

Dynamische Frailties in Zählprozessen mit Anwendung auf Ratingmigrationen

Dissertation

zur Erlangung des akademischen Grades
Doktor der Naturwissenschaften

Der Fakultät Statistik
der Technischen Universität Dortmund
vorgelegt von

Ronja Anna Walter
aus Frankfurt am Main

Dortmund, 2011

Tag der mündlichen Prüfung: 16. September 2011

Gutachter: Prof. Dr. Walter Krämer
Prof. Dr. Rafael Weißbach
Prof. Dr. Roland Fried

Inhaltsverzeichnis

1. Einleitung	3
2. Ratinghistorien und dynamische Frailties	6
2.1. Ratingverläufe und Zählprozesse	6
2.2. Frailties	8
2.3. Dynamische Frailties	11
3. Likelihood für einen Zählprozess mit dynamischer Frailty	15
3.1. Laplace-Approximation	15
3.2. Betaverteilte Frailty	18
3.3. AR(1)-Frailty	27
4. Simulationsstudien	36
4.1. Numerische Verfahren	36
4.2. Schätzen bei Ignorieren der Frailty	38
4.3. Likelihoodschätzer bei betaverteilter Frailty	40
4.4. Likelihoodschätzer bei einer AR(1)-Frailty	46
5. Anwendung auf ein Portfolio der WestLB AG	49
5.1. Daten	49
5.2. Ergebnisse	51
6. Zusammenfassung und Diskussion	54
Literatur	61
Anhang	62
A. R-Code	62
A.1. Simulationen	62
A.2. Datenauswertung	75

1. Einleitung

Das Kreditgeschäft ist ein wichtiger Teil des Bankwesens. Dabei ist vor allem im Zuge von Basel II und der Aufarbeitung der Finanzkrise die Ermittlung des Kreditrisikos von besonderer Bedeutung. Um dieses Kreditrisiko möglichst gut einschätzen zu können, entwickeln Banken eigene Ratingsysteme, die nicht wie bei externen Ratingagenturen zeitdiskret, sondern zeitstetig berechnet werden. Das bedeutet, dass man zu jedem Zeitpunkt das genaue Rating eines Kontrahenten im Portfolio angeben kann. Der Ratingverlauf eines solchen Kontrahenten lässt sich als zeitstetiger Markoff-Prozess ansehen. Somit ist eine Modellierung des Kreditrisikos möglich.

Ein homogener Markoff-Prozess mit konstanten Migrationsintensitäten ist immer noch der übliche Weg Ratingmigrationen zu modellieren (Casjens et al., 2007; Bluhm et al., 2002, S. 197ff). Immer mehr Arbeiten zielen aber auf das Verwerfen dieser Annahme auf Grund von mikro- und makroökonomischen Einflussvariablen ab, siehe Lando und Skødeberg (2002), Altman und Kao (1992) und Bangia et al. (2002). Betrachtet man zugehörige Ratinghistorien in Kalenderzeit, ist klar, dass die Migrationswahrscheinlichkeiten und somit auch das Kreditrisiko von der Konjunktur abhängen. Die Beachtung dieser Korrelation der Migrationen, die z. B. Schönbucher (2003) bestätigt, und der Abhängigkeit von einem zusätzlichen Faktor wird vor allem dazu führen, Risiken realistischer einzuschätzen. Explizit sollte sich die Einschätzung des Risikos durch die Hereinnahme dieses zusätzlichen Faktors verbessern. Koopman et al. (2009) zeigen, dass ein solcher Einfluss auch zu einer anderen Einschätzung der statistischen Signifikanz der sonst betrachteten Makrovariablen führt. Stefanescu et al. (2009) modellieren ihn mit Hilfe von Brownsch'schen Bewegungen in einem Bayesian Hierarchical Model.

In dieser Arbeit wird die Abhängigkeit mit Hilfe eines latenten Faktors, einer so-

genannten Frailty, modelliert. Die Frailty sollte sich auf jeden Fall über die Zeit verändern, um die Veränderung der Konjunktur abzubilden. Idealerweise sollte zusätzlich eine Abhängigkeit über die Zeit möglich sein um den Verlauf der Konjunktur realistischer nachzubilden. Bei der Schätzung ergibt sich die Likelihood in diesen Fällen immer zu einem (je nach Modellierung) hochdimensionalen Integral. Delloye et al. (2006), Koopman et al. (2008) und Duffie et al. (2009) haben das Problem im Umfeld von Ratingmigrationen mit Hilfe von aufwändigen numerischen Verfahren gelöst. Hier soll eine Laplace-Approximation, ähnlich wie in Gagliardini und Gourieroux (2008) oder Harding und Hausman (2007), zur Annäherung des Integrals verwendet werden. Pinheiro und Bates (1995) sehen das für lineare Modelle mit zufälligen Effekten als einen sinnvollen Ansatz. So wird nur noch ein einfacher Algorithmus, wie z.B. ein Newton-Raphson Algorithmus oder das Verfahren von Nelder und Mead, zur Bestimmung des Maximums der Likelihoodfunktion benötigt. Die Arbeit stellt zwei Modelle vor: eine betaverteilte Frailty, die zeitdynamisch ist, und eine Frailty, die einem AR(1) Prozess entspricht und somit zusätzlich über die Zeit korreliert.

In beiden Fällen zeigen Simulationen, dass die Basisintensität gut zu schätzen ist, dass aber nur verzerrte bzw. im Falle der AR(1)-Frailty gar keine Schätzer für die Parameter der Frailtyverteilung zu berechnen sind. Die Verzerrung der betaverteilten Frailty resultiert hauptsächlich aus einem Teil des Korrekturterms der Laplace-Approximation, der nicht von den Daten abhängt. Bei einer Anwendung auf Daten wird jedoch deutlich, dass das Einsetzen verschiedener Werte für die als bekannt angenommenen Verteilungsparameter kaum Einfluss auf den Schätzer der Basisintensität hat.

Das folgende Kapitel beschäftigt sich zunächst mit den für die Schätzung bei Ratinghistorien benötigten Zählprozessen. Außerdem wird das Prinzip der Frailty vorgestellt. Kapitel 3 beschreibt dann die Laplace-Approximation und die Berechnung der Likelihood der zwei verschiedenen Modelle: Ratingmigrationen mit einer betaverteilten Frailty und einer AR(1)-Frailty. Die Ergebnisse werden mit Hilfe von Simulatio-

nen in Kapitel 4 überprüft. Unter anderem werden die so gefundenen Schätzer mit einem einfachen, der die Frailty ignoriert, verglichen. In Kapitel 5 werden die Modelle auf Daten der WestLB angewandt. Kapitel 6 fasst die gefundenen Erkenntnisse zusammen und diskutiert sie.

2. Ratinghistorien und dynamische Frailties

Dieses Kapitel stellt zunächst die grundlegende Theorie zu Zählprozessen und deren Verknüpfung zu Ratinghistorien vor. Es folgt eine Erläuterung des Begriffs der Frailty und eine Einführung in bisherige Arbeiten zu dem Thema. Außerdem werden Modelle mit Frailties in Verbindung mit Ratinghistorien vorgestellt.

2.1. Ratingverläufe und Zählprozesse

Betrachten wir den Markoff-Prozess $X = \{X_t, t \in [0, T]\}$ auf einem Wahrscheinlichkeitsraum $(\Omega, \mathfrak{F}, P)$ und mit einer Filtration \mathfrak{F}_t mit $t \in [0, T]$, wobei $X_t \in \{1, \dots, k\}$ das Rating eines Kontrahenten zum Zeitpunkt t angibt. Das Rating 1 beschreibt dabei die beste Klasse, k dagegen die absorbierende Klasse „Ausfall“. Ein solcher Markoff-Prozess wird durch die Migrationsmatrizen

$$P(s, t) = (p_{hj}(s, t))_{h,j=1,\dots,k} \in \mathbb{R}^{k \times k}; \quad s, t \in [0, T], s \leq t,$$

beschrieben. Die $p_{hj}(s, t)$ geben dabei die bedingte Wahrscheinlichkeit dafür an, bis zum Zeitpunkt t in Klasse j zu wechseln, wenn man zum Zeitpunkt s in Klasse h ist. Durch die Intensitäten

$$q_{hj}(t) = \lim_{u \rightarrow 0^+} \frac{p_{hj}(t, t+u)}{u}$$

ist der Generator des Prozesses definiert. Bei einem homogenen Prozess sind diese konstant über die Zeit. Die Migrationsmatrizen können dann als eine Funktion

von $Q = (q_{hj})_{h,j=1,\dots,k}$ dargestellt werden. Im inhomogenen Fall ist dies durch die kumulativen Intensitäten $A_{hj}(t) = \int_0^t q_{hj}(s)ds$ möglich, vgl. Andersen et al. (1993). Für den absorbierenden Zustand k gilt, dass $p_{kj}(s, t) = q_{kj}(t) = 0$, $\forall s, t \in [0, T]$ mit $j \neq k$.

Möchte man nun die Migrationsmatrizen bzw. Intensitäten schätzen, verwendet man Zählprozesse. Wenn man Ratinghistorien von n Kontrahenten der Form $X_i = \{X_t^i, t \in [0, T]\}$, $i = 1, \dots, n$ hat, erhält man die nötige Information auch durch die Startratings X_0^1, \dots, X_0^n mit den Prozessen

$$N_{hj}(t) = \#\{s \in [0, t] | \exists i = 1, \dots, n \text{ mit } X_{s-}^i = h, X_s^i = j\}, t \in [0, T], j \neq h,$$

die die Anzahl der Migrationen im Portfolio von Rating h zu j bis zum Zeitpunkt t angeben. Zusätzlich werden die Prozesse $Y_h(t)$ verwendet. Diese bezeichnen die Anzahl der Kontrahenten in Rating h zum Zeitpunkt t . Außerdem betrachtet man meist den Intensitätsprozess

$$E(dN_{hj}(t) | \mathcal{F}_{t-}) = \lambda_{hj}(t) = q_{hj}(t)Y_h(t). \quad (2.1)$$

Für eine parametrische Schätzung ist es in diesem Zusammenhang am einfachsten, einen Maximum-Likelihood-Schätzer zu verwenden. Es kann hier jedoch nur eine partielle Likelihood gefunden werden (Andersen et al., 1993, Kapitel II.7.3), da meist Zensierungen vorliegen. Sie ergibt sich durch

$$L(q_{hj}(t)) = \prod_{t \in [0, T]} \prod_{j \neq h} (q_{hj}(t)Y_h(t))^{\Delta N_{hj}(t)} \exp \left(- \sum_{j \neq h} \int_0^T q_{hj}(t)Y_h(t)dt \right) \quad (2.2)$$

mit $\lambda(t)$ dem Vektor der $\lambda_{hj}(t)$, $h, j = 1, \dots, k, j \neq h$ und $\Delta N_{hj}(t) = N_{hj}(t) - N_{hj}(t^-) \in \{0, 1\}$. Diese sind genau dann eins, falls $N_{hj}(t)$ springt. $\prod_{j \neq h}$ entspricht dabei $\prod_{h=1}^{k-1} \prod_{\substack{j=1 \\ j \neq h}}^k$ bzw. $\sum_{j \neq h} = \sum_{h=1}^{k-1} \sum_{\substack{j=1 \\ j \neq h}}^k$. Hieraus folgt

$$\log(L) = \int_0^T \sum_{j \neq h} \log(\lambda_{hj}(t)) dN_{hj}(t) - \sum_{j \neq h} \int_0^T \lambda_{hj}(t) dt.$$

Für stationäre bzw. homogene Ratingverläufe ($q_{hj}(t) \equiv q_{hj}$) findet sich ein einfacher Maximum-Likelihood-Schätzer (ML-Schätzer) für die Intensität:

$$\hat{q}_{hj} = \frac{N_{hj}(T)}{\int_0^T Y_h(t) dt}.$$

Im Falle, dass $q_{h,h+1}(t) = q_{h,h-1}(t) = q$, $h = 1, \dots, k-1$ und alle anderen $q_{hj} = 0$, wenn also nur zwischen benachbarten Klassen migriert werden kann und alle Intensitäten gleich und konstant über die Zeit sind, ergibt sich der ML-Schätzer zu

$$\hat{q} = \frac{\sum_{j \neq h} N_{hj}(T)}{\int_0^T Y_1(t) dt + 2 \sum_{h=2}^{k-1} \int_0^T Y_h(t) dt}. \quad (2.3)$$

Möchte man nun modellieren, dass die Migrationen einzelner Individuen (Firmen) miteinander korrelieren, verwendet man hierfür Frailties. Dieses Prinzip wird im nächsten Abschnitt vorgestellt.

