Varianz von  $Y$ . Dabei wenden wir die Formeln (3.24) und (3.25) für Erwartungswert und Varianz der Binomialverteilung und die Rechenregeln (3.4) und (3.8) an.

$$\mathbb{E}(Y) = \frac{1}{n}\mathbb{E}(X) = p = 0,0195$$

$$\mathbb{V}(Y) = \frac{1}{n^2}\mathbb{V}(X) = \frac{p(1-p)}{n} = 0,000239.$$

Erwartungswert und Varianz von  $Y$  setzen wir nun in die Formeln (3.36) und (3.37) zur Bestimmung der beiden Parameter der Beta-Verteilung ein:

$$\alpha = 1,54$$

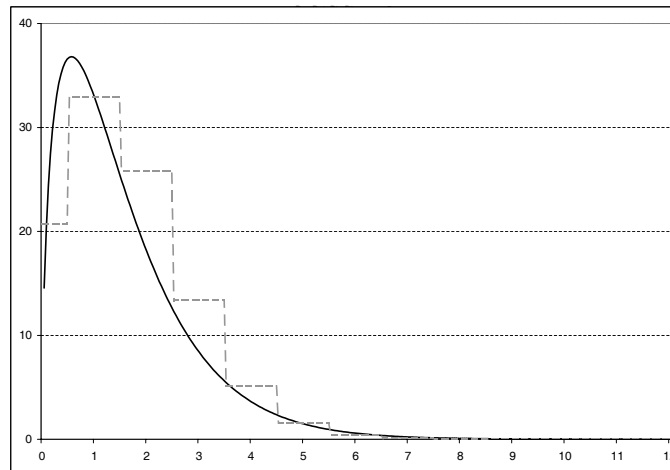
$$\beta = 77,46.$$

Damit ist die Beta-Verteilung ausreichend spezifiziert um die  $\mathcal{B}(80; 0,0195)$ -Verteilung über die Momentenmethode zu approximieren.

In Abbildung 4.1 ist die Dichte der Beta-Verteilung und die Wahrscheinlichkeitsfunktion der Binomialverteilung zu sehen. Dazu haben wir den Träger mit 80 multipliziert, was einer Rücktransformation von  $Y$  nach  $X$  entspricht, also der Darstellung der Anzahl der Ausfälle anstelle deren Anteil. Aus Übersichtlichkeitsgründen ist wie schon bei vorhergehenden Beispielen zu Portfolio II in der Darstellung der Träger bei 12 abgeschnitten. Wir erkennen einen ähnlichen Verlauf der beiden Verteilungen, der jedoch deutliche Unterschiede zwischen den beiden Verteilungen nicht verdecken kann. Offenbar kann die diskrete Binomialverteilung noch nicht durch die stetige Beta-Verteilung approximiert werden. In Abschnitt 3.2.2 hatten wir gesehen, dass für große  $n$  die Binomialverteilung ihren diskreten Charakter mehr und mehr verliert. Aufgrund des kleinen  $p$  im Beispiel sind jedoch nur die möglichen Realisationen  $\leq 7$  relevant, in dem Sinne, dass in diesem Bereich die Verteilung gut approximiert werden muss. Offenbar ist es besser zu fordern, dass  $np$  groß ist, um die Binomialverteilung durch eine stetige Verteilung zu approximieren.

## 4.4 Grenzwertsätze

Viele Beweise in der Stochastik und viele Anwendungen der Statistik verwenden Erkenntnisse aus Grenzwertsätzen, um den Einsatz bestimmter Verteilungsmodelle,



**Abbildung 4.1.** Dichtefunktion der Beta-Verteilung für die Anzahl der Ausfälle im Portfolio II (zum Vergleich ist die Wahrscheinlichkeitsfunktion der korrespondierenden Binomialverteilung im Portfolio II angegeben)

insbesondere oftmals der Normalverteilung, zumindest näherungsweise zu rechtfertigen. An Beispielen haben wir in Abschnitt 3.2.3 die Annäherung der Binomialverteilung durch die Poisson-Verteilung untersucht und festgestellt, dass für kleine Werte von  $p$  diese Näherung gerechtfertigt werden kann.

