Seminar Quantitatives Risikomanagement

Kreditrisikomanagement II

Fabian Wunderlich

Mathematisches Institut
der
Universität zu Köln

Sommersemester 2009
Betreuung: Prof. Schmidli, J. Eisenberg

Contents

Chapter 1

Mixture Modelle

Bei Mixture Modellen geht man von folgenden grundsätzlichen Annahmen aus:
• Das Kreditausfallrisiko eines Schuldners hängt von einigen makroökonomischen Faktoren ab.
• Bedingt auf diese Faktoren sind die Ausfallrisiken einzelner Firmen unabhängig.
Natürlich gibt es trotzdem Abhängigkeiten zwischen den Ausfallwahrscheinlichkeiten, diese kommen aber daher, dass alle von denselben ökonomischen Faktoren abhängen.

1.1 Definition: Bernoulli Mixture Modelle

Sei p < m und $\Psi = (\Psi_1, ..., \Psi_p)$ ein p-dimensionaler Zufallsvektor. Der Zufallsvektor $Y = (Y_1, ..., Y_m)$ genügt einem Bernoulli Mixture Modell mit Faktorvektor Ψ wenn es für $1 \leq i \leq m$ Funktionen $p_i : \mathbb{R}^p \rightarrow [0, 1]$ gibt, so dass die Komponenten von Y bedingt auf Ψ unabhängige Bernoulli Zufallsvariablen sind und folgendes erfüllen: $\mathrm{P}(Y_i = 1|\Psi = \psi) = p_i(\psi)$

Für $y = (y_1, ..., y_m)$ aus $\{0, 1\}^m$ gilt:

$$P(Y = y|\Psi = \psi) = \prod_{i=1}^{m} p_i(\psi)^{y_i} (1 - p_i(\psi))^{1-y_i} \tag{1.1}$$

Da der Ausfall eines Kredites eher selten ist, kann man die Bernoulli Zufallsvariablen durch Poisson Zufallsvariablen approximieren. Diese Modelle nennt man dann Poisson Mixture Modelle. Dieser Ansatz führt natürlich dazu, dass eine Firma "mehr als einmal" in einem bestimmten Zeitraum Pleite gehen kann, dies allerdings mit einer sehr kleinen Wahrscheinlichkeit. Wir bezeichnen die Anzahl der Kreditausfälle einer Firma i mit $\tilde{Y}_i \in \{0, 1, 2, ...\}$ Die formale Definition ist analog zu der eines Bernoulli Mixture Modells.

1.2 Definition: Poisson Mixture Modelle

Seien p und Ψ wie in Definition 1.1. Der Zufallsvektor $\tilde{Y} = \left(\tilde{Y}_1, ..., \tilde{Y}_m\right)$ genügt einem Poisson Mixture Modell mit Faktorvektor Ψ wenn es für $1 \leq i \leq m$ Funktionen $\lambda_i : \mathbb{R}^p \to (0, \infty)$ gibt, so dass $\tilde{Y}$ bedingt auf Ψ ein Vektor von unabhängigen Poisson-verteilten Zufallsvariablen mit Parametern $\lambda_i(\psi)$ ist.

Ein Beispiel eines in der Praxis genutzten Poisson Mixture Modells ist *CreditRisk+*, dass in Abschnitt 1.4 näher betrachtet wird.

1.3 Bernoulli Mixture Modelle mit einem Faktor

In vielen praktischen Situationen ist es sinnvoll Modelle mit nur einem Einflussfaktor anzunehmen. In vielen Fällen hat man nicht die nötigen Informationen um ein Modell mit mehreren Faktoren abzustimmen, während Modelle mit nur einem Faktor relativ einfach an statistische Kreditausfalldaten angepasst werden können. Außerdem ist bei solchen Modellen das Verhalten für große Portfolios einfach zu verstehen.
In diesem Abschnitt soll also Ψ eine Zufallsvariable mit Werten in $\mathbb{R}$ sein. $p_i : \mathbb{R} \to [0, 1]$ sind Funktionen, so dass der Ausfallindikator Y, bedingt auf Ψ, ein Vektor von unabhängigen Bernoulli Zufallsvariablen mit $P(Y_i = 1|\Psi = \psi) = p_i(\psi)$ ist. Wir werden uns jetzt einige Spezialfälle anschauen.

1.3.1 Austauschbare Bernoulli Mixture Modelle

Ein Spezialfall tritt auf, wenn alle Funktionen p_i gleich sind. Das bedeutet alle Firmen haben dasselbe Ausfallrisiko. In diesem Fall nennt man das

Bernoulli Mixture Modell austauschbar. Der Grund für den Namen ist, dass man die Elemente von Y tauschen kann, ohne etwas zu ändern.
Es ist hier sinnvoll die Zufallsvariable $Q := p_1(\psi)$ einzuführen. Die Verteilungsfunktion wird mit $G(q)$ bezeichnet. Man nennt Q Mixing-Variable und G(q) Mixing-Verteilung. Bedingt auf $Q = q$ ist die Anzahl der Kreditausfälle M die Summe von m unabhängigen Bernoulli Zufallsvariablen mit Parameter q und hat daher folgende Binomialverteilung:

$$P(M = k|Q = q) = \binom{m}{k} q^k (1-q)^{m-k} \tag{1.2}$$

Die nicht bedingte Verteilung von M erhält man durch Integrieren über q, also:

$$P(M = k) = \binom{m}{k} \int_0^1 q^k (1-q)^{m-k} \, dG(q) \tag{1.3}$$

Beispiele:
Die folgenden Mixing-Verteilungen werden regelmäßig bei Bernoulli Mixture Modellen benutzt.