2.2. Frailties

Falls Ratingverläufe im jeweiligen Zeitverlauf stationär sind, die Intensitäten also über die Zeit konstant, ist es unerheblich, ob sie in Portfoliozeit (Zeit seit Eintritt in das Portfolio) oder in Kalenderzeit betrachtet werden. Der Maximum-Likelihood-Schätzer ist hier immer gleich. Möchte man diese Stationarität in Portfoliozeit testen, kann man dies z.B. wie in Weißbach und Walter (2010) tun. Bei der Betrachtung in Kalenderzeit kann von einer Stationarität kaum ausgegangen werden. Es ist einleuchtend, dass die Konjunktur - neben mikroökonomischen Kovariablen - einen

erheblichen Einfluss auf Migrations- und Ausfallwahrscheinlichkeiten hat und das Risiko einzelner Firmen über diese korreliert ist. Diese Konjunktur kann man auf zwei Arten in das Modell aufnehmen, einmal durch bestimmte (zeitdynamische) makroökonomische Kovariablen, wie z.B. dem Bruttoinlandsprodukt, oder durch einen latenten, unbeobachtbaren Prozess. Makroökonomische Variablen repräsentieren jedoch nicht die Konjunktur selbst, sondern werden nur von ihr beeinflusst. So scheint es sinnvoller, die Konjunktur als latente Variable aufzufassen.

In der Überlebenszeitanalyse werden solche latenten Variablen Frailties genannt. Sie können so modelliert werden, dass sie für ein einzelnes Individuum die Überlebenswahrscheinlichkeit (oder in unserem Fall die Ausfallwahrscheinlichkeit) erhöhen oder verringern. Die Ausfallintensität für ein Individuum l , in der Überlebenszeitanalyse meist Hazardrate genannt, sieht dann folgendermaßen aus:

$$q_l(t) = q^0(t)\psi_l,$$

wobei $q^0(t)$ die Basisintensität für alle Individuen, t in diesem Fall das Alter und ψ_l die Realisation einer unbeobachtbaren Zufallsvariablen Ψ ist. Sie wird als Frailty des Individuums i bezeichnet. In einem solchen Fall kann die Frailty etwa eine unentdeckte Veranlagung zu einer Krankheit oder ähnlichem repräsentieren. Es gilt, dass je größer die Frailty, desto kleiner die Überlebenswahrscheinlichkeit bzw. desto größer die Ausfallwahrscheinlichkeit. Für $\psi_l > 1$ erhöht sich die Intensität und damit die Ausfallwahrscheinlichkeit im Gegensatz zur Basisintensität $q_0(t)$, für $0 < \psi_l < 1$ wird die Intensität kleiner als die Basisintensität und für $\psi_l = 1$ entspricht sie dieser. Vor allem werden Frailties aber verwendet um Korrelationen zwischen den Überlebenszeiten einzelner Individuen zu modellieren. In diesem Fall sind die Realisationen von Ψ für verschiedene Untergruppen identisch, und man nennt sie auch „shared frailty“. Das Ziel ist, Ähnlichkeiten zwischen Gruppen, wie etwa Familienmitgliedern, abzubilden. Diese haben aufgrund gleicher Gene, die aber - genau wie die Konjunktur - nicht beobachtbar sind, zum Beispiel bei Krankheiten eine ähnliche

Überlebenswahrscheinlichkeit. Dies ist bei zeitdiskreten Paneldaten äquivalent zu einem Modell mit zufälligen Effekten, da auch solche als unbeobachtbare Zufallsvariablen den Zusammenhang innerhalb von Gruppen modellieren. Die Ausfallintensität für die Gruppe m hat dann die Form

$$q_m(t) = q^0(t)\psi_m,$$

wobei hier ψ_m die Frailty der Gruppe m bezeichnet.

Die Frailty sollte positiv sein und so eine entsprechende Verteilung besitzen, damit keine negative Intensität auftreten kann. Standardverteilungen sind dabei die Gamma-Verteilung oder eine Lognormalverteilung (vgl. Vaupel et al., 1979; Pickles und Crouchley, 1995). Die Gammaverteilung ist beliebt, weil sie nur positive Werte erzeugt und bei gleichen Parametern den Erwartungswert eins hat. Das bedeutet, dass die Basisintensität der Erwartungswert der Intensität ist. Die Standard-Lognormalverteilung hat den Vorteil, dass der Median eins beträgt und so jeweils die Hälfte der Wahrscheinlichkeitsmasse auf $(0, 1)$ bzw. $(1, \infty)$ liegt.

Die Likelihood bei gammaverteilten Frailties berechnen Nielsen et al. (1992), zur Berechnung der ML-Schätzer ist aber noch ein EM-Algorithmus notwendig. Murphy (1995) gelingt es, für diese Schätzer die Konsistenz und asymptotische Normalität zu zeigen.

Hougaard (1995) ist mit gammaverteilten Frailties nicht einverstanden, da vor allem die Annahme gleicher Parameter nicht realistisch sei. Er schlägt deswegen eine inverse Gaußverteilung oder Power Variance Function Exponentialfamilien vor. Manton et al. (1986) zeigen hier jedoch, dass die angenommene Verteilung der Frailty bei medizinischen Daten wesentlich weniger wichtig ist als das Modell der Basisintensität.

Zwischen den Gruppen korrelierte Gamma-Frailties behandelt unter anderem Parner (1998), der auch die Konsistenz und asymptotische Normalität der Schätzer bei zusätzlich vorliegenden Kovariablen zeigt.

2.3. Dynamische Frailties

Nun können die Frailties auch zeitabhängig (dynamisch) und über die Zeit korreliert sein. In der Interpretation als Konjunktur sind genau diese Eigenschaften gewünscht. Über die Zeit korrelierte Frailties für Poisson-Zählprozesse betrachten z.B. Henderson und Shimakura (2003), deren Frailties multivariat-gammaverteilt sind, und Fiocco et al. (2009), die sich mit gammaverteilten Frailties beschäftigen. Hasan et al. (2007) diskutieren exponentialverteilte, dynamische Frailties, Yue und Chan (1997) verwenden einen multiplikativen Random Walk, der betaverteilte Innovationen besitzt, Manda und Meyer (2005) und Yau und McGilchrist (1998) befassen sich jeweils mit log-AR(1)-Frailties.

Auch für den speziellen Fall der Ratingverläufe gibt es schon Vorschläge Frailties zusätzlich zu mikro- und makroökonomischen Variablen in die Migrationsintensitäten aufzunehmen. Diese soll genau die Eigenschaften der Zeitabhängigkeit und der Korrelation über die Zeit besitzen. Delloye et al. (2006) schlagen gammaverteilte Frailties vor. Die Intensität für Firma l hat hier die Form

$$q_{hjl}(t) = q_{hj}^0(t)\psi_{lt} \exp[\beta'_{hj}z_{hjl}(t)],$$

mit z_{hjl} als Vektor der zu Firma l gehörenden Kovariablen und

$$\begin{aligned}\psi_{lt} &= \psi_{l(t-1)}\tilde{\psi}_{lt}, \quad t = 2, \dots, \tau \\ \psi_{l1} &= \tilde{\psi}_{l1}, \\ \tilde{\psi}_{lt} &\stackrel{u.i.v.}{\sim} \text{Gamma}(\alpha, \alpha)\end{aligned}$$

als Frailtyprozess. Dabei gehen sie davon aus, dass die Frailties über ein Jahr konstant sind, τ gibt also die Anzahl der betrachteten Jahre an und entspricht so meist T . Es kann aber auch als Anzahl der Perioden, in denen sich die Frailty verändert, gesehen werden (z.B. wahlweise monatlich oder quartalsweise).

Hier liegt eine Mischung aus Cox-Modell und Modell mit zeitabhängiger Frailty vor. Die Frailty repräsentiert in diesem Fall weniger die Konjunktur als vielmehr un-

beobachtbare oder vergessene, firmenspezifische Kovariablen. Möchte man die volle Likelihood mit unbeobachtbaren Daten konstruieren, bekommt man, wie in allen folgenden Fällen, das Problem eines hochdimensionalen (τ -dimensionalen) Integrals. Genauer ergibt sich die Likelihood zu:

$$\begin{aligned} & L(q_{hjl}, \beta, \alpha) \\ &= E_\psi(L(q_{hjl}, \beta|\psi)) \\ &= C(\beta) E_\psi \left(\prod_{t=1}^{\tau} \prod_{l=1}^n \psi_{lt}^{\sum_{h \neq j} \Delta N_{hjl}(t)} \exp \left[- \sum_{l=1}^n \psi_{lt} \sum_{h \neq j} \int_{t-1}^t Y_{hl}(u) e^{\beta_{hj} + \beta'_{hjl} z_{hjl}(u)} du \right] \right), \end{aligned}$$

wobei $C(\beta)$ der Teil der Likelihood ist, der nur vom Parameter β abhängt. Im von den Autoren betrachteten Fall ist $\tau = 23$ und es existiert keine geschlossene Form der Likelihood. Eine explizite Darstellung der ML-Schätzer ist so nicht möglich. Deswegen verwenden die Autoren zur Schätzung von β eine Mischung aus EM-Algorithmus und MCMC Methoden. Das Hinzunehmen der Frailty verändert den Schätzer des Vektor β kaum, die Ausfallwahrscheinlichkeit von durch S&P als gut bewerteter Firmen (Investmentgrade) wird aber eher unterschätzt. Die Korrelationen zwischen Ausfällen, also die gegenseitige „Ausfallansteckung“, im Frailty-Modell entsprechen jedoch den von S&P historisch belegten.