Wir stellen im Folgenden als Repräsentant der Grenzwertsätze den zentralen Grenzwertsatz vor. Durch Anwendung von Grenzwertsätzen erhalten wir Faustregeln, wann und wie eine bestimmte Verteilung durch eine andere angenähert werden kann.

#### 4.4.1 Zentraler Grenzwertsatz

Bereits der Name *zentraler Grenzwertsatz* macht dessen herausragende Bedeutung für die Stochastik und Statistik deutlich. Der Grenzwertsatz ist zentral, da er die Normalverteilung als Grenzverteilung in vielen Anwendungen rechtfertigt. Allerdings setzt er in der nachfolgend vorgestellten Version Unabhängigkeit und die gleiche Verteilung der involvierten Zufallsvariablen voraus, was in der Anwendung nicht immer gegeben ist. Es gibt jedoch auch allgemeinere Versionen des zentralen Grenzwertsatzes, in denen diese Annahmen gelockert werden.

**Zentraler Grenzwertsatz:**  $X_1, \dots, X_n$  seien unabhängig identisch verteilte Zufallsvariablen mit

$$\mathbb{E}(X_i) = \mu \quad \text{und} \quad \mathbb{V}(X_i) = \sigma^2 > 0.$$

Dann konvergiert die Verteilungsfunktion  $F_n(z) = \mathbb{P}(Z_n \leq z)$  der standardisierten Summe

$$Z_n = \frac{X_1 + \cdots + X_n - n\mu}{\sqrt{n}\sigma} = \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{X_i - \mu}{\sigma}$$

für  $n \rightarrow \infty$  an jeder Stelle  $z \in \mathbb{R}$  gegen die Verteilungsfunktion  $\Phi(z)$  der Standardnormalverteilung:

$$F_n(z) \rightarrow \Phi(z).$$

Wir schreiben dafür kurz

$$Z_n \stackrel{a}{\sim} \mathcal{N}(0, 1). \quad (4.4)$$

Die Schreibweise  $\stackrel{a}{\sim}$  steht für *approximativ* (näherungsweise für große  $n$ ) oder *asymptotisch* (für  $n \rightarrow \infty$ ) verteilt. Dies besagt also, dass standardisierte Summen unabhängiger und identisch verteilter Zufallsvariablen als näherungsweise normalverteilt betrachtet werden können. Da  $X_1, \dots, X_n$  alle die gleiche Verteilung haben, besitzen sie insbesondere auch alle den gleichen Erwartungswert und die gleiche Varianz. Implizit wird in obiger Formulierung des zentralen Grenzwertsatzes vorausgesetzt, dass diese beiden Momente existieren, also endlich sind. Mit den Formeln (3.1) und (3.9) lassen sich leicht Erwartungswert und Varianz der Zufallsvariablen  $Y_n = X_1 + \cdots + X_n$  bestimmen:

$$\mathbb{E}(Y_n) = n\mu$$

$$\mathbb{V}(Y_n) = n\sigma^2$$

Somit wird deutlich, dass  $Z_n$  die standardisierte Version von  $Y_n$  ist. Machen wir die Standardisierung rückgängig, so erhalten wir näherungsweise die Verteilung von  $Y_n$ :

$$Y_n = X_1 + \cdots + X_n \stackrel{a}{\sim} \mathcal{N}(n\mu, n\sigma^2).$$

Die *Approximationsgeschwindigkeit*, also die Güte der Näherung in Abhängigkeit von  $n$ , kann recht gut sein, hängt jedoch von der Verteilung der  $X_i$  ab. Wir führen dies exemplarisch und nur grafisch am Beispiel der Binomialverteilung vor. Gemäß Formeln (3.21) und (3.22) sind binomialverteilte Zufallsvariablen die Summe unabhängiger und identisch Bernoulli-verteilter Zufallsvariablen mit Erwartungswert  $p$  und Varianz  $p(1-p)$ . Somit sind die Voraussetzungen des zentralen Grenzwertsatzes erfüllt und die Binomialverteilung lässt sich durch die Normalverteilung approximieren. Dieser Zusammenhang wird auch Grenzwertsatz von de Moivre genannt, den wir wegen seiner Bedeutung explizit formulieren.