Beta Mixing-Verteilung
Q $\sim$ Beta(a,b) für Parameter a>0 und b>0.

Probit-Normal Mixing-Verteilung
$Q = \Phi(\mu + \sigma\psi)$ für $\psi \sim$ N(0,1), $\mu \in \mathbb{R}$ und $\sigma > 0$, wobei Φ die Standartnormalverteilung ist.

Logit-Normal Mixing-Verteilung
Q $= F(\mu + \sigma\psi)$ für $\psi \sim$ N(0,1), $\mu \in \mathbb{R}$ und $\sigma > 0$, wobei $F(x) = (1 + e^{-x})^{-1}$ die Verteilungsfunktion der sogenannten logistischen Verteilung ist.

1.3.2 Modelle mit einem Faktor und weiteren Einflussgrößen

Hier betrachtet man Bernoulli Mixture Modelle bei denen Ψ nur eine einzige gemeinsame Variable ist. Allerdings erlaubt man, dass andere Einflussgrößen das Kreditausfallrisiko für einzelne Firmen beeinflussen. Dieses können Indikatoren für die Rating-Klasse, den Industrie-Sektor oder ausgewählte Größen aus der Firmenbilanz sein.

Der Vektor $x_i \in \mathbb{R}^k$ soll diese gegebenen Einflussgrößen enthalten. Ein typisches Modell wäre, für die Kreditausfallwahrscheinlichkeiten $p_i(\psi)$ folgendes anzunehmen:

$$p_i(\psi) = h(\mu + \beta x_i + \sigma\psi) \tag{1.4}$$

wobei $h : \mathbb{R} \to (0,1)$ eine streng monoton wachsende Verknüpfungsfunktion, wie z.B. $\Phi(x)$ ist. Der Vektor β enthält Regressionsparameter, $\mu \in \mathbb{R}$ ist ein Achsenschnittpunkt und $\sigma > 0$ ein skalierender Parameter.
Falls $x_i = x$ für alle i, d.h. alle Risiken haben dieselben Einflußgrößen, so haben wir wieder volle Austauschbarkeit.

1.3.3 Modell für mehrere austauschbare Gruppen

Mit der Struktur (1.4) lassen sich auch Modelle mit mehreren austauschbaren Gruppen realisieren. Wir definieren dazu einige Gruppen, innerhalb derer die Risiken austauschbar sind. Diese Gruppen können z.B. interne oder externe Rating-Klassen repräsenieren. Wir setzen also die Einflussgrößen x_i einfach als k-dimensionale Einheitsvektoren der Form $x_i = e_{r(i)}$, wobei $r(i) \in \{1, ..., k\}$ die Rating-Klasse der Firma i ist. Dann kann man das Model aus (1.4) wie folgt schreiben:

$$p_i(\psi) = h\left(\mu + \beta_{r(i)} + \sigma\psi\right) \tag{1.5}$$

für r=1,...,k also für jede Rating-Klasse. Setzt man diese Spezifizierung in (1.1) ein, so kann man die bedingte Verteilung des Kreditausfallvektors finden. Angenommen es gibt m_r Schuldner in der Rating-Klasse r und bezeichne die Anzahl der Kreditausfälle mit M_r. Die bedingte Verteilung des Vektors $M = (M_1, ..., M_k)$ ist dann gegeben durch:

$$P(M = l|\Psi = \psi) = \prod_{r=1}^{k} \binom{m_r}{l_r} \left(h\left(\mu + \beta_{r(i)} + \sigma\psi\right)\right)^{l_r} \left(1 - h\left(\mu + \beta_{r(i)} + \sigma\psi\right)\right)^{m_r - l_r} \tag{1.6}$$

wobei $l = (l_i, ..., l_k)$ ist.

1.4 CreditRisk+

CreditRisk+ ist ein industrielles Kreditriskomodell, dass 1997 von Credit-Suisse herausgebracht wurde. Es hat die Struktur eines Poisson Mixture

Modells, bei dem der Faktorvektor Ψ aus p unabhängigen, gammaverteilten Zufallsvariablen besteht. Die Verteilungsannahmen und Struktur des Modells machen es möglich die Verteilung von M relativ explizit zu berechnen.

Struktur von CreditRisk+

CreditRisk+ ist ein Poisson Mixture Modell.