Eine andere Verteilung für den latenten Faktor wählen dagegen Koopman et al. (2008). Mit den Migrationsintensitäten

$$q_{hjl}(t) = q_{hjl}^0(t) \exp [\eta_{hj} + \beta'_{hjl} z_l(t) + \alpha_{hj} \psi(t)]$$

und den unbekanntem Koeffizienten η_{hj} und α_{hj} sieht die Frailty $\psi(t) = \psi_i$, $t \in (t_{i-1}, t_i]$ folgendermaßen aus:

$$\begin{aligned} \psi_i &= \rho^{t_i - t_{i-1}} \psi_{i-1} + \sigma_i \varepsilon_i, \\ \varepsilon_i &\stackrel{u.i.v.}{\sim} N(0, 1), \\ \sigma_i^2 &= \begin{cases} (1 - \rho^{2(t_i - t_{i-1})}) / (1 - \rho^2), & \text{wenn } -1 < \rho < 1 \\ t_i - t_{i-1}, & \text{wenn } \rho = 1. \end{cases} \end{aligned}$$

Hierbei sind die $t_i, i = 1, \dots, \tau$ Zeitpunkte, zu denen ein Schuldner migriert, wobei τ der Gesamtzahl der Migrationen entspricht, also auch der Perioden, in denen sich die Frailty ändert. Der latente Prozess ist also ein AR(1)-Prozess, der sich für $-1 < \rho < 1$ einem Ornstein-Uhlenbeck-Prozess und für $\rho = 1$ einer Brownschen Bewegung annähert. Da hier eine Frailty für alle Firmen gilt, ist sie auch tatsächlich als Konjunktur zu interpretieren. Um mit dem Problem des hochdimensionalen Integrals umgehen zu können, formulieren die Autoren das Zählprozessmodell in ein State-Space-Modell um und schätzen es mit dort verwendeten MCMC-Methoden und dem Kalmanfilter. Simulationen zeigen, dass die Schätzer sehr nahe den wahren Werten mit relativ kleiner Standardabweichung sind. Der latente Prozess zeigt hier einen signifikanten Einfluss auf die Verweildauer in einer Ratingklasse, wobei der Einfluss auf Herabstufungen größer ist als auf Heraufstufungen.

Nicht einen AR(1)-Prozess, sondern direkt einen Ornstein-Uhlenbeck Prozess wählen Duffie et al. (2009). Sie betrachten nur die Wahrscheinlichkeit auszufallen. Die Intensitäten sehen dann folgendermaßen aus:

$$q_l(t) = \exp[\eta + \beta' z_l(t) + \alpha\psi(t)],$$

wobei z_l von den Autoren noch zusätzlich in firmenspezifische und generelle makroökonomische Variablen aufgeteilt wird. Sie wählen die Frailty als Ornstein-Uhlenbeck-Prozess wie folgt:

$$\begin{aligned} d\psi(t) &= -\kappa\psi(t)dt + dB(t), \\ \psi(0) &= 0, \end{aligned}$$

mit $B(t)$ einer Brown'schen Bewegung. Zur Bestimmung der Parameterschätzer findet eine Mischung aus EM-Algorithmus und Gibbs Sampler Verwendung. Der Vergleich des Modells mit und ohne Frailty zeigt, dass sie einen signifikanten Einfluss auf die Überlebenszeit hat. Die Parameterschätzer der gewählten Einflussvariablen und deren Signifikanzen ändern sich aber auch in diesem Fall kaum.

In allen drei Fällen sind komplizierte und umfangreiche numerische Verfahren nötig

um Schätzer berechnen zu können. Das folgende Kapitel beschäftigt sich mit einem alternativen Weg, die Likelihood für ein Modell mit zeitdynamischer Frailty zu berechnen.

3. Likelihood für einen Zählprozess mit dynamischer Frailty

Dieses Kapitel beschäftigt sich mit zwei alternativen Wegen eine Frailty zu modellieren. Die im vorigen Kapitel vorgestellten Modelle benötigen alle mehrschrittige numerische Verfahren, die Kombinationen aus zwei oder mehr Algorithmen und sehr programmier- und rechenaufwändig sind. In diesem Kapitel wird mit Hilfe der Laplace-Approximation jeweils eine Likelihood konstruiert. Zur Schätzung benötigt man dann nur noch einen einfachen Newton-Raphson Algorithmus oder das Verfahren nach Nelder und Mead. Die Laplace-Approximation wird zuerst vorgestellt. Der folgende Abschnitt zieht eine betaverteilte Frailty in Betracht, die jedoch keinerlei autoregressive Eigenschaften hat. Deswegen wird später auch auf ein AR(1)-Prozess als Frailty zurückgegriffen, dieser kann jedoch negative Intensitäten verursachen.

3.1. Laplace-Approximation

Möchte man zeitvariable bzw. dynamische Frailties betrachten, so ergibt sich, wie schon im vorigen Kapitel ausgeführt, die Likelihood zu einem hochdimensionalen Integral. Hat dieses Integral eine bestimmte Form, lässt es sich mit Hilfe einer Laplace-Approximation (vgl. Butler, 2007, S. 86) annähern. Hierbei ist vor allem ein Exponentialterm wichtig.

Satz 3.1 *Es gilt:*

$$\int_M h(x)e^{-g(x)} dx \approx \frac{2\pi^{m/2}}{\sqrt{|H_g(\tilde{x})|}} h(\tilde{x})e^{-g(\tilde{x})}, \quad (3.1)$$

wobei $M \subset \mathbb{R}^m$, $g(x) > 0 \forall x \in M$ stetig differenzierbar, $h(x) > 0 \forall x \in M$ stetig

differenzierbar, $\tilde{x} := \operatorname{argmin}_{x \in M} g(x)$ und $|H_g(\cdot)|$ die Determinante der Hessematrix zur Funktion $g(\cdot)$ ist.

Der Beweis hierzu ist einfach mit einer Taylorentwicklung von $g(x)$ um \tilde{x} zu führen. Der Korrekturterm $\frac{2\pi^{m/2}}{\sqrt{|H_g(\tilde{x})|}}$ entsteht dann durch ein normalverteilungsähnliches Integral. Offensichtlich bestimmt vor allem das Minimum von $g(\cdot)$, in dem dann das Maximum von $e^{-g(\cdot)}$ liegt, das Integral. Wie exakt diese Approximation ist, hängt also vor allem davon ab, wie „quadratisch“ $g(\cdot)$ in der Umgebung von \tilde{x} und wie spitz $e^{-g(\cdot)}$ an der Stelle \tilde{x} ist. Außerdem wird davon ausgegangen, dass die größte Masse des Integrals innerhalb der Menge M und nicht an ihren Rändern liegt, das bedeutet insbesondere, dass Randmaxima nicht zulässig sind. Diese multivariate Laplace-Approximation geht auf Hsu (1948) zurück.

Die univariate Laplace-Approximation wurde z. B. von Daniels (1954), Barndorff-Nielsen und Cox (1979) auf statistische und ökonometrische Problemen angewendet, genauer entwickeln sie eine verallgemeinerte Edgeworth-Entwicklung. Tierney und Kadane (1986) und Efstathiou et al. (1998) berechnen mit ihrer Hilfe a posteriori Momente und Verteilungen. Eine erste statistische Anwendungen der multivariaten Laplace-Approximation nahmen Glynn (1980) vor. Wolfinger (1993) hat sie für lineare Modelle mit zufälligen Effekten verwendet. Neuere Anwendungen stammen von Butler und Wood (2002), Harding und Hausman (2007) und Gagliardini und Gouriéroux (2008).

Damit diese Approximation für Zählprozesse mit latenten Variablen anwendbar ist, muss die Dichte der latenten Variable über einen Exponentialterm mit negativem Exponenten verfügen. Dieser darf zusätzlich auch nichtlinear sein. Diese Forderungen werden anhand einer einfachen gammaverteilten Frailty erläutert.

Der Intensitätsprozess der Zählprozesse sei gegeben durch

$$\lambda_{hj}(t|\psi) = \sum_{i=0}^{\tau} \mathbb{1}_{(t_i, t_{i+1}]}(t) \psi_i q_{hj}(t) Y_h(t).$$

Die Zeitpunkte t_i sind hierbei Zeitpunkte, in denen ein Schuldner migriert, also gilt $\Delta N_{hj}(t_i) = N_{hj}(t_i) - N_{hj}(t_i-) = 1$. Zusätzlich ist $t_0 = 0$ und $t_{\tau+1} = T$. Es gilt außerdem

$$\psi_i \stackrel{u.i.v.}{\sim} \text{Gamma}(\alpha, \alpha).$$

Die Frailty ändert sich also zu jedem Zeitpunkt, in dem ein Schuldner zwischen zwei beliebigen Klassen migriert. Die einzelnen Ausprägungen sind zunächst unabhängig voneinander.

Dann hat die Likelihood, gegeben ψ , die Form

$$L(q|\psi) = \prod_{i=0}^{\tau} \prod_{h \neq j} (q_{hj}(t) \psi_i Y_h(t_i))^{\Delta N_{hj}(t)} e^{-\sum_{i=0}^{\tau} \sum_{h \neq j} q_{hj}(t) \psi_i \int_{t_i}^{t_{i+1}} Y_h(t) dt}.$$

Die gesuchte Likelihood ergibt sich so zu

$$\begin{aligned} & L(q_{hj}(t), \alpha) \\ &= \int_{\mathbb{R}_+^{\tau+1}} \prod_{i=0}^{\tau} \prod_{h \neq j} (q_{hj}(t) \psi_i Y_h(t_i))^{\Delta N_{hj}(t)} e^{-\sum_{i=0}^{\tau} \sum_{h \neq j} q_{hj}(t) \psi_i \int_{t_i}^{t_{i+1}} Y_h(t) dt} \\ & \quad \prod_{i=0}^{\tau} \left(\frac{\psi_i^{\alpha-1}}{\alpha^\alpha \Gamma(\alpha)} e^{-\frac{\psi_i}{\alpha}} \right) d\psi \\ &= \int_{\mathbb{R}_+^{\tau+1}} \prod_{i=0}^{\tau} \frac{\psi_i^{\alpha-1}}{\alpha^\alpha \Gamma(\alpha)} \prod_{h \neq j} (q_{hj}(t) \psi_i Y_h(t_i))^{\Delta N_{hj}(t)} e^{-\sum_{i=0}^{\tau} (\frac{\psi_i}{\alpha} + \sum_{h \neq j} q_{hj}(t) \psi_i \int_{t_i}^{t_{i+1}} Y_h(t) dt)} d\psi. \end{aligned}$$

Soll nun eine Laplace-Approximation angewendet werden um das hochdimensionale Integral aufzulösen, erhalten wir mit der Notation wie zuvor

$$h(\psi) = \prod_{i=0}^{\tau} \frac{\psi_i^{\alpha-1}}{\beta^\alpha \Gamma(\alpha)} \prod_{h \neq j} (q_{hj}(t) \psi_i Y_h(t_i))^{\Delta N_{hj}(t)}$$

und

$$g(\psi) = \sum_{i=0}^{\tau} \left(\sum_{h \neq j} q_{hj}(t) \psi_i \int_{t_i}^{t_{i+1}} Y_h(t) dt + \frac{\psi_i}{\alpha} \right).$$

Offensichtlich ist $g(\cdot)$ linear in ψ und somit existiert nur am Rand der Menge M

(hier: $\mathbb{R}_+^{\tau+1}$) ein Minimum. Dieses ist aber für die Approximation nicht geeignet, somit kann sie nicht angewendet werden.