**Grenzwertsatz von de Moivre:**  $X_1, \dots, X_n$  seien unabhängig identisch bernoulli-verteilte Zufallsvariablen mit

$$X_i \sim \mathcal{B}(1, p).$$

$H_n = X_1 + \cdots + X_n$  nennen wir *absolute Häufigkeit*. Für  $n \rightarrow \infty$  konvergiert die Verteilung der standardisierten absoluten Häufigkeiten

$$\frac{H_n - np}{\sqrt{np(1-p)}}$$

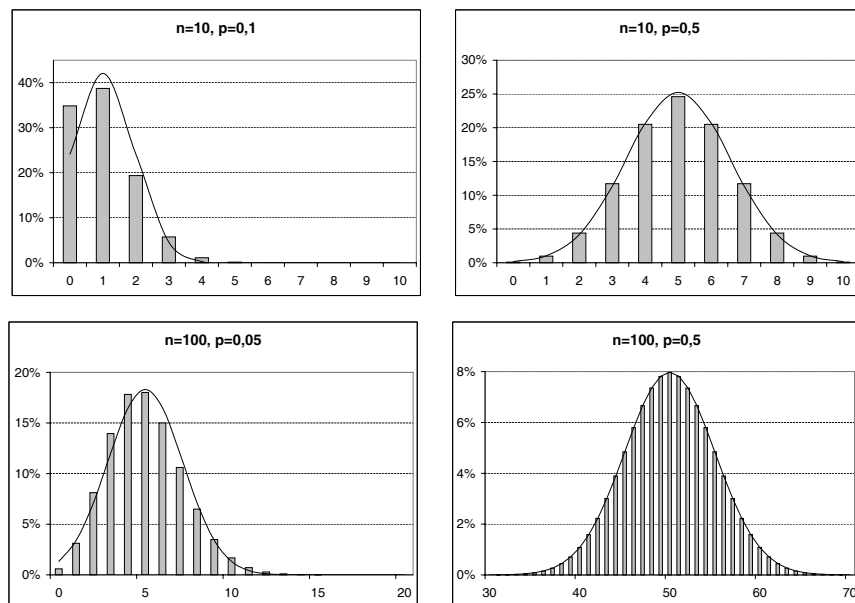
gegen die Standardnormalverteilung. Für großes  $n$  gilt

$$H_n \stackrel{a}{\sim} \mathcal{N}(np, np(1-p)).$$

Für die *relative Häufigkeit*  $\frac{H_n}{n}$  gilt entsprechend

$$\frac{H_n}{n} \stackrel{a}{\sim} \mathcal{N}\left(p, \frac{p(1-p)}{n}\right).$$

Dieses Ergebnis verwenden wir nun, um die Beispiele für die Binomialverteilung in den Abbildungen 3.3 und 3.4 mit der Normalverteilung anzunähern (vgl. Abbildung 4.2). Für die Verteilungen mit  $n = 100$  ist gegenüber Abbildung 3.4 der Wertebereich aus Übersichtlichkeitsgründen eingeschränkt.



**Abbildung 4.2.** Die Wahrscheinlichkeitsfunktion der Binomialverteilung für verschiedene Parameter und die angepasste Dichte der zugehörigen näherungsweise Normalverteilung

In Abbildung 4.2 ist zu erkennen, dass für größere  $n$  und für  $p$  näher an 0,5 die Annäherung der Binomialverteilung durch die Normalverteilung besser wird. Die bessere Anpassungsgüte bei größeren  $n$  wird uns durch den Grenzwertsatz von de Moivre verdeutlicht. Die bessere Anpassungsgüte für  $p$  nahe 0,5 ist durch die Symmetrie der Normalverteilung um  $\mu$  und der Binomialverteilung um den Erwartungswert für

$p = 0,5$  begründbar. Anders ausgedrückt: für kleine Werte von  $p$  nahe der 0 und große Werte von  $p$  nahe der 1 ist die Wahrscheinlichkeitsfunktion der Binomialverteilung schief, was der Symmetrie der Normalverteilung widerspricht. Diese Schiefe kann für große  $n$  weitgehend egalisiert werden. Es lässt sich feststellen, dass die Anpassungsgüte für  $n = 10$  und  $p = 0,5$  sowie für die beiden Beispiele mit  $n = 100$  wohl schon recht gut ist, bis auf Ungenauigkeiten in der Anpassung für  $p = 0,05$ . Jedoch sollte bei Anwendung der Näherung überprüft werden, ob die erforderliche Anpassungsgüte erreicht wird.