Die Parameter $\lambda_i(\Psi)$ der bedingten Poissonverteilung von Firma i ist gegeben durch $\lambda_i(\Psi) = k_i \omega_j' \Psi$, wobei

• $k_i > 0$ eine Konstante

• $\omega_i = (\omega_{i1}, ..., \omega_{i1})'$ positive Faktorgewichte mit $\sum_j \omega_{ij} = 1$

• $\psi_1, ..., \psi_p$ unabhängige $Ga(\alpha_j, \beta_j)$-verteilte Faktoren mit Parametern $\alpha_j = \beta_j = \sigma_j^{-2}$ für $\sigma_j > 0$

Durch diese Parametrisierung erreicht man, dass gilt:

• $E(\psi_j) = 1$

• $var(\psi_j) = \sigma_j^2$

• $E(\lambda_i(\Psi)) = k_i E(\omega_i' \Psi) = k_i$

Als Poisson Mixture Modell ist in diesem Modell die Ausfallwahrscheinlichkeit gegeben durch: $P(Y_i = 1) = P(\tilde{Y}_i > 0) = E(P(\tilde{Y}_i > 0)|\Psi)$. Weil $\tilde{Y}_i$ bedingt auf Ψ Poissonverteilt ist, gilt:

$$E(P(\tilde{Y}_i > 0|\Psi)) = E(1 - e^{-k_i wi' \Psi}) \approx k_i E(w_i' \Psi) = k_i \tag{1.7}$$

Somit ist k_i ungefähr gleich der Ausfallwahrscheinlichkeit von Firma i.

Man interessiert sich für die Verteilung von $\tilde{M} = \sum_{i=1}^m \tilde{Y}_i$. Mit der Struktur von CreditRisk+ ist es tatsächlich möglich einfache Rekursionsformeln für die Wahrscheinlichkeiten $P(\tilde{M} = k)$ herzuleiten.

1.5 Asymptotisches Verhalten für große Porfolios

Hier möchten wir einige asymptotische Betrachtungen für Bernoulli Mixture Modelle anstellen. Die Ergebnisse davon sind können hilfreich sein um die Verlustverteilungen in großen Portfolios anzunähern.

Da wir nicht (nur) an den Kreditausfällen als solches, sondern an dem entstehenden Gesamtverlust interessiert sind, müssen wir hier auch die Verluste bei gegebenem Kreditausfall betrachten. Sei also $(e_i)_{i\in\mathbb{N}}$ eine unendliche Folge von gegebenen positiven Risikopositionen. Sei $(Y_i)_{i\in\mathbb{N}}$ die zugehörige Folge von Kreditausfallindikatoren. Sei $(\delta_i)_{i\in\mathbb{N}}$ eine Folge von Zufallsvariablen mit

Werten in $(0,1]$, die die prozentualen Verluste bei auftretendem Kreditausfalls repräsentieren. Bei einem Portfolio der größe m wäre der Verlust hier: $L^{(m)} = \sum_{i=1}^{m} L_i$, wobei $L_i = e_i \delta_i Y_i$ die einzelnen Verluste der Risikopositionen sind.

Wir müssen nun noch drei technische Annahmenfür unsere Betrachtungen treffen:

(A1) Es gibt einen p-dimensionalen Zufallsvektor Ψ und Funktionen $l_i : \mathbb{R}^p \rightarrow [0,1]$, so dass $(L_i)_{i \in \mathbb{N}}$ bedingt auf Ψ eine Folge unabhängiger Zufallsvariablen mit $l_i(\psi) = E(L_i | \Psi = \psi)$

(A2) Es gibt eine Funktion $\bar{l} : \mathbb{R}^p \rightarrow \mathbb{R}^+$, so dass:
$\lim\limits_{m \to \infty} \frac{1}{m} E\left(L^{(m)} | \Psi = \psi\right) = \lim\limits_{m \to \infty} \frac{1}{m} \sum_{i=1}^{m} l_i(\psi) = \bar{l}(\psi)$
für alle $\psi \in \mathbb{R}^p$. Wir nennen $\bar{l}(\psi)$ die asymptotische, bedingte Verlustfunktion.

(A3) Es gibt ein C $< \infty$, so dass $\sum_{i=1}^{m} \left(\frac{e_i}{i}\right)^2 < C$ für alle m.
Die dritte Annahme verhindert, dass die Werte der Risikopositionen systematisch mit dem Portfolio wachsen.