Da der aus der Likelihood des Zählprozesses stammende Teil der Funktion $g(\cdot)$ immer linear in ψ ist, ist es notwendig, dass dies für den aus der Verteilung des latenten Faktors stammenden Teil nicht gilt. Idealerweise sollte er quadratisch sein, um eine möglichst genaue Approximation zu erhalten.

3.2. Betaverteilte Frailty

Das Ziel ist es jetzt, ohne die Hilfe numerischer Methoden eine Likelihood mit dynamischer Frailty herzuleiten. Dabei werden keine weiteren mikro- oder makroökonomischen Variablen berücksichtigt, da die Basisintensität bei einer Cox-Regression unabhängig geschätzt wird. Außerdem werden die jeweiligen Basisintensitäten als konstant über die Zeit angenommen. Das geschieht vor allem um das Vorgehen in einem möglichst einfachen Fall zu überprüfen.

Die Bedingung an eine Frailty ist, dass sie möglichst um den Wert eins schwankt. Um der Interpretation als Konjunktur gerecht zu werden, soll in der Modellierung zusätzlich für die Heraufstufung ein positiver, für die Herabstufung ein negativer Zusammenhang mit der Frailty angenommen werden. So bietet es sich zum Beispiel an, auf $(-1, 1)$ betaverteilte Variablen zu verwenden.

Die Intensitäten der Zählprozesse bei gegebenen Beobachtungen der Frailty haben dann die Form

$$q_{hj}(t|\psi) = \begin{cases} q \sum_{i=0}^{\tau} \mathbf{1}_{(t_i, t_{i+1}]}(t)(1 + \psi_i), & \text{für } h - j = -1 \\ q \sum_{i=0}^{\tau} \mathbf{1}_{(t_i, t_{i+1}]}(t)(1 - \psi_i), & \text{für } h - j = +1 \\ 0, & \text{sonst.} \end{cases}$$

Die Frailty ψ_i soll für $i = 0, \dots, \tau$ die folgende Dichte besitzen:

$$\begin{aligned} f(\psi_i) &= \mathbb{1}_{(-1,1)}(\psi_i) \frac{\Gamma(2\alpha)}{2\Gamma(\alpha)^2} \left(\frac{\psi_i + 1}{2}\right)^{\alpha-1} \left(\frac{-\psi_i + 1}{2}\right)^{\alpha-1} \\ &= \mathbb{1}_{(-1,1)}(\psi_i) \frac{\Gamma(2\alpha)}{2\Gamma(\alpha)^2} \exp \left[(\alpha - 1) \left(\log \left(\frac{\psi_i + 1}{2} \right) + \log \left(\frac{-\psi_i + 1}{2} \right) \right) \right]. \end{aligned}$$

Die beiden Parameter der verallgemeinerten Beta-Verteilung werden gleich gesetzt um eine Symmetrie um Null zu erzeugen. Außerdem soll gelten, dass $\alpha > 1$ ist, damit eine normalverteilungähnliche Form der Dichte erreicht und nicht der Großteil der Wahrscheinlichkeitsmasse auf die Ränder verteilt ist.

Bei vielen verschiedenen Ratingklassen enthält das Modell sehr viele zu schätzende Parameter. Es wird hier zur Modellierung davon ausgegangen, dass nur Migrationen zwischen benachbarten Ratings möglich sind, wie auch in Delloye et al. (2006) betrachtet. Dies scheint eine logische Annahme, da sich ein Rating eigentlich langsam und schrittweise verbessert bzw. verschlechtert. Alle Sprünge über mehr als eine Ratingklasse können so als Fehler in den Daten angesehen werden. Eine zusätzliche Annahme ist, dass die Basisintensitäten für alle Migrationen gleich sind. Weißbach und Mollenhauer (2011) zeigen, dass dies auch sinnvoll ist.

Geht man wieder davon aus, dass Beobachtungen von ψ_i gegeben sind, ergibt sich die partielle Likelihood mit (2.2) zu

$$\begin{aligned} &L(q|\psi) \\ &= \prod_{i=0}^{\tau} \prod_{h=j+1} [q(1 + \psi_i)Y_h(t_i)]^{\Delta N_{hj}(t_{i+1})} \prod_{h=j-1} [q(1 - \psi_i)Y_h(t_i)]^{\Delta N_{hj}(t_{i+1})} \\ &\exp \left[- \left(\sum_{i=0}^{\tau} q(1 + \psi_i) \sum_{h=j+1} \int_{t_i}^{t_{i+1}} Y_h(t) dt + \sum_{i=0}^{\tau} q(1 - \psi_i) \sum_{h=j-1} \int_{t_i}^{t_{i+1}} Y_h(t) dt \right) \right]. \end{aligned}$$

Die Summe $\sum_{h=j+1}$ steht für $\sum_{h=1, j=h-1}^{k-1}$, äquivalentes gilt bei $\sum_{h=j-1}, \prod_{h=j+1}$ und $\prod_{h=j-1}$.

Da hier die Frailty ψ aber unbekannt und zufällig ist, benötigen wir die marginale

partielle Likelihood:

$$\begin{aligned}
L(q, \alpha) &= \int_{\mathbb{R}^{\tau+1}} L(q \mid N_{hj}(t), Y_h(t), \psi) f_\alpha(\psi) d\psi \\
&= \frac{\Gamma(2\alpha)^{\tau+1}}{2^{\tau+1} \Gamma(\alpha)^{2(\tau+1)}} \\
&\quad \int_{(-1,1)^{\tau+1}} \prod_{i=0}^{\tau} \prod_{h=j+1}^{\tau} [q(1 + \psi_i) Y_h(t_i)]^{\Delta N_{hj}(t_{i+1})} \prod_{h=j-1}^{\tau} [q(1 - \psi_i) Y_h(t_i)]^{\Delta N_{hj}(t_{i+1})} \\
&\quad \exp \left[- \left(\sum_{i=0}^{\tau} q(1 + \psi_i) \sum_{h=j+1}^{\tau} \int_{t_i}^{t_{i+1}} Y_h(t) dt \right. \right. \\
&\quad \left. \left. + \sum_{i=0}^{\tau} q(1 - \psi_i) \sum_{h=j-1}^{\tau} \int_{t_i}^{t_{i+1}} Y_h(t) dt \right. \right. \\
&\quad \left. \left. - \sum_{i=0}^{\tau} (\alpha - 1) \left(\log \left(\frac{\psi_i + 1}{2} \right) + \log \left(\frac{-\psi_i + 1}{2} \right) \right) \right] d\psi.
\end{aligned}$$

Diese ist nach Gill (1992) bei auf ψ bedingter nichtinformativer Zensierung äquivalent zur partiellen marginalen Likelihood. Bei der partiellen marginalen Likelihood handelt es sich um die Likelihood für die q_{hj} , gegeben den $N_{hj}(t)$, im „marginalen“ Experiment bzw. Experiment mit fehlenden Daten. Je nach Größe von n und T bzw. τ liegt bei der marginalen partiellen Likelihood ein hochdimensionales Integral vor, das mit Hilfe einer Laplace-Approximation angenähert werden kann, vgl. Abschnitt 3.1.

Die in unserem Fall zu minimierende Funktion sieht folgendermaßen aus:

$$\begin{aligned}
g(\psi) &= \sum_{i=0}^{\tau} q(1 + \psi_i) \sum_{h=j+1}^{\tau} \int_{t_i}^{t_{i+1}} Y_h(t) dt + \sum_{i=0}^{\tau} q(1 - \psi_i) \sum_{h=j-1}^{\tau} \int_{t_i}^{t_{i+1}} Y_h(t) dt \\
&\quad - \sum_{i=0}^{\tau} (\alpha - 1) \left(\log \left(\frac{\psi_i + 1}{2} \right) + \log \left(\frac{-\psi_i + 1}{2} \right) \right).
\end{aligned}$$

Die Funktion $g(\psi)$ ist auf $(-1, 1)$ positiv, da $\log \left(\frac{\psi_i + 1}{2} \right) < 0$ bzw. $\log \left(\frac{-\psi_i + 1}{2} \right) < 0$.

Die interessierenden Ableitungen ergeben sich also folgendermaßen:

$$\frac{\partial g(\psi)}{\partial \psi_i} = q \left(\sum_{h=j+1} \int_{t_i}^{t_{i+1}} Y_h(t) dt - \sum_{h=j-1} \int_{t_i}^{t_{i+1}} Y_h(t) dt \right) - (\alpha - 1) \left[\frac{1}{\psi_i - 1} + \frac{1}{\psi_i + 1} \right]$$

$$\frac{\partial^2 g(\psi)}{(\partial \psi_i)^2} = (\alpha - 1) \left[\frac{1}{(\psi_i - 1)^2} + \frac{1}{(\psi_i + 1)^2} \right] > 0, \quad i = 0, \dots, \tau$$

$$\frac{\partial^2 g(\psi)}{\partial \psi_i \partial \psi_l} = 0, \quad i \neq l \in \{0, \dots, \tau\}.$$

Damit ist die Hessematrix positiv definit und alle Nullstellen der 1. Ableitung sind Minimalstellen. Es gibt zwei dieser Minimalstellen:

$$\hat{\psi}_i^{(1,2)} = -\frac{\alpha - 1}{q \int_{t_i}^{t_{i+1}} Y_1(t) dt} \pm \sqrt{\frac{(\alpha - 1)^2}{q^2 (\int_{t_i}^{t_{i+1}} Y_1(t) dt)^2} + 1},$$

wobei nur eine im Intervall $(-1, 1)$ liegen kann. Da für $x > 0$ gilt

$$|-x - \sqrt{x^2 + 1}| > 1,$$

ist unsere gesuchte Nullstelle

$$\hat{\psi}_i = -\frac{\alpha - 1}{q \int_{t_i}^{t_{i+1}} Y_1(t) dt} + \sqrt{\frac{(\alpha - 1)^2}{q^2 (\int_{t_i}^{t_{i+1}} Y_1(t) dt)^2} + 1}. \quad (3.2)$$