Abbildung 4.2 belegt auch, dass der Übergang von einer diskreten zu einer stetigen Verteilung nicht ganz unproblematisch ist. Diesem Problem kann man mit der *Stetigkeitskorrektur* entgegen wirken, die zum Beispiel in [29] beschrieben ist.

#### 4.4.2 Faustregeln für die Approximation von Verteilungen

Wir geben einen Überblick über Faustregeln (vgl. [29]), ab wann die Näherungen von ausgewählten Verteilungen durch eine andere in der Regel ausreichend ist. Jedoch müssen diese Näherungen in der jeweiligen Anwendung auf den Prüfstand gestellt werden, da zum Beispiel die Forderung nach einer besonders guten Anpassung für Quantile zu Wahrscheinlichkeiten nahe 1 nicht erfüllt sein muss, selbst wenn die Näherung im Zentrum der Verteilung sehr gut erscheint.

Verteilung	Näherung	Faustregel	Parameter
Binomialverteilung $\mathcal{B}(n, p)$	Poissonverteilung $\mathcal{P}(\lambda)$	$n > 30$ $p \leq 0,05$	$\lambda = np$
Binomialverteilung $\mathcal{B}(n, p)$	Normalverteilung $\mathcal{N}(\mu, \sigma^2)$	$np \geq 5$ $n(1-p) \geq 5$	$\mu = np$ $\sigma^2 = np(1-p)$
Poissonverteilung $\mathcal{P}(\lambda)$	Normalverteilung $\mathcal{N}(\mu, \sigma^2)$	$\lambda \geq 10$	$\mu = \lambda$ $\sigma^2 = \lambda$
t-Verteilung $t(n)$	Normalverteilung $\mathcal{N}(0, 1)$	$n \geq 30$	$\mu = 0$ $\sigma^2 = 1$

### 4.5 Analytische Bestimmung der Verlustverteilung

Der Portfolioverlust  $L_{PF}$  sei durch eine Modellgleichung oder eine Folge von Modellgleichungen beschreibbar:

$$L_{PF} = g(X_1, \dots, X_p, y_1, \dots, y_q). \quad (4.5)$$

Dabei seien  $X_1, \dots, X_p$  stochastische und  $y_1, \dots, y_q$  deterministische oder feste Einflussgrößen. Die Verteilungen und stochastischen Abhängigkeitsstrukturen der  $X_1, \dots, X_p$  sowie die Werte der  $y_1, \dots, y_q$  seien bekannt. Ebenfalls bekannt sei die Funktion  $g()$ , über die sich aus den stochastischen und nicht stochastischen Einflussgrößen der Portfolioverlust bestimmt. Da der Portfolioverlust von stochastischen Komponenten abhängt, ist er selber stochastisch. Ideal ist es also, wenn man über die



Modellgleichung (4.5) die Verteilung des Portfolioverlustes bestimmen kann. Allerdings ist dies oft nur unter vereinfachenden Annahmen möglich. Daher gibt es unter den anspruchsvolleren Portfoliomodellen nur wenige, bei denen dies gelingt. Das bekannteste Beispiel hierfür ist das Portfoliomodell CreditRisk<sup>+</sup>, welches wir in Abschnitt 6.2 besprechen. Aber auch hier sind die resultierende Verlustverteilung und die zugehörige Rechenvorschrift alles andere als trivial.

Den Ansatz zum CreditRisk<sup>+</sup>-Modell haben wir bereits in dem Abschnitt 3.2.5 zur Gammaverteilung beschrieben. Dort geht aus der Mischung der Poisson- mit der Gammaverteilung die negative Binomialverteilung hervor. Wir beschreiben ebenfalls in Abschnitt 3.2.5, dass die  $\chi^2$ -Verteilung ein Spezialfall der Gammaverteilung ist. Die  $\chi^2$ -Verteilung ist aber auch die Verteilung der Summe unabhängiger, quadrierter standardnormalverteilter Zufallsvariablen:

$$X_1^2 + \dots + X_n^2 \sim \chi^2(n) \quad \text{falls} \quad X_1, \dots, X_n \stackrel{iid}{\sim} \mathcal{N}(0, 1).$$