Proposition 1: Sei $L^{(m)} = \sum_{i=1}^{m} L_i$ eine Folge, die die obigen 3 Annahmen erfüllt. Bezeichne die bedingte Verteilung der Folge $(L_i)_{i \in \mathbb{N}}$ gegeben $\Psi = \psi$ mit $P(\cdot | \Psi = \psi)$. Dann gilt:
$\lim\limits_{m \to \infty} \frac{1}{m} L^{(m)} = \bar{l}(\psi)$, $P(\cdot | \Psi = \psi)$ fast sicher

Proposition 2: Sei $L^{(m)} = \sum_{i=1}^{m} L_i$ eine Folge, die die obigen 3 Annahmen erfüllt, wobei die Mixing-Variable Ψ eindimensional ist und Verteilungsfunktion G hat. Angenommen die bedingte, asymptotische Verlustfunktion $\bar{l}(\psi)$ ist monoton wachsend und rechtsstetig. Nehme weiterhin an, dass G monoton wachsend in $q_\alpha(\Psi)$ ist, d.h. $G(q_\alpha(\Psi) + \delta) > \alpha$ für alle $\delta > 0$. Dann gilt:
$\lim\limits_{m \to \infty} \frac{1}{m} q_\alpha\left(L^{(m)}\right) = \bar{l}(q_\alpha(\Psi))$

Die beiden Propositionen zeigen folgendes:
Der durchschnittliche Portfolioverlust wird wesentlich durch die bedingte, asymptotische Verlustfunktion $\bar{l}$ und die Realisation des Faktorvektors Ψ bestimmt. Bei Bernoulli Mixture Modellen mit nur einem Faktor ist die Flanke der Verlustverteilung wesentlich durch die Flanke der Mixingverteilung festgelegt.

1.6 Grenzmodelle als Mixture Modelle

Grenzmodelle und Mixture Modelle sind grundsätzlich unterschiedlich. Allerdings kann man zu den meisten Grenzmodellen zugehörige Mixture Modelle konstruieren, die diese Grenzmodelle repräsentieren. Das ist deshalb nützlich, weil die Mixture Modelle gegenüber den Grenzmodellen einige Vorteile haben:
• Bernoulli Mixture Modelle bieten sich an, um diese mit Monte Carlo Methoden zu untersuchen.
• Bernoulli Mixture Modelle sind besser geeignet, um sie mit Hilfe von empirischen Daten anzupassen.
• Das Verhalten von Bernoulli Mixture Modellen für große Portfolios ist relativ einfach zu verstehen, da es stark von der Verteilung der ökonomischen Faktoren abhängt.
Die folgende Bedingung stellt sicher, dass ein Grenzmodell als ein Bernoulli Mixture Modell geschrieben werden kann.

1.6.1 Definition: bedingte Unabhängigkeitsstruktur

Ein Zufallsvektor $X = (X_1, ..., X_m)$ hat eine p-dimensionale bedingte Unabhängigkeitsstruktur mit bedingender Variable Ψ, falls p<m ist und es einen p-dimensionalen Zufallsvektor $\Psi = (\psi_1, ..., \psi_p)$ gibt, so dass die Zufallsvariablen $X_1, ..., X_m$ bedingt auf Ψ unabhängig sind.

1.6.2 Lemma: Repräsentation von Grenzmodellen durch Mixture Modelle

Sei (X, D) ein Grenzmodell für einen m-dimensionalen Zufallsvektor X und X habe eine p-dimensionale bedingte Unabhängigkeitsstruktur mit bedingender Variable Ψ. Dann genügen die Kreditausfallindikatoren $Y_i = I_{\{X_i \leq d_{i1}\}}$ einem Bernoulli Mixture Modell mit Faktor Ψ, wobei die bedingten Kreditausfallwahrscheinlichkeiten durch $p_i(\psi) = P(X_i \leq d_{i1} | \Psi = \psi)$ gegeben sind.

Beweis. Für $y \in \{0,1\}^m$ definieren wir 2 Mengen:
$B := \{1 \leq i \leq m : y_i = 1\}$ und deren Komplement:
$B^c = \{1, ..., m\} \backslash B$
Wir erhalten dann:
$P(Y = y | \Psi = \psi) = P\left(\bigcap_{i \in B} \{X_i \leq d_{i1}\} \bigcap_{i \in B^c} \{X_i > d_{i1}\} | \Psi = \psi\right)$
$= \prod_{i \in B} P(X_i \leq d_{i1} | \Psi = \psi) \prod_{i \in B^c} (1 - P(X_i \leq d_{i1} | \Psi = \psi))$
Somit sind die Y_i, bedingt auf $\Psi = \psi$ unabhängige Bernoulli Zufallsvariablen mit Erfolgswahrscheinlichkeit $p_i(\psi) = P(X_i \leq d_{i1} | \Psi = \psi)$ ∎

1.7 Modellrisiken

Wir betrachten hier die Modellrisiken für ein austauschbares Bernoulli Mixture Modell und gehen dabei davon aus, dass π und ρ_Y fest und bekannt sind. Wir wissen, dass die Flanke von M wesentlich von der Flanke der Mixing-Variable Q bestimmt wird. In der Tabelle 8.3 ist die rechte Flanke der Probit-normal, Beta und Logit-normal Verteilung, sowie die der zur Clayton Copula gehörenden Mixture-Verteilung abgebildet. In allen Fällen liegen die Werte $\pi = 0.049$ und $\pi_2 = 0.00313$ zu Grunde. Das entspricht etwa der Standart&Poor's Rating-Klasse B. Man sieht, dass die Funktionen erst nach dem 99%-Quantil deutlich voneinander abweichen. Natürlich heißt das nicht, dass Bernoulli Mixture Modelle frei von Modellrisiken sind, allerdings zeigt es, dass die Wahl der Mixing-Verteilung von eher geringer Bedeutung ist.