Die Likelihood bzw. log-Likelihood sehen dann mit einer Laplace-Approximation

durch Einsetzen von (3.2) wie folgt aus:

$$\begin{aligned}
& L(q, \alpha) \\
& \approx \frac{\Gamma(2\alpha)^{\tau+1}}{2^{\tau+1}\Gamma(\alpha)^{2(\tau+1)}} \frac{\sqrt{2\pi^{\tau+1}}}{\sqrt{\prod_{i=0}^{\tau} (\alpha-1) \left(\frac{1}{(\hat{\psi}_i-1)^2} + \frac{1}{(\hat{\psi}_i+1)^2} \right)}} \\
& \prod_{i=0}^{\tau} \prod_{h=j+1} \left[q \left(1 - \frac{\alpha-1}{q \int_{t_i}^{t_{i+1}} Y_1(t) dt} + \sqrt{\frac{(\alpha-1)^2}{q^2 (\int_{t_i}^{t_{i+1}} Y_1(t) dt)^2} + 1}} \right) Y_h(t_i) \right]^{\Delta N_{hj}(t_{i+1})} \\
& \prod_{h=j-1} \left[q \left(1 + \frac{\alpha-1}{q \int_{t_i}^{t_{i+1}} Y_1(t) dt} - \sqrt{\frac{(\alpha-1)^2}{q^2 (\int_{t_i}^{t_{i+1}} Y_1(t) dt)^2} + 1}} \right) Y_h(t_i) \right]^{\Delta N_{hj}(t_{i+1})} \\
& \exp \left[- \left(\sum_{i=0}^{\tau} q \left(1 - \frac{\alpha-1}{q \int_{t_i}^{t_{i+1}} Y_1(t) dt} + \sqrt{\frac{(\alpha-1)^2}{q^2 (\int_{t_i}^{t_{i+1}} Y_1(t) dt)^2} + 1}} \right) \sum_{h=j+1} \int_{t_i}^{t_{i+1}} Y_h(t) dt \right. \right. \\
& \left. \left. + \sum_{i=0}^{\tau} q \left(1 + \frac{\alpha-1}{q \int_{t_i}^{t_{i+1}} Y_1(t) dt} - \sqrt{\frac{(\alpha-1)^2}{q^2 (\int_{t_i}^{t_{i+1}} Y_1(t) dt)^2} + 1}} \right) \sum_{h=j-1} \int_{t_i}^{t_{i+1}} Y_h(t) dt \right) \right] \\
& \left(\frac{1 - \frac{\alpha-1}{q \int_{t_i}^{t_{i+1}} Y_1(t) dt} + \sqrt{\frac{(\alpha-1)^2}{q^2 (\int_{t_i}^{t_{i+1}} Y_1(t) dt)^2} + 1}}{2} \right)^{\alpha-1} \\
& \left(\frac{1 + \frac{\alpha-1}{q \int_{t_i}^{t_{i+1}} Y_1(t) dt} - \sqrt{\frac{(\alpha-1)^2}{q^2 (\int_{t_i}^{t_{i+1}} Y_1(t) dt)^2} + 1}}{2} \right)^{\alpha-1}
\end{aligned}$$

und

$$\begin{aligned}
& l(q, \alpha) = \log(L(q, \alpha)) \\
& = (\tau+1) \log(\Gamma(2\alpha)) - \log(2^{\tau+1}) - 2(\tau+1) \log(\Gamma(\alpha)) + (\tau+1) \log(\sqrt{2\pi}) \\
& - (\tau+1) \log(\sqrt{\alpha-1}) - \sum_{i=0}^{\tau} \log \left(\sqrt{\frac{1}{(\hat{\psi}_i-1)^2} + \frac{1}{(\hat{\psi}_i+1)^2}} \right) \\
& + \sum_{i=0}^{\tau} (N_+(t_{i+1}) - N_+(t_i)) \log \left(1 - \frac{\alpha-1}{q \int_{t_i}^{t_{i+1}} Y_1(t) dt} + \sqrt{\frac{(\alpha-1)^2}{q^2 (\int_{t_i}^{t_{i+1}} Y_1(t) dt)^2} + 1}} \right) \\
& + \sum_{i=0}^{\tau} (N_-(t_{i+1}) - N_-(t_i)) \log \left(1 + \frac{\alpha-1}{q \int_{t_i}^{t_{i+1}} Y_1(t) dt} - \sqrt{\frac{(\alpha-1)^2}{q^2 (\int_{t_i}^{t_{i+1}} Y_1(t) dt)^2} + 1}} \right)
\end{aligned}$$

$$\begin{aligned}
& + N_{\bullet}(T) \log(q) + \sum_{i=0}^{\tau} \sum_{h \neq j} \log(Y_h(t_i)^{\Delta N_{hj}(t_{i+1})}) \\
& - \sum_{i=0}^{\tau} q \left(1 - \frac{\alpha - 1}{q \int_{t_i}^{t_{i+1}} Y_1(t) dt} + \sqrt{\frac{(\alpha - 1)^2}{q^2 (\int_{t_i}^{t_{i+1}} Y_1(t) dt)^2} + 1} \right) \sum_{h=j+1} \int_{t_i}^{t_{i+1}} Y_h(t) dt \\
& - \sum_{i=0}^{\tau} q \left(1 + \frac{\alpha - 1}{q \int_{t_i}^{t_{i+1}} Y_1(t) dt} - \sqrt{\frac{(\alpha - 1)^2}{q^2 (\int_{t_i}^{t_{i+1}} Y_1(t) dt)^2} + 1} \right) \sum_{h=j-1} \int_{t_i}^{t_{i+1}} Y_h(t) dt \\
& + (\alpha - 1) \left[\log \left(1 - \frac{\alpha - 1}{q \int_{t_i}^{t_{i+1}} Y_1(t) dt} + \sqrt{\frac{(\alpha - 1)^2}{q^2 (\int_{t_i}^{t_{i+1}} Y_1(t) dt)^2} + 1} \right) \right. \\
& \left. + \log \left(1 + \frac{\alpha - 1}{q \int_{t_i}^{t_{i+1}} Y_1(t) dt} - \sqrt{\frac{(\alpha - 1)^2}{q^2 (\int_{t_i}^{t_{i+1}} Y_1(t) dt)^2} + 1} \right) - 2 \log(2) \right],
\end{aligned}$$

wobei $N_+(t) = \sum_{h=2, j=h+1}^{k-1} N_{hj}(t)$, $N_-(t) = \sum_{h=1, j=h-1}^{k-1} N_{hj}(t)$ und $N_{\bullet}(t) = N_+(t) + N_-(t)$.

Um in diesem Fall schätzen zu können benötigt man weiterhin numerische Methoden, da keine geschlossene Darstellung der Schätzer möglich ist. Um ein Maximum zu finden bieten sich numerische Verfahren an, die im Abschnitt 4.1 beschrieben werden. Für diese werden die ersten und zweiten Ableitungen der log-Likelihood benötigt. Mit

$$\begin{aligned}
a_i &= \frac{(\alpha - 1)}{q^2 \int_{t_i}^{t_{i+1}} Y_1(t) dt} - \frac{(\alpha - 1)^2}{q^3 (\int_{t_i}^{t_{i+1}} Y_1(t) dt)^2 \sqrt{\frac{(\alpha - 1)^2}{q^2 (\int_{t_i}^{t_{i+1}} Y_1(t) dt)^2} + 1}}, \\
b_i &= 1 + \sqrt{\frac{(\alpha - 1)^2}{q^2 (\int_{t_i}^{t_{i+1}} Y_1(t) dt)^2} + 1} - \frac{(\alpha - 1)}{q \int_{t_i}^{t_{i+1}} Y_1(t) dt}, \\
c_i &= 1 - \sqrt{\frac{(\alpha - 1)^2}{q^2 (\int_{t_i}^{t_{i+1}} Y_1(t) dt)^2} + 1} + \frac{(\alpha - 1)}{q \int_{t_i}^{t_{i+1}} Y_1(t) dt}
\end{aligned}$$

ist

$$\begin{aligned}
\frac{\partial l(q, \alpha)}{\partial q} &= \sum_{i=0}^{\tau} \left(\frac{2a_i}{b_i^3} - \frac{2a_i}{c_i^3} \right) \left(\frac{2}{b_i^2} + \frac{2}{c_i^2} \right) \\
&+ \sum_{i=0}^{\tau} \left(\frac{N_+(t_{i+1}) - N_+(t_i)}{b_i} - \frac{(N_-(t_{i+1}) - N_-(t_i))}{c_i} \right) a_i
\end{aligned}$$

$$\begin{aligned}
& + \frac{N_{\bullet}(T)}{q} - \sum_{i=0}^{\tau} q a_i \left(\sum_{h=j+1} \int_{t_i}^{t_{i+1}} Y_h(t) dt - \sum_{h=j-1} \int_{t_i}^{t_{i+1}} Y_h(t) dt \right) \\
& - b_i \sum_{h=j+1} \int_{t_i}^{t_{i+1}} Y_h(t) dt - c_i \sum_{h=j-1} \int_{t_i}^{t_{i+1}} Y_h(t) dt + (\alpha - 1) \sum_{i=0}^{\tau} \left(\frac{a_i}{b_i} - \frac{a_i}{c_i} \right)
\end{aligned}$$

und mit

$$d_i = \frac{1}{q \int_{t_i}^{t_{i+1}} Y_1(t) dt} - \frac{(\alpha - 1)}{q^2 \left(\int_{t_i}^{t_{i+1}} Y_1(t) dt \right)^2 \sqrt{\frac{(\alpha-1)^2}{q^2 \left(\int_{t_i}^{t_{i+1}} Y_1(t) dt \right)^2} + 1}}$$

ergibt sich

$$\begin{aligned}
\frac{\partial l(q, \alpha)}{\partial \alpha} &= -1.3863 + (\tau + 1)\Phi(2\alpha) - 2(\tau + 1)\Phi(\alpha) - \frac{\tau + 1}{2(\alpha - 1)} \\
& + \sum_{i=0}^{\tau} \left(\frac{2d_i}{c_i^3} - \frac{2d_i}{b_i^3} \right) \left(\frac{2}{b_i^2} + \frac{2}{c_i^2} \right) \\
& - \sum_{i=0}^{\tau} \left(\frac{N_+(t_{i+1}) - N_+(t_i)}{b_i} - \frac{(N_-(t_{i+1}) - N_-(t_i))}{c_i} \right) d_i \\
& + \sum_{i=0}^{\tau} q d_i \left(\sum_{h=j+1} \int_{t_i}^{t_{i+1}} Y_h(t) dt - \sum_{h=j-1} \int_{t_i}^{t_{i+1}} Y_h(t) dt \right) \\
& + \sum_{i=0}^{\tau} (\log(b_i) + \log(c_i)) + (\alpha - 1) \sum_{i=0}^{\tau} \left(\frac{d_i}{c_i} - \frac{d_i}{b_i} \right),
\end{aligned}$$

wobei Φ die Digammafunktion ist. Die zweiten Ableitungen ergeben sich mit

$$\begin{aligned}
e_i &= \frac{2(\alpha - 1)}{q^3 \int_{t_i}^{t_{i+1}} Y_1(t) dt} - \frac{3(\alpha - 1)^2}{q^4 \left(\int_{t_i}^{t_{i+1}} Y_1(t) dt \right)^2 \sqrt{\frac{(\alpha-1)^2}{q^2 \left(\int_{t_i}^{t_{i+1}} Y_1(t) dt \right)^2} + 1}} \\
& + \frac{(\alpha - 1)^4}{q^6 \left(\int_{t_i}^{t_{i+1}} Y_1(t) dt \right)^4 \left(\frac{(\alpha-1)^2}{q^2 \left(\int_{t_i}^{t_{i+1}} Y_1(t) dt \right)^2} + 1 \right)^{3/2}}
\end{aligned}$$