Dies hat zwar keinen direkten Bezug zum Kreditrisiko, zeigt aber, dass es sehr wohl möglich ist, analytisch die Verteilung zusammengesetzter Zufallsvariablen, nichts anderes ist der Portfolioverlust, zu bestimmen. Es zeigt aber auch, dass es bereits in den beiden einfachen Beispielen, Mischung von Verteilungen und Summe von Zufallsvariablen, recht schnell kompliziert wird und es einige Übung und Erfahrung erfordert, solche Umformungen und Herleitungen vorzunehmen. Dies überlassen wir dann auch denjenigen, die hauptberuflich mit solchen Dingen beschäftigt sind. Wir sollten zumindest ansatzweise verstehen, wie solche Umformungen funktionieren.

Ein wichtiges Werkzeug bei der Ermittlung von Verteilungen zusammengesetzter Zufallsvariablen ist die *Faltung* (engl: convolution). Über die Faltung von Verteilungen kann die Verteilung einer Zufallsvariablen hergeleitet werden, die als Summe stochastisch unabhängiger Zufallsvariablen darstellbar ist. Das bedeutet, dass wir für stochastisch unabhängige  $X_1, \dots, X_n$  über Faltung die Verteilung von

$$Y = a_0 + a_1 X_1 + \dots + a_n X_n$$

bestimmen können. Was bedeutet das für unsere Fragestellung, Verlustverteilungen zu bestimmen? Wenn die Ausfallindikatoren  $D_1, \dots, D_m$  der Kreditnehmer eines Portfolios als stochastisch unabhängig angenommen werden und wir

$$a_i = \text{EAD}_i \cdot \text{LGD}_i, \quad i = 1, \dots, m$$

$$L_i = a_i D_i$$

setzen, dann ist unser Portfolioverlust als

$$L_{PF} = L_1 + \dots + L_m$$

darstellbar. Die Verteilungen der Verluste  $L_i$  gehen direkt aus den Verteilungen der Ausfälle  $D_i$  hervor, da die  $a_i$  als deterministisch angenommen werden. Die Verteilung von  $L_{PF}$  können wir dann über die Faltung der Verteilungen von  $L_1, \dots, L_m$  bestimmen. Die Verteilung von  $L_{PF}$  unter diesen Annahmen zu bestimmen ist auch gar nicht abwegig, da wir somit eine Referenzverlustverteilung bestimmen können.

Genau dies haben wir im übrigen in Abschnitt 2.3.14 getan. Dort haben wir die Verlustverteilung von Portfolio II, unter der Annahme der Unabhängigkeit, in Tabelle 2.8 angegeben

Auch im Modell CreditRisk<sup>+</sup> wird die Faltung von Verteilungen verwendet. Dies ist möglich, da die Sektoren, über die die Abhängigkeiten zwischen den Kreditnehmern gesteuert werden, als stochastisch unabhängig angenommen werden. Für weitere Ausführungen verweisen wir auf Abschnitt 6.2 und dort auf den Unterabschnitt 6.2.3 unter „Berechnung der Verlustverteilung“ sowie auf die Referenzen in Abschnitt 6.2.

Für die Faltung von Verteilungen existieren Formeln, nach denen man strikt vorgehen kann. Wir geben die Faltungsformel für zwei unabhängige stetige Zufallsvariablen  $X_1$  und  $X_2$  an, deren Dichten  $f_{X_1}$  und  $f_{X_2}$  existieren und bekannt sind. Dann ist die Dichte von  $X_1 + X_2$

$$f_{X_1+X_2}(z) = \int_{-\infty}^{\infty} f_{X_1}(z-u)f_{X_2}(u)du.$$

Ähnliche Formeln lassen sich auch für  $X_1 - X_2$ ,  $X_1 \cdot X_2$  und  $\frac{X_1}{X_2}$  angeben. Natürlich lässt sich dies auch auf Summen von  $X_1, \dots, X_n$  verallgemeinern. Ebenso existieren Faltungsformeln für stochastisch abhängige  $X_1, \dots, X_n$ . Jedoch muss hierfür die gemeinsame Verteilung der  $X_1, \dots, X_n$  bekannt sein. Wesentlich leichter erhält man die Faltung von Verteilungen über erzeugende Funktionen (vgl. Anhang E). Denn die Verteilung der Summe von unabhängigen Zufallsvariablen  $X_1, \dots, X_n$  kann auch als Produkt der erzeugenden Funktionen der  $X_1, \dots, X_n$  dargestellt werden.