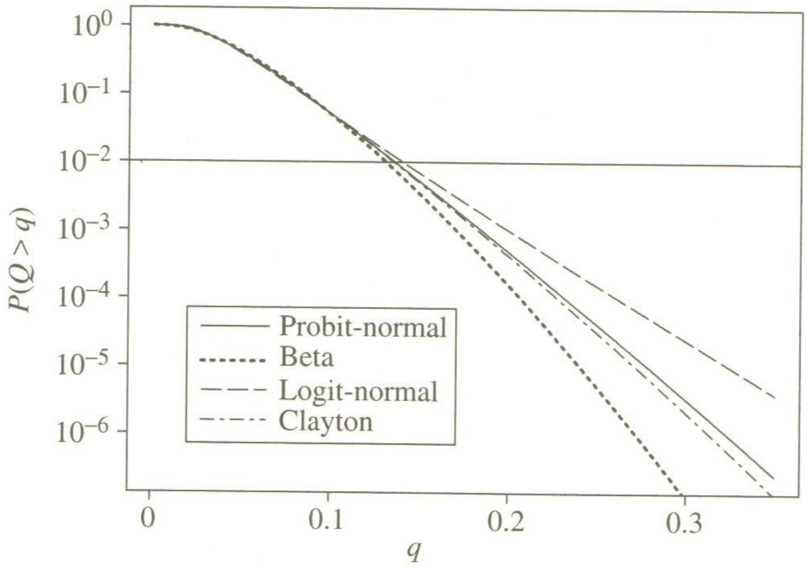

Figure 8.3. Tail of the mixing distribution of Q in four different exchangeable Bernoulli-mixture models: beta; probit-normal (one-factor KMV/CreditMetrics); logit-normal (CreditPortfolioView); Clayton. In all cases the first two moments have the values $\pi = 0.049$ and $\pi_2 = 0.003\,13$, which correspond roughly to Standard & Poor's rating category B; the actual parameter values can be found in Table 8.6. The horizontal line at 10^{-2} shows that the models only really start to differ around the 99th percentile of the mixing distribution.

Ein anderer Gesichtspunkt des Modellrisikos ist die Modellierung des Verlustes bei gegebenem Kreditausfall. Normalerweise wird in den Modellen angenommen, dass der Verlust bei gegebenem Kreditausfall unabhängig vom Ereigniss des Kreditausfalls ist. In empirischen Studien wurde allerdings gezeigt, dass das nicht richtig ist. Tatsächlich ist z.B. in wirtschaftlichen Abschwungzeiten nicht nur die Ausfallwahrscheinlichkeit höher, sondern auch der Verlust bei gegebenem Ausfall. Das kann dazu führen, dass das benötigte Risikokapital für Banken in diesen Modellen erheblich unterschätzt wird.

Chapter 2

Monte Carlo Methoden

Idee der Monte Carlo Simulation ist es analytisch schwer lösbare Aufgaben durch häufige Simulation von Zufallsvarianlen mit bekannter Dichte anzunähern. In diesem Abschnitt betrachten wir ein Bernoulli Mixture Modell für ein Kreditportfolio. Wir nehmen dabei an, dass der Gesamtverlust von der Form $L = \sum_{i=1}^{m} L_i$ ist, wobei die einzelnen Verluste L_i bedingt unabhängig bzgl. einem gegebenem Faktorvektor Ψ sind. Angenommen wir wollen den Expected Shortfall und die Beiträge zum Expected Shortfall zum Konfidenzniveau α berechnen. Dann müssen wir folgende bedingte Erwartungen berechnen:

$$E\left(L|L \geq q_\alpha\left(L\right)\right) \ und \ E\left(L_i|L \geq q_\alpha\left(L\right)\right) \tag{2.1}$$

Solche Berechnungen kann man mit Hilfe von Monte Carlo Integration oder Importance Sampling durchführen.

2.1 Monte Carlo Integration und Importance Sampling

Sei (Ω, F, P) ein Wahrscheinlichkeitsraum und X eine Zufallsvariable mit Dichte f. Wir betrachten die Berechnung von folgendem Erwartungswert:

$$\theta = E\left(h\left(X\right)\right) = \int_{-\infty}^{\infty} h\left(x\right) f\left(x\right) dx \tag{2.2}$$

für eine bekannte Funktion h. Um die Wahrscheinlichkeit eines Ereignisses zu berechnen, betrachten wir eine Funktion der Form $h\left(x\right) = I_{x \in A}$ für eine

Menge $A \subset \mathbb{R}$. Für die Berechnung eines Expected Shortfalls betrachten wir eine Funktion der Form $h(x) = xI_{x \geq c}$ für ein $c \in \mathbb{R}$. Wenn die Komplexität der Verteilung von X keine einfache Berechnung von (8) ?!?!?!?!?! zulässt, können wir auf einen Monte Carlo Ansatz zurückgreifen, für den es lediglich notwendig ist Zufallsvariablen mit Dichte f zu erzeugen.