zu

$$\begin{aligned}
\frac{\partial^2 l(q, \alpha)}{(\partial q)^2} &= \sum_{i=0}^{\tau} \partial^2 \log \left(\frac{1}{(\hat{\psi} - 1)^2} + \frac{1}{(\hat{\psi} + 1)^2} \right) / (\partial q)^2 \\
&\quad - \sum_{i=0}^{\tau} \left(\frac{N_+(t_{i+1}) - N_+(t_i)}{b_i^2} + \frac{(N_-(t_{i+1}) - N_-(t_i))}{c_i^2} \right) a_i^2 \\
&\quad - \sum_{i=0}^{\tau} \left(\frac{N_+(t_{i+1}) - N_+(t_i)}{b_i} - \frac{(N_-(t_{i+1}) - N_-(t_i))}{c_i} \right) e_i \\
&\quad - \frac{N_{\bullet}(T)}{q^2} + \sum_{i=0}^{\tau} q e_i \left(\sum_{h=j+1}^{\tau} \int_{t_i}^{t_{i+1}} Y_h(t) dt - \sum_{h=j-1}^{\tau} \int_{t_i}^{t_{i+1}} Y_h(t) dt \right) \\
&\quad - 2 \sum_{i=0}^{\tau} a_i \left(\sum_{h=j+1}^{\tau} \int_{t_i}^{t_{i+1}} Y_h(t) dt - \sum_{h=j-1}^{\tau} \int_{t_i}^{t_{i+1}} Y_h(t) dt \right) \\
&\quad - (\alpha - 1) \sum_{i=0}^{\tau} \left(\frac{a_i^2}{b_i^2} + \frac{a_i^2}{c_i^2} + \frac{e_i}{b_i} - \frac{e_i}{c_i} \right),
\end{aligned}$$

mit Φ' der Trigammafunktion und

$$\begin{aligned}
f_i &= \frac{1}{q^2 \left(\int_{t_i}^{t_{i+1}} Y_1(t) dt \right)^2 \sqrt{\frac{(\alpha-1)^2}{q^2 \left(\int_{t_i}^{t_{i+1}} Y_1(t) dt \right)^2} + 1}} \\
&\quad - \frac{(\alpha - 1)^2}{q^4 \left(\int_{t_i}^{t_{i+1}} Y_1(t) dt \right)^4 \left(\frac{(\alpha-1)^2}{q^2 \left(\int_{t_i}^{t_{i+1}} Y_1(t) dt \right)^2} + 1 \right)^{3/2}}
\end{aligned}$$

erhalten wir

$$\begin{aligned}
\frac{\partial^2 l(q, \alpha)}{(\partial \alpha)^2} &= (\tau + 1) \Phi'(2\alpha) - 2(\tau + 1) \Phi'(\alpha) + \frac{\tau + 1}{2(\alpha - 1)^2} \\
&\quad - \sum_{i=0}^{\tau} \partial^2 \log \left(\frac{1}{(\hat{\psi} - 1)^2} + \frac{1}{(\hat{\psi} + 1)^2} \right) / (\partial \alpha)^2 \\
&\quad - \sum_{i=0}^{\tau} \left(\frac{N_+(t_{i+1}) - N_+(t_i)}{b_i^2} + \frac{(N_-(t_{i+1}) - N_-(t_i))}{c_i^2} \right) d_i^2 \\
&\quad + \sum_{i=0}^{\tau} \left(\frac{N_+(t_{i+1}) - N_+(t_i)}{b_i} - \frac{(N_-(t_{i+1}) - N_-(t_i))}{c_i} \right) f_i
\end{aligned}$$

$$\begin{aligned}
& - \sum_{i=0}^{\tau} q f_i \left(\sum_{h=j+1} \int_{t_i}^{t_{i+1}} Y_h(t) dt - \sum_{h=j-1} \int_{t_i}^{t_{i+1}} Y_h(t) dt \right) \\
& + 2 \sum_{i=0}^{\tau} \left(\frac{d_i}{c_i} - \frac{d_i}{b_i} \right) - (\alpha - 1) \sum_{i=0}^{\tau} \left(\frac{d_i^2}{b_i^2} - \frac{d_i^2}{c_i^2} - \frac{f_i}{b_i} + \frac{f_i}{c_i} \right)
\end{aligned}$$

und mit

$$\begin{aligned}
g_i &= \frac{1}{q^2 \int_{t_i}^{t_{i+1}} Y_1(t) dt} - \frac{2(\alpha - 1)}{q^3 \left(\int_{t_i}^{t_{i+1}} Y_1(t) dt \right)^2 \sqrt{\frac{(\alpha-1)^2}{q^2 \left(\int_{t_i}^{t_{i+1}} Y_1(t) dt \right)^2} + 1}} \\
& + \frac{(\alpha - 1)^3}{q^5 \left(\int_{t_i}^{t_{i+1}} Y_1(t) dt \right)^4 \left(\frac{(\alpha-1)^2}{q^2 \left(\int_{t_i}^{t_{i+1}} Y_1(t) dt \right)^2} + 1 \right)^{3/2}}
\end{aligned}$$

ist schließlich

$$\begin{aligned}
\frac{\partial^2 l(q, \alpha)}{\partial \alpha \partial q} &= - \sum_{i=0}^{\tau} \partial^2 \log \left(\frac{1}{(\hat{\psi} - 1)^2} + \frac{1}{(\hat{\psi} + 1)^2} \right) / (\partial \alpha \partial q) \\
& + \sum_{i=0}^{\tau} \left(\frac{N_+(t_{i+1}) - N_+(t_i)}{b_i^2} + \frac{(N_-(t_{i+1}) - N_-(t_i))}{c_i^2} \right) a_i d_i \\
& + \sum_{i=0}^{\tau} \left(\frac{N_+(t_{i+1}) - N_+(t_i)}{b_i} + \frac{(N_-(t_{i+1}) - N_-(t_i))}{c_i} \right) g_i \\
& - \sum_{i=0}^{\tau} q g_i \left(\sum_{h=j+1} \int_{t_i}^{t_{i+1}} Y_h(t) dt - \sum_{h=j-1} \int_{t_i}^{t_{i+1}} Y_h(t) dt \right) \\
& + \sum_{i=0}^{\tau} d_i \left(\sum_{h=j+1} \int_{t_i}^{t_{i+1}} Y_h(t) dt - \sum_{h=j-1} \int_{t_i}^{t_{i+1}} Y_h(t) dt \right) \\
& + \sum_{i=0}^{\tau} \left(\frac{a_i}{b_i} - \frac{a_i}{c_i} \right) + (\alpha - 1) \sum_{i=0}^{\tau} \left(\frac{a_i d_i}{b_i^2} - \frac{a_i d_i}{c_i^2} + \frac{g_i}{b_i} - \frac{g_i}{c_i} \right).
\end{aligned}$$

Im nächsten Abschnitt wird noch eine alternative Modellierung der latenten Variablen vorgestellt.

3.3. AR(1)-Frailty

Die im vorherigen Abschnitt vorgestellte Modellierung weist leider keinerlei Autokorrelation zwischen den Werten der Frailty zu verschiedenen Zeitpunkten auf. Deswegen wird hier als Alternative ein um eins schwankender AR(1)-Prozess verwendet. Für die Basisintensitäten werden dabei wieder die gleichen Vereinfachungen vorgenommen wie im vorhergehenden Abschnitt. Dann haben die Intensitäten der Zählprozesse bei gegebenen Beobachtungen der Frailty die Form

$$q_{hj}(t|\psi) = \begin{cases} q \sum_{i=0}^{\tau} \mathbf{1}_{(t_i, t_{i+1}]}(t)(1 + \psi_i), & \text{für } h - j = -1 \\ q \sum_{i=0}^{\tau} \mathbf{1}_{(t_i, t_{i+1}]}(t)(1 - \psi_i), & \text{für } h - j = +1 \\ 0, & \text{sonst,} \end{cases}$$

wobei

$$\begin{aligned} \psi_i &= \rho\psi_{i-1} + \varepsilon_i = \sum_{l=1}^i \rho^{i-l} \varepsilon_l, \quad i = 1, \dots, \tau \\ \psi_0 &= \varepsilon_0 = 0 \\ \varepsilon_i &\stackrel{u.v.}{\sim} N(0, (t_i - t_{i-1})\sigma^2) \end{aligned} \quad (3.3)$$

mit $u.v.$ $\hat{=}$ „unabhängig verteilt“. Die Varianz der ε_i ist nicht nur von einem Varianzparameter σ^2 abhängig, sondern ähnlich wie von Koopman et al. (2008) vorgeschlagen auch von der Länge des vorhergehenden Intervalls. Dies soll die Probleme, die durch die verschiedenen Länge der Intervalle des AR(1) Prozesses entstehen, beheben. Die Modellierung der Frailty als AR(1) Prozess und nicht als log-AR(1)-Prozess birgt offensichtlich das Problem, dass die Intensitäten theoretisch negativ werden können und somit der Zählprozess nicht definiert wäre. Für einen log-AR(1)-Prozess wäre aber die Laplace-Approximation nicht anwendbar, da kein explizites Minimum der Funktion $g(\cdot)$ gefunden werden kann. Man kann aber davon ausgehen, dass in der Realität negative Intensitäten nicht auftauchen. Deswegen wird davon ausgegangen, dass das Problem keine großen Auswirkungen auf die Schätzung hat.