Neben der Faltung gibt es noch weitere allgemeine Ansätze etwa bei bekannten (gemeinsamen) Dichten der  $X_1, \dots, X_p$  in Modellgleichung (4.5) die Dichte bzw. Verteilung von  $L_{PF}$  zu bestimmen. Hiermit sind die Transformationssätze für Dichten gemeint, die jedoch strenge mathematische Eigenschaften der Dichten fordern. Nach unserer Erfahrung kommen die Transformationssätze für Dichten im Zusammenhang mit Portfoliomodellen nur selten zum Einsatz.

## 4.6 Monte-Carlo-Simulation

Oft kann die Verlustverteilung nicht analytisch bestimmt werden. Mit Hilfe von Monte-Carlo-Simulationen ist es möglich, bei vorgegebenen Modellannahmen und -gleichungen, die Verlustverteilung zu schätzen. Dabei mag der Begriff *Monte-Carlo* suggerieren, dass entweder viel Glück im Spiel ist oder mit hoher Geschwindigkeit ans Ziel gelangt wird. Keines von beidem ist der Fall. Monte-Carlo ist hierbei als Synonym für zufällig oder stochastisch zu verstehen. Häufig verzichtet man auf den Zusatz *Monte-Carlo* und spricht von *Simulation* oder genauer von *stochastischer Simulation*.

Die folgenden Ausführungen sollen das prinzipielle Vorgehen bei der Monte-Carlo-Simulation illustrieren, wobei erwähnt werden sollte, dass Monte-Carlo-Techniken ganze Bücher füllen, z.B. um effizientere Algorithmen zur Simulation von Zufallszahlen darzustellen.

Wie im Abschnitt 4.5 gehen wir davon aus, dass der Portfolioverlust  $L_{PF}$  durch eine Modellgleichung oder eine Folge von Modellgleichungen der Form (4.5) beschreibbar ist:

$$L_{PF} = g(X_1, \dots, X_p, y_1, \dots, y_q). \quad (4.6)$$

In Vektorschreibweise verkürzt sich diese Modellgleichung noch etwas. Mit

$$\begin{aligned} \mathbf{X} &= X_1, \dots, X_p, \\ \mathbf{y} &= y_1, \dots, y_q \quad \text{ist} \\ L_{PF} &= g(\mathbf{X}, \mathbf{y}). \end{aligned} \quad (4.7)$$

Wie zuvor erwähnt, ist es oft nicht möglich, die Verteilung von  $L_{PF}$  aus der Modellgleichung (4.6) und den Annahmen zu den stochastischen Komponenten  $X_1, \dots, X_p$  (Verteilungen und Abhängigkeitsstruktur) analytisch zu bestimmen. In aller Regel hingegen ist möglich, bei bekannten *Realisationen*  $x_1, \dots, x_p$  der  $X_1, \dots, X_p$  den zugehörigen realisierten Portfolioverlust  $l_{PF}$  zu berechnen, ggf. mit Hilfe numerischer Methoden.

Die Grundidee der Monte-Carlo-Simulation ist nun recht einfach zu beschreiben. Es werden mit Hilfe von *Zufallszahlen* viele Realisationen der  $X_1, \dots, X_p$  erzeugt und für jedes Realisationstupel wird der zugehörige Portfolioverlust berechnet. Damit liegen ebenso viele realisierte Portfolioverluste vor, wie Realisationen der  $X_1, \dots, X_p$ . Aus den realisierten Portfolioverlusten kann dann die empirische Verteilungsfunktion (vgl. Abschnitt 4.1.5) ermittelt werden, die als Schätzer für die Verteilung des stochastischen Portfolioverlustes  $L_{PF}$  dient.