2.1.1 Algorithmus: Monte Carlo Integration

(1) Erzeuge $X_1, ..., X_n$ unabhängig mit Dichte f.
(2) Berechne den Standart Monte Carlo Schätzer $\widehat{\theta}_n^{MC} = \frac{1}{n}\sum_{i=1}^{n} h(X_i)$

Wegen dem starken Gesetz der großen Zahlen konvergiert der Monte Carlo Schätzer gegen θ. Folgende Überlegung zeigt, dass der Monte Carlo aber nur bedingt geeignet ist um seltene Ereignisse zu schätzen:
Ist beispielsweise $\alpha = 0.99$, dann werden durchschnittlich nur 1% aller Simulationen zu einem Gesamtverlust führen, der größer als $q_\alpha = q_{0.99}(L)$ ist. Das Problem ist, dass dann fast alle Simulationen verschwendet werden, weil sie zu einem Verlust $L \geq q_\alpha(L)$ führen (der in diesem Fall uninteressant ist). Der Monte Carlo Schätzer wird daher instabil und stark schwankend sein, es sei denn die Anzahl der Simulationen ist sehr groß.
Um dieses Problem zu umgehen, benutzt man das sogenannte Importance Sampling, eine Technik der Varianzreduktion, die gut für diese Probleme geeignet ist. Die Grundidee ist eine andere Schreibweise des Integrals (2.2). Dafür benutzt man eine andere Wahrscheinlichkeitsdichte g und definiert das Likelihood-Verhältnis r(x) mit $r(x) := \frac{f(x)}{g(x)}$ falls g(x)>0 und r(x)=0 sonst. Dann kann das Integral wie folgt umgeschrieben werden:

$$\theta = \int_{-\infty}^{\infty} h(x)r(x)g(x)dx = E_g(h(X)r(X)) \tag{2.3}$$

wobei E_g den Erwartungswert bzgl. der Dichte g bezeichnet. Somit können wir das Integral mit folgendem Algorithmus approximieren:

2.1.2 Algorithmus: Importance Sampling

(1) Erzeuge $X_1, ..., X_n$ unabhängig mit Dichte g.
(2) Berechne den Importance Sampling Schätzer $\widehat{\theta}_n^{IS} = \frac{1}{n}\sum_{i=1}^{n} h(X_i)r(X_i)$

Die Dichte g wird häufig als Importance-Sampling-Dichte bezeichnet. Die Kunst beim Importance Sampling ist es, diese Dichte so zu wählen, dass die Varianz des IS-Schätzers für festes n deutlich kleiner als die des MC-Schätzers wird. Die beiden Varianzen sind gegeben durch:

$$var_g(\widehat{\theta}_n^{IS}) = \frac{1}{n}(E_g(h(X)^2 r(X)^2) - \theta^2) \tag{2.4}$$

$$var(\widehat{\theta}_n^{MC}) = \frac{1}{n}(E(h(X)^2) - \theta^2) \tag{2.5}$$

Theoretisch kann die Varianz von θ_n^{IS} auf 0 reduziert werden, womit der richtige Wert schon bei der ersten Simulation erreicht würde. In der Praxis ist das nicht möglich, weil dies voraussetzen würde schon vorher den gesuchten Erwartungswert zu kennen.

Wir betrachten nun die Schätzung einer Wahrscheinlichkeit für ein seltenes Ereignis. In diesem Fall ist dann also $h(x) = I_{\{x \geq c\}}$ für ein c, dass deutlich größer als der Mittelwert von X ist. Wir erhalten dann $E(h(X)^2) = E(h(X)) = P(X \geq c)$ und mit (2.3):

$$E_g(h(X)^2 r(X)^2) = E_g(r(X)^2; X \geq c) = E(r(X); X \geq c) \tag{2.6}$$

Da wir genau diesen Ausdruck klein bekommen wollen, müssen wir g so wählen, dass das Likelihood-Verhältnis $r(x) = \frac{f(x)}{g(x)}$ klein wird für $x \geq c$. Mit anderen Worten versuchen wir das Ereignis $\{X \geq c\}$ bzgl. der Dichte g wahrscheinlicher zu machen, als bzgl. der Dichte f. Eine Möglichkeit das zu erreichen ist die Methode des Exponential Tilting.

Grobe Idee des Exponential Titlting

(1) Definiere die IS-Dichte $g_t(x) := \frac{e^{tx} f(x)}{M_X(t)}$ wobei $t \in \mathbb{R}$ und $M_X(t)$ die Momentenerzeugende Funktion von X ist.

(2) Definiere nun $\mu_t := E_{g_t}(X)$, d.h. der Erwartungswert von X bezüglich der IS-Dichte.

Nun muss man natürlich noch den Wert t für die IS-Dichte wählen. Man kann zeigen, dass t für unsere Zwecke optimal ist, wenn man es wie folgt wählt:

(3) Wähle t, so dass $\mu_t = c$ erfüllt ist.