Die Likelihood gegeben der Frailty ergibt sich dann zu

$$\begin{aligned}
& L(q \mid \varepsilon_i) \\
&= \prod_{i=0}^{\tau} (1 + \sum_{l=1}^i \rho^{i-l} \varepsilon_l)^{\Delta N_+(t_{i+1})} (1 - \sum_{l=1}^i \rho^{i-l} \varepsilon_l)^{\Delta N_-(t_{i+1})} \prod_{h \neq j} (q Y_h(t_i))^{\Delta N_{hj}(t_{i+1})} \\
&\quad \exp \left[- \left\{ (1 + \sum_{l=1}^i \rho^{i-l} \varepsilon_l) \sum_{h=j+1}^i q \int_{t_i}^{t_{i+1}} Y_h(t) dt \right. \right. \\
&\quad \left. \left. + (1 - \sum_{l=1}^i \rho^{i-l} \varepsilon_l) \sum_{h=j-1}^i q \int_{t_i}^{t_{i+1}} Y_h(t) dt + \sum_{h \neq j} q \int_0^{t_1} Y_h(t) dt \right\} \right].
\end{aligned}$$

Mit der Dichte

$$f(\varepsilon) = \prod_{i=1}^{\tau} \frac{1}{\sqrt{2\pi(t_i - t_{i-1})}\sigma} \exp \left[-\frac{\varepsilon_i^2}{2\sigma^2(t_i - t_{i-1})} \right]$$

ist die volle Likelihood dann gegeben durch

$$\begin{aligned}
& L(q, \sigma^2, \rho) \\
&= \int_{\mathbb{R}^{\tau+1}} \prod_{i=0}^{\tau} \frac{1}{\sqrt{2\pi(t_i - t_{i-1})}\sigma} (1 + \sum_{l=1}^i \rho^{i-l} \varepsilon_l)^{\Delta N_+(t_{i+1})} (1 - \sum_{l=1}^i \rho^{i-l} \varepsilon_l)^{\Delta N_-(t_{i+1})} \\
&\quad \prod_{h \neq j} (q Y_h(t_i))^{\Delta N_{hj}(t_{i+1})} \\
&\quad \exp \left[- \left\{ (1 + \sum_{l=1}^i \rho^{i-l} \varepsilon_l) \sum_{h=j+1}^i q \int_{t_i}^{t_{i+1}} Y_h(t) dt \right. \right. \\
&\quad \left. \left. + (1 - \sum_{l=1}^i \rho^{i-l} \varepsilon_l) \sum_{h=j-1}^i q \int_{t_i}^{t_{i+1}} Y_h(t) dt \right. \right. \\
&\quad \left. \left. + \sum_{h \neq j} q \int_0^{t_1} Y_h(t) dt + \frac{\varepsilon_i^2}{2\sigma^2(t_i - t_{i-1})} \right\} \right] d\varepsilon.
\end{aligned}$$

Auch hier liegt ein hochdimensionales Integral vor. Es sind wieder die Voraussetzungen für eine Laplace-Approximation gegeben. Die zu minimierende Funktion hat

aber dieses Mal die Form:

$$\begin{aligned}
g(\varepsilon) &= \sum_{i=0}^{\tau} (1 + \sum_{l=1}^i \rho^{i-l} \varepsilon_l) \sum_{h=j+1} q \int_{t_i}^{t_{i+1}} Y_h(t) dt \\
&+ \sum_{i=1}^{\tau} (1 - \sum_{l=1}^i \rho^{i-l} \varepsilon_l) \sum_{h=j-1} q \int_{t_i}^{t_{i+1}} Y_h(t) dt \\
&+ \sum_{i=1}^{\tau} \frac{\varepsilon_i^2}{2\sigma^2(t_i - t_{i-1})}, \tag{3.4}
\end{aligned}$$

mit den zugehörigen Ableitungen

$$\begin{aligned}
\frac{\partial g(\varepsilon)}{\partial \varepsilon_i} &= \sum_{l=i}^{\tau} \rho^{l-i} \left[\sum_{h=j+1} q \int_{t_l}^{t_{l+1}} Y_h(t) dt - \sum_{h=j-1} q \int_{t_l}^{t_{l+1}} Y_h(t) dt \right] + \frac{\varepsilon_i}{\sigma^2(t_i - t_{i-1})}, \\
\frac{\partial g(\varepsilon)^2}{\partial \varepsilon_i^2} &= \frac{1}{\sigma^2(t_i - t_{i-1})}, \\
\frac{\partial g(\varepsilon)^2}{\partial \varepsilon_i \partial \varepsilon_j} &= 0, \quad i \neq j.
\end{aligned}$$

Die Hessematrix ist also positiv definit. Als Minimum der Funktion $g(\varepsilon)$ erhält man dann

$$\hat{\varepsilon}_i = \sigma^2(t_i - t_{i-1}) \sum_{l=i}^{\tau} \rho^{l-i} \left[\sum_{h=j+1} q \int_{t_l}^{t_{l+1}} Y_h(t) dt - \sum_{h=j-1} q \int_{t_l}^{t_{l+1}} Y_h(t) dt \right]. \tag{3.5}$$

Somit ergibt sich die marginale partielle Likelihood approximativ durch Einsetzen von (3.5):

$$\begin{aligned}
&L(q, \sigma^2, \rho) \\
&\approx \prod_{i=1}^{\tau} (1 + \sigma^2 q \sum_{d=1}^i \rho^{i-d} (t_d - t_{d-1}) \sum_{l=d}^{\tau} \rho^{l-d} \int_{t_l}^{t_{l+1}} Y_1(t) dt)^{\Delta N_+(t_{i+1})}
\end{aligned}$$

$$\begin{aligned}
& (1 - \sigma^2 q \sum_{d=1}^i \rho^{i-d} (t_d - t_{d-1}) \sum_{l=d}^{\tau} \rho^{l-d} \int_{t_l}^{t_{l+1}} Y_1(t) dt)^{\Delta N_-(t_{i+1})} \\
& q^{N_{\bullet}(T)} \prod_{i=0}^{\tau} \prod_{h \neq j} (Y_h(t_i))^{\Delta N_{hj}(t_{i+1})} \exp \left(- \sum_{h \neq j} q \int_0^{t_1} Y_h(t) dt \right) \\
& \exp - \left[q \sum_{i=1}^{\tau} (1 + \sigma^2 q \sum_{d=1}^i \rho^{i-d} (t_d - t_{d-1}) \sum_{l=d}^{\tau} \rho^{l-d} \int_{t_l}^{t_{l+1}} Y_1(t) dt) \sum_{h=j+1} \int_{t_i}^{t_{i+1}} Y_h(t) dt \right. \\
& + q \sum_{i=1}^{\tau} (1 - \sigma^2 q \sum_{d=1}^i \rho^{i-d} (t_d - t_{d-1}) \sum_{l=d}^{\tau} \rho^{l-d} \int_{t_l}^{t_{l+1}} Y_1(t) dt) \sum_{h=j-1} \int_{t_i}^{t_{i+1}} Y_h(t) dt \\
& \left. + \sigma^2 q^2 / 2 \sum_{i=1}^{\tau} (t_i - t_{i-1}) \left[\sum_{l=i}^{\tau} \rho^{l-i} \int_{t_l}^{t_{l+1}} Y_1(t) dt \right]^2 \right].
\end{aligned}$$

Es gilt hierbei, dass $N_+(t) = \sum_{h=j+1} N_{hj}(t)$ und $N_-(t) = \sum_{h=j-1} N_{hj}(t)$. Die log-Likelihood hat also die Form

$$\begin{aligned}
& l(q, \sigma^2, \rho) = \log(L(q, \sigma^2, \rho)) \\
& = \sum_{i=1}^{\tau} \left((N_+(t_{i+1}) - N_+(t_i)) \log(1 + \sigma^2 q \sum_{d=1}^i \rho^{i-d} (t_d - t_{d-1}) \sum_{l=d}^{\tau} \rho^{l-d} \int_{t_l}^{t_{l+1}} Y_1(t) dt) \right. \\
& \quad \left. + (N_-(t_{i+1}) - N_-(t_i)) \log(1 - \sigma^2 q \sum_{d=1}^i \rho^{i-d} (t_d - t_{d-1}) \sum_{l=d}^{\tau} \rho^{l-d} \int_{t_l}^{t_{l+1}} Y_1(t) dt) \right) \\
& + \sum_{i=0}^{\tau} \log \left(\prod_{h \neq j} (Y_h(t_i))^{\Delta N_{hj}(t_i)} \right) + N_{\bullet}(T) \log(q) - q \int_0^{t_1} Y_{\bullet}(t) dt \\
& - \left[q \sum_{i=1}^{\tau} (1 + \sigma^2 q \sum_{d=1}^i \rho^{i-d} (t_d - t_{d-1}) \sum_{l=d}^{\tau} \rho^{l-d} \int_{t_l}^{t_{l+1}} Y_1(t) dt) \sum_{h=j+1} \int_{t_i}^{t_{i+1}} Y_h(t) dt \right. \\
& + q \sum_{i=1}^{\tau} (1 - \sigma^2 q \sum_{d=1}^i \rho^{i-d} (t_d - t_{d-1}) \sum_{l=d}^{\tau} \rho^{l-d} \int_{t_l}^{t_{l+1}} Y_1(t) dt) \sum_{h=j-1} \int_{t_i}^{t_{i+1}} Y_h(t) dt \\
& \left. + \sigma^2 q^2 / 2 \sum_{i=1}^{\tau} (t_i - t_{i-1}) \left[\sum_{l=i}^{\tau} \rho^{l-i} \int_{t_l}^{t_{l+1}} Y_1(t) dt \right]^2 \right].
\end{aligned}$$

Auch hier lassen sich Schätzer nur mit numerischen Methoden berechnen. Die nötigen Ableitungen ergeben sich mit

$$\tilde{a}_i = \sum_{d=1}^i \rho^{i-d} (t_d - t_{d-1}) \sum_{l=d}^{\tau} \rho^{l-d} \int_{t_l}^{t_{l+1}} Y_1(t) dt$$

zu

$$\begin{aligned} \frac{\partial l(q, \sigma^2, \rho)}{\partial q} &= \sum_{i=1}^{\tau} \left(\frac{(N_+(t_{i+1}) - N_+(t_i)) \sigma^2 \tilde{a}_i}{1 + \sigma^2 q \tilde{a}_i} - \frac{(N_-(t_{i+1}) - N_-(t_i)) \sigma^2 \tilde{a}_i}{1 - \sigma^2 q \tilde{a}_i} \right) \\ &\quad + \frac{N_{\bullet}(T)}{q} - \int_0^{t_1} Y_{\bullet}(t) dt \\ &\quad - \sum_{i=1}^{\tau} (1 + 2\sigma^2 q \tilde{a}_i) \sum_{h=j+1}^{\tau} \int_{t_i}^{t_{i+1}} Y_h(t) dt \\ &\quad - \sum_{i=1}^{\tau} (1 - 2\sigma^2 q \tilde{a}_i) \sum_{h=j-1}^{\tau} \int_{t_i}^{t_{i+1}} Y_h(t) dt \\ &\quad - \sigma^2 q \sum_{i=1}^{\tau} (t_i - t_{i-1}) \left[\sum_{l=i}^{\tau} \rho^{l-i} \int_{t_l}^{t_{l+1}} Y_1(t) dt \right]^2, \end{aligned}$$

bzw.