Formal stellt sich die Monte-Carlo-Simulation wie folgt dar:

1. Erzeuge  $n$  zufällige Realisationen von  $\mathbf{X}$ :  $\mathbf{x}_1, \dots, \mathbf{x}_n$ .
2. Berechne  $n$  Realisationen des Portfolioverlusts:

$$l_{PF,i} = g(\mathbf{x}_i, \mathbf{y}), \quad i = 1, \dots, n.$$

3. Bestimme aus den Realisationen des Portfolioverlusts die empirische Verteilungsfunktion des Portfolioverlusts.

Eine besondere Bedeutung kommt der Wahl der Simulationensanzahl  $n$  zu. Mit Hilfe von Computern ist es möglich, eine einzelne Realisation des Portfolioverlustes in kürzester Zeit zu berechnen. Aber auch sehr kurze Zeitabschnitte können sich zu einem beträchtlichen Zeitaufwand summieren. Das heißt, in erster Linie wird  $n$  so gewählt, dass die empirische Verteilungsfunktion des Portfolioverlustes der wahren Verlustverteilung mindestens zu einer vorgegebenen Genauigkeit nahe kommt. Die Rechenzeit wird vor allem dann zum Problem, wenn die Verlustverteilung öfter, z.B. mit verschiedenen Parametern  $\mathbf{y}$  bestimmt werden muss. Und dies ist natürlich eher die Regel als die Ausnahme. So muss die Verlustverteilung z.B. zur Portfoliooptimierung und bei Szenarioanalysen mehr als einmal geschätzt werden.

In diesem Buch wenden wir die Monte-Carlo-Simulation explizit zur Bestimmung der Verlustverteilung des Portfolios II im Rahmen des Ein-Faktor-Modells an (vgl. Abschnitt 6.1.6).

An die Zufallszahlen werden einige Anforderungen gestellt, wie wir bereits auf Seite 45 beschrieben haben. Die wichtigsten Anforderungen sind die folgenden:

- Die Realisationen der Zufallszahlen erscheinen als zufällig.
- Die Realisationen der Zufallszahlen erscheinen als stochastisch unabhängig.

In der Praxis bedeuten diese beiden Anforderungen, dass die Zufallszahlen allen Tests auf ihre Zufälligkeit und Unabhängigkeit Stand halten.

Viele Programmiersprachen und Programme, z.B. Excel, halten Routinen zur Erzeugung von Zufallszahlen bereit. Wobei es gar nicht schwierig ist, einen Zufallszahlengenerator selber zu programmieren. Algorithmen hierzu sowie Ausführungen zur Simulation generell finden sich z.B. in [24] und [67]. Dabei reicht es, gleichverteilte Zufallszahlen zur Verfügung zu haben. Realisationen aus den Verteilungen der  $X_1, \dots, X_p$  erhält man über Transformationen der gleichverteilten Zufallszahlen. Dies ist in

$$F(X) \sim U, \quad \text{mit } U \sim \text{uni}(0, 1), \quad X \sim \mathcal{F}$$

begründet, falls  $F$  die Verteilungsfunktion von  $X$  ist. Denn durch die einfache Umformung

$$F^{-1}(U) \sim \mathcal{F}$$

erhält man, dass aus gleichverteilten Zufallszahlen beliebige Zufallszahlen mit bekannter Verteilungsfunktion  $F$  und Verteilung  $\mathcal{F}$  konstruiert werden können. Hierbei muss allerdings die Einschränkung beachtet werden, dass  $F^{-1}$  existieren und bekannt sein muss.

Neben den stark algorithmischen Werken [24] und [67] eignet sich [70] als Begleiter bei der Planung und Umsetzung von Monte-Carlo-Simulationen. Die numerical recipes [67] sind auch für andere Programmiersprachen erhältlich.

Ein wichtiges Hilfsmittel für die Transformationen von unabhängigen Zufallszahlen zu einem Zufallsvektor, der aus stochastisch abhängigen Komponenten besteht, sind die Copulas, die wir im folgenden Abschnitt beschreiben.

## 4.7 Copulas

Es reicht im Allgemeinen nicht, die (eindimensionalen) Randverteilungen und die Korrelationsmatrix eines Zufallsvektors zu kennen, um die gemeinsame Verteilung der Komponenten dieses Zufallsvektors anzugeben. Das Problem ist, dass die Korrelationsmatrix nicht ausreichend geeignet ist, um komplexe Abhängigkeitsstrukturen anzugeben. *Copulas* beinhalten als Information genau die Abhängigkeitsstrukturen eines Zufallsvektors. Daher werden Copulas auch „dependence function“ genannt (vgl. [35]). Formal sind Copulas mehrdimensionale Verteilungsfunktionen, mit Gleichverteilungen auf dem Einheitsintervall als eindimensionale Randverteilungen. Eine wichtige Eigenschaft ist, dass sie auf beliebige Randverteilungen anwendbar

sind. Dennoch sind speziell für bestimmte Verteilungsfamilien geeignete Copulas entstanden, z.B. die t-Copula für t-Verteilungen (vgl. [20]).