Diese Wahl von t ist nicht verwunderlich, denn somit erreicht man, dass das eigentlich seltene Ereignis $\{X \geq c\}$ bzgl. der IS-Dichte völlig normal (d.h. häufig) ist.

Exponential Tilting bei der Standartnormalverteilung
Besonders einfach ist das Exponential Tilting bei einer Standartnormalverteilten ZV X $\sim$ N(0,1). Hier ist die Verteilung von X nach dem Exponential Tilting einfach X $\sim$ N(t,1), hier wird also eim Grunde genommen nur der Durchschnittswert von X verschoben.

Ähnlichen Methoden kann man auch auf die Verlustfunktionen von Bernoulli Mixture Modellen anwenden. Somit kann man seltene Ereignisse oder beispielsweise den Expected Shortfall berechnen.

Chapter 3

Statistischer Rückschluss für Mixture Modelle

In diesem Kapitel betrachten wir die statistische Schätzung von Modelparametern aufgrund von historischen Ausfalldaten.

Für die Kalibrierung von industriell genutzten Kreditrisikomodellen für Portfolios, werden die Modelparameter in der Regel nicht nur durch formale statistische Schätzung aufgrund von historischen Ausfall- und Kreditmigrationsdaten berechnet. Der Hauptgrund dafür ist, dass (speziell für Firmen mit hohem Rating) nicht genug relevante Ausfalldaten vorliegen.
Die Ausfallwahrscheinlichkeit einer einzelnen Firma wird normalerweise durch eine historische Ausfallrate einer ähnlichen Firma geschätzt. Ähnlich bedeutet hier beispielsweise dieselbe Rating-Klasse oder einen ähnlichen Distanco-to-Default Wert.
Dass die meisten industriellen Modelle wenig formal statistisch sind, liegt an der Wahl der Modelparameter. Die meisten Modelle enthalten plausible Faktoren und Strukturen, die die Abhängigkeiten der Ausfälle beschreiben. Die Parameter dieser Faktoren werden aber meistens einfach aufgrund von ökonomischen Argumenten festgelegt, oder durch die Analyse verwandter Variablen abgeleitet. Diese Art der Parameterwahl lässt natürlich die Frage offen, wie sehr man diesen Parametern vertrauen kann und wieviel Modellrisiko bleibt.
In diesem letzten Kapitel betrachten wir Methoden für rein statistische

Schätzung aller Modelparameter. Wie oben schon erwähnt ist dies nur sinnvoll möglich für Risiken, bei denen sowohl für Ausfallwahrscheinlichkeit, als auch für Ausfallabhängigkeiten genügend Daten vorhanden sind.
Grundlage der folgenden Überlegungen sind Daten über Kreditausfall verschiedener Gruppen von Firmen über mehrere Zeiträume.

3.1 Austauschbare Bernoulli Mixture Models

Hier geht es um die Schätzung von Ausfallwahrscheinlichkeit und Ausfallabhängigkeiten für homogene Gruppen von Firmen. Angenommen wir haben historische Daten von n Zeiträumen (normalerweise Jahre). Für t=1,...,n sei:

- m_t die feste Anzahl der Firmen zu Beginn des Zeitraumes
- M_t eine Zufallsvariable, die die Anzahl der in diesem Zeitraum zahlungsunfähig gewordenen Firmen angibt.

Außerdem nehmen wir an, dass die Kreditausfälle durch ein austauschbares Bernoulli Mixture Modell erzeugt wurden und dass die Mixing-Variablen $Q_1, ..., Qn$ identisch verteilt sind.
Wir betrachten zwei Methoden um die fundamentalen Parameter der Mixing-Verteilung π, π_2 und ρ_Y zu schätzen.

Ein einfacher Momentenschätzer
Für $1 \leq t \leq n$ seien $Y_{t,1}, ..., Y_{t,m_t}$ die Ausfallindikatoren für die m_t Firmen in der Gruppe. Wir definieren folgende Zufallsvariable:

$$\binom{M_t}{k} := \sum_{\{i_1,...,i_k\} \subset \{1,...,m_t\}} Y_{t,i_1}, ..., Y_{t,i_k} \tag{3.1}$$

Diese repräsentiert die Anzahl der Untergruppen von k Schuldnern, die in einer Periode zahlungsunfähig werden. Wir betrachten den Erwartungswert dieser Zufallsvariable und erhalten:
$E(\binom{M_t}{k}) = \binom{m_t}{k}\pi_k$
Durch umformen ergibt sich:
$\pi_k = E(\binom{M_t}{k})/\binom{m_t}{k}$

Somit können wir den gesuchten Wert π_k schätzen, in dem wir den Erwartungswert in der Formel durch den empirischen Durchschnitt der Daten aus n Jahren ersetzen

$$\hat{\pi_k} = \frac{1}{n}\sum_{t=1}^{n}\frac{\binom{M_t}{k}}{\binom{m_t}{k}} = \frac{1}{n}\sum_{t=1}^{n}\frac{M_t(M_t-1)\cdots(M_t-k+1)}{m_t(m_t-1)\cdots(m_t-k+1)} \tag{3.2}$$

Für k=1 erhalten wir den Standartschätzer der Ausfallwahrscheinlichkeit:

$$\hat{\pi} = \frac{1}{n}\sum_{t=1}^{n}\frac{M_t}{m_t} \tag{3.3}$$

Wir wissen, dass für die Korrelation gilt: $\rho_Y = \frac{\pi_2-\pi^2}{\pi-\pi^2}$. Folglich können wir die Korrelation leicht schätzen durch $\hat{\rho_Y} = \frac{\hat{\pi_2}-\hat{\pi}^2}{\hat{\pi}-\hat{\pi}^2}$

Maximum Likelihood Schätzer

Um einen Maximum Likelihood Schätzer zu realisieren, müssen wir für Q_t eine einfache Form (so wie Beta, Logit-Normal oder Probit-Normal) annehmen. Die gemeinsame Wahrscheinlichkeitsfunktion der Ausfallzähler $M_1, ..., M_n$ gegeben $m_1, ..., m_n$ kann durch (1.3) berechnet werden, wobei wir annehmen müssen, dass die Q_t in verschiedenen Jahren unabhängig sind. Dieser Ausdruck wird dann bzgl. der Parameter der Mixing-Verteilung (a und b bei Beta; μ und σ bei Logit-Normal und Probit-Normal) maximiert.

In der Praxis ist es am Einfachsten die Beta Mixing-Verteilung zu wählen. In diesem Fall ist die Wahrscheinlichkeit, dass von den m_t Firmen genau M_t Firmen zahlungsunfähig werden, gegeben durch: $\binom{m_t}{M_t}\frac{\beta(a+M_t,b+m_t-M_t)}{\beta(a,b)}$ Der zu maximierende Likelihood hat deshalb folgende Form:

$$L(a,b;data) = \prod_{t=1}^{n}\binom{m_t}{M_t}\frac{\beta(a+M_t,b+m_t-M_t)}{\beta(a,b)} \tag{3.4}$$

Diesen kann man numerisch bzgl. a und b maximieren und erhält den gesuchten Likelihood-Schätzer.

Um die beiden Ansätze zu vergleichen, simulieren wir Daten von 20 Jahren für verschiedene Modelle. Danach schätzen wir die Parameter π, π_2 und ρ_Y und errechnen den Schätzfehler für jeden Parameter (RRMSE). Durch Δ wird der prozentuale Anstieg des Fehlers verglichen mit der besseren Methode ausgedrückt. Tabelle 8.7 zeigt die Ergebnisse.

Table 8.7. Each part of the table relates to a block of 5000 simulations using a particular exchangeable Bernoulli mixture model with parameter values roughly corresponding to a particular S&P rating class. For each parameter of interest, an estimated RRMSE is tabulated for both estimation methods: moment estimation using (8.61) and ML estimation based on the beta model. Methods can be compared by using Δ, the percentage increase of the estimated RRMSE with respect to the better method (i.e. the RRMSE minimizing method) for each parameter. Thus, for each parameter the better method has $\Delta = 0$. The table clearly shows that MLE is better in all but one case.

			Moment		MLE-beta	
Group	True model	Parameter	RRMSE	Δ	RRMSE	Δ
CCC	Beta	π	0.101	0	0.101	0
CCC	Beta	π_2	0.202	0	0.201	0
CCC	Beta	ρ_Y	0.332	5	0.317	0
CCC	Probit-normal	π	0.100	0	0.100	0
CCC	Probit-normal	π_2	0.205	1	0.204	0
CCC	Probit-normal	ρ_Y	0.347	11	0.314	0
CCC	Logit-normal	π	0.101	0	0.101	0
CCC	Logit-normal	π_2	0.209	1	0.208	0
CCC	Logit-normal	ρ_Y	0.357	11	0.320	0
B	Beta	π	0.130	0	0.130	0
B	Beta	π_2	0.270	0	0.269	0
B	Beta	ρ_Y	0.396	8	0.367	0
B	Probit-normal	π	0.130	0	0.130	0
B	Probit-normal	π_2	0.286	3	0.277	0
B	Probit-normal	ρ_Y	0.434	19	0.364	0
B	Logit-normal	π	0.131	0	0.132	0
B	Logit-normal	π_2	0.308	7	0.289	0
B	Logit-normal	ρ_Y	0.493	26	0.392	0
BB	Beta	π	0.199	0	0.199	0
BB	Beta	π_2	0.435	0	0.438	1
BB	Beta	ρ_Y	0.508	7	0.476	0
BB	Probit-normal	π	0.197	0	0.197	0
BB	Probit-normal	π_2	0.492	10	0.446	0
BB	Probit-normal	ρ_Y	0.607	27	0.480	0
BB	Logit-normal	π	0.196	0	0.196	0
BB	Logit-normal	π_2	0.572	24	0.462	0
BB	Logit-normal	ρ_Y	0.752	45	0.517	0