Copulas haben Ende der 90er Jahre einen wahren Boom erlebt, nachdem sie viele Jahrzehnte weniger beachtet waren. Dies hat nach unserer Ansicht im wesentlichen zwei Gründe.

- Mehrdimensionale Verteilungen wurden in der praktischen Anwendung, und hier ist vor allem die Risikoanalyse zu nennen, immer wichtiger. Aber zu den meisten eindimensionalen Verteilungen existieren keine wohl eingeführten mehrdimensionalen Verallgemeinerungen. Hier boten sich Copulas als einfache und effektive Methode an, mehrdimensionale Verteilungsfamilien zu erzeugen.
- Durch ihre Eigenschaft mehrdimensionale Varianten gleichverteilter Zufallsvariablen zu sein, eignen sie sich besonders für Monte-Carlo-Simulationen, da Zufallszahlen einer bestimmten Verteilung oftmals aus gleichverteilten Zufallszahlen berechnet werden.

**Definition 17** Eine  $p$ -dimensionale Copula (kurz:  $p$ -Copula)  $C_p$  ist die Verteilungsfunktion eines  $p$ -dimensionalen Zufallsvektors mit Gleichverteilungen auf dem Einheitsintervall als eindimensionale Randverteilungen.

$$C_p = (u_1, \dots, u_p) : [0; 1]^p \longrightarrow [0; 1].$$

Copulas haben die folgenden Eigenschaften:

$$\begin{aligned} C_p(1, \dots, 1, u_i, 1, \dots, 1) &= u_i && \forall u_i \in [0, 1], \forall i \in \{1, \dots, p\} \\ C_p(u_1, \dots, u_p) &= 0 && \text{falls für ein } i \in \{1, \dots, p\} : u_i = 0. \end{aligned}$$

Es lässt sich zeigen, dass zu jeder  $p$ -dimensionalen Verteilungsfunktion  $F$  mit den eindimensionalen Randverteilungsfunktionen  $F_1, \dots, F_p$  eine  $p$ -Copula existiert, so dass

$$F(x_1, \dots, x_p) = C_p(F_1(x_1), \dots, F_p(x_p)).$$

Es gilt sogar, dass die Copula  $C_p$  auf dem Träger eindeutig ist. Somit kann eine  $p$ -dimensionale Verteilung in ihre eindimensionalen Randverteilungen, welche keine Informationen über Abhängigkeiten enthalten, und eine Abhängigkeitsfunktion, die keine Information über die Randverteilungen enthält, zerlegt werden. Dies kann vor allem bei der Simulation  $p$ -dimensionaler Zufallsvektoren hilfreich sein.

Für eine Einführung und Anwendung von Copulas sowie der Zitation historischer Quellen zu Copulas vgl. [44]. Für eine ausführliche Behandlung der Copulas vgl. auch [19].

### Normalcopula

$$C_p(u_1, \dots, u_p) = \Phi_p(\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_p), \mathbf{R}),$$

wobei  $\Phi_p$  die Verteilungsfunktion der  $p$ -dimensionalen Standardnormalverteilung,  $\Phi$  die Verteilungsfunktion der eindimensionalen Standardnormalverteilung und  $\mathbf{R}$  die zugehörige Korrelationsmatrix ist. Die Anwendung dieser Copula mit normalverteilten Randverteilungen ergibt wieder eine  $p$ -dimensionalen Normalverteilung mit Korrelationsmatrix  $\mathbf{R}$ .

An dem Beispiel der Normalcopula ist zu erkennen, wie man Copulas aus bekannten mehrdimensionalen Verteilungen herleiten kann und diese auf andere Randverteilungen anwenden kann. So findet sich als Beispiel hierfür in [20] die t-Copula, die aus einer Variante einer mehrdimensionalen t-Verteilung abgeleitet wird und z.B. auf normalverteilte Randverteilungen angewendet werden kann.