$$\begin{aligned} \frac{\partial l(q, \sigma^2, \rho)}{\partial \sigma^2} &= \sum_{i=1}^{\tau} \left(\frac{(N_+(t_{i+1}) - N_+(t_i)) q \tilde{a}_i}{1 + \sigma^2 q \tilde{a}_i} - \frac{(N_-(t_{i+1}) - N_-(t_i)) q \tilde{a}_i}{1 - \sigma^2 q \tilde{a}_i} \right) \\ &\quad - q^2 \sum_{i=1}^{\tau} \tilde{a}_i \left(\sum_{h=j+1}^{\tau} \int_{t_i}^{t_{i+1}} Y_h(t) dt - \sum_{h=j-1}^{\tau} \int_{t_i}^{t_{i+1}} Y_h(t) dt \right) \\ &\quad - q^2/2 \sum_{i=1}^{\tau} (t_i - t_{i-1}) \left[\sum_{l=i}^{\tau} \rho^{l-i} \int_{t_l}^{t_{l+1}} Y_1(t) dt \right]^2 \end{aligned}$$

und mit

$$\begin{aligned} \tilde{b}_i &= \sum_{d=1}^i \rho^{i-d} (t_d - t_{d-1}) \sum_{l=d+1}^{\tau} (l-d) \rho^{l-d-1} \int_{t_l}^{t_{l+1}} Y_1(t) dt \\ \tilde{c}_i &= \sum_{d=1}^i (i-d) \rho^{i-d-1} (t_d - t_{d-1}) \sum_{l=d+1}^{\tau} \rho^{l-d} \int_{t_l}^{t_{l+1}} Y_1(t) dt \end{aligned}$$

gilt

$$\begin{aligned}
& \frac{\partial l(q, \sigma^2, \rho)}{\partial \rho} \\
= & \sum_{i=1}^{\tau} \left(\frac{(N_+(t_{i+1}) - N_+(t_i))q\sigma^2\tilde{b}_i}{1 + \sigma^2q\tilde{a}_i} + \frac{(N_+(t_{i+1}) - N_+(t_i))q\sigma^2\tilde{c}_i}{1 + \sigma^2q\tilde{a}_i} \right. \\
& \left. - \frac{(N_-(t_{i+1}) - N_-(t_i))\sigma^2q\tilde{b}_i}{1 - \sigma^2q\tilde{a}_i} - \frac{(N_-(t_{i+1}) - N_-(t_i))\sigma^2q\tilde{c}_i}{1 - \sigma^2q\tilde{a}_i} \right) \\
& - \sigma^2q^2 \sum_{i=1}^{\tau} (\tilde{b}_i + \tilde{c}_i) \left(\sum_{h=j+1}^{\tau} \int_{t_i}^{t_{i+1}} Y_h(t) dt - \sum_{h=j-1}^{\tau} \int_{t_i}^{t_{i+1}} Y_h(t) dt \right) \\
& - \sigma^2q^2 \sum_{i=1}^{\tau} (t_i - t_{i-1}) \left[\sum_{l=i+1}^{\tau} (l-i)\rho^{l-i-1} \int_{t_i}^{t_{i+1}} Y_1(t) dt \sum_{l=i}^{\tau} \rho^{l-i} \int_{t_i}^{t_{i+1}} Y_1(t) dt \right].
\end{aligned}$$

Die zweiten Ableitungen sehen folgendermaßen aus:

$$\begin{aligned}
\frac{\partial^2 l(q, \sigma^2, \rho)}{(\partial q)^2} &= - \sum_{i=1}^{\tau} \left(\frac{(N_+(t_{i+1}) - N_+(t_i))(\sigma^2\tilde{a}_i)^2}{(1 + \sigma^2q\tilde{a}_i)^2} + \frac{(N_-(t_{i+1}) - N_-(t_i))(\sigma^2\tilde{a}_i)^2}{(1 - \sigma^2q\tilde{a}_i)^2} \right) \\
& - \sum_{i=1}^{\tau} 2\sigma^2\tilde{a}_i \left(\sum_{h=j+1}^{\tau} \int_{t_i}^{t_{i+1}} Y_h(t) dt - \sum_{h=j-1}^{\tau} \int_{t_i}^{t_{i+1}} Y_h(t) dt \right) \\
& - \frac{N_{\bullet}(T)}{q^2} - \sigma^2 \sum_{i=1}^{\tau} (t_i - t_{i-1}) \left[\sum_{l=i}^{\tau} \rho^{l-i} \int_{t_i}^{t_{i+1}} Y_1(t) dt \right]^2,
\end{aligned}$$

$$\frac{\partial^2 l(q, \sigma^2, \rho)}{(\partial \sigma^2)^2} = - \sum_{i=1}^{\tau} \left(\frac{(N_+(t_{i+1}) - N_+(t_i))(q\tilde{a}_i)^2}{(1 + \sigma^2q\tilde{a}_i)^2} + \frac{(N_-(t_{i+1}) - N_-(t_i))(q\tilde{a}_i)^2}{(1 - \sigma^2q\tilde{a}_i)^2} \right)$$

und mit

$$\begin{aligned}\tilde{d}_i &= \sum_{d=1}^i \rho^{i-d} (t_d - t_{d-1}) \sum_{l=d+2}^{\tau} (l-d)(l-d-1) \rho^{l-d-2} \int_{t_i}^{t_{i+1}} Y_1(t) dt \\ \tilde{e}_i &= \sum_{d=1}^i (i-d) \rho^{i-d-1} (t_d - t_{d-1}) \sum_{l=d+1}^{\tau} (l-d) \rho^{l-d-1} \int_{t_i}^{t_{i+1}} Y_1(t) dt \\ \tilde{f}_i &= \sum_{d=1}^i (i-d)(i-d-1) \rho^{i-d-2} (t_d - t_{d-1}) \sum_{l=d+1}^{\tau} \rho^{l-d} \int_{t_i}^{t_{i+1}} Y_1(t) dt\end{aligned}$$

ergibt sich

$$\begin{aligned}& \frac{\partial l(q, \sigma^2, \rho)}{(\partial \rho)^2} \\ &= \sum_{i=1}^{\tau} \left(\frac{(N_+(t_{i+1}) - N_+(t_i)) (q\sigma^2 \tilde{d}_i + 2q\sigma^2 \tilde{e}_i + q\sigma^2 \tilde{f}_i)}{1 + \sigma^2 q \tilde{a}_i} \right. \\ & \quad + \frac{(N_+(t_{i+1}) - N_+(t_i)) (q\sigma^2 \tilde{b}_i + q\sigma^2 \tilde{c}_i)^2}{(1 + \sigma^2 q \tilde{a}_i)^2} - \frac{(N_-(t_{i+1}) - N_-(t_i)) (\sigma^2 q \tilde{b}_i + \sigma^2 q \tilde{c}_i)^2}{(1 - \sigma^2 q \tilde{a}_i)^2} \\ & \quad \left. - \frac{(N_-(t_{i+1}) - N_-(t_i)) (\sigma^2 q \tilde{d}_i + 2\sigma^2 q \tilde{e}_i + \sigma^2 q \tilde{f}_i)}{1 - \sigma^2 q \tilde{a}_i} \right) \\ & \quad - \sigma^2 q^2 \sum_{i=1}^{\tau} (\tilde{d}_i + 2\tilde{e}_i + \tilde{f}_i) \sum_{h=j+1}^{\tau} \int_{t_i}^{t_{i+1}} Y_h(t) dt \\ & \quad + \sigma^2 q^2 \sum_{i=1}^{\tau} (\tilde{d}_i + 2\tilde{e}_i + \tilde{f}_i) \sum_{h=j-1}^{\tau} \int_{t_i}^{t_{i+1}} Y_h(t) dt \\ & \quad - \sigma^2 q^2 \sum_{i=1}^{\tau} (t_i - t_{i-1}) \left[\left(\sum_{l=i+1}^{\tau} (l-i) \rho^{l-i-1} \int_{t_i}^{t_{i+1}} Y_1(t) dt \right)^2 \right. \\ & \quad \left. + \left(\sum_{l=i}^{\tau} \rho^{l-i} \int_{t_i}^{t_{i+1}} Y_1(t) dt \right) \left(\sum_{l=i}^{\tau} (l-i)(l-i-1) \rho^{l-i-2} \int_{t_i}^{t_{i+1}} Y_1(t) dt \right) \right].\end{aligned}$$

Die gemischten Ableitungen berechnen sich zu

$$\begin{aligned} \frac{\partial^2 l(q, \sigma^2, \rho)}{\partial \sigma^2 \partial q} &= \sum_{i=1}^{\tau} \left(\frac{(N_+(t_{i+1}) - N_+(t_i)) \tilde{a}_i}{(1 + \sigma^2 q \tilde{a}_i)^2} + \frac{(N_-(t_{i+1}) - N_-(t_i)) \tilde{a}_i}{(1 - \sigma^2 q \tilde{a}_i)^2} \right) \\ &\quad - 2q \sum_{i=1}^{\tau} \tilde{a}_i \left(\sum_{h=j+1}^{\tau} \int_{t_i}^{t_{i+1}} Y_h(t) dt - \sum_{h=j-1}^{\tau} \int_{t_i}^{t_{i+1}} Y_h(t) dt \right) \\ &\quad - q \sum_{i=1}^{\tau} (t_i - t_{i-1}) \left[\sum_{l=i}^{\tau} \rho^{l-i} \int_{t_i}^{t_{l+1}} Y_1(t) dt \right]^2, \end{aligned}$$

$$\begin{aligned} &\frac{\partial l(q, \sigma^2, \rho)}{\partial \rho \partial q} \\ &= \sum_{i=1}^{\tau} \left(\frac{(N_+(t_{i+1}) - N_+(t_i)) \sigma^2 \tilde{b}_i + (N_+(t_{i+1}) - N_+(t_i)) \sigma^2 \tilde{c}_i}{(1 + \sigma^2 q \tilde{a}_i)^2} \right. \\ &\quad \left. - \frac{(N_-(t_{i+1}) - N_-(t_i)) \sigma^2 \tilde{b}_i + (N_-(t_{i+1}) - N_-(t_i)) \sigma^2 \tilde{c}_i}{(1 - \sigma^2 q \tilde{a}_i)^2} \right) \\ &\quad - 2\sigma^2 q \sum_{i=1}^{\tau} (\tilde{b}_i + \tilde{c}_i) \sum_{h=j+1}^{\tau} \int_{t_i}^{t_{i+1}} Y_h(t) dt \\ &\quad + 2\sigma^2 q \sum_{i=1}^{\tau} (\tilde{b}_i + \tilde{c}_i) \sum_{h=j-1}^{\tau} \int_{t_i}^{t_{i+1}} Y_h(t) dt \\ &\quad - 2\sigma^2 q \sum_{i=1}^{\tau} (t_i - t_{i-1}) \left[\sum_{l=i+1}^{\tau} (l-i) \rho^{l-i-1} \int_{t_i}^{t_{l+1}} Y_1(t) dt \sum_{l=i}^{\tau} \rho^{l-i} \int_{t_i}^{t_{l+1}} Y_1(t) dt \right] \end{aligned}$$

und

$$\begin{aligned} &\frac{\partial l(q, \sigma^2, \rho)}{\partial \rho \partial \sigma^2} \\ &= \sum_{i=1}^{\tau} \left(\frac{(N_+(t_{i+1}) - N_+(t_i)) q \tilde{b}_i + (N_+(t_{i+1}) - N_+(t_i)) q \tilde{c}_i}{(1 + \sigma^2 q \tilde{a}_i)^2} \right. \\ &\quad \left. - \frac{(N_-(t_{i+1}) - N_-(t_i)) q \tilde{b}_i + (N_-(t_{i+1}) - N_-(t_i)) q \tilde{c}_i}{(1 - \sigma^2 q \tilde{a}_i)^2} \right) \end{aligned}$$

$$\begin{aligned}
& -q^2 \sum_{i=1}^{\tau} (\tilde{b}_i + \tilde{c}_i) \sum_{h=j+1} \int_{t_i}^{t_{i+1}} Y_h(t) dt \\
& + q^2 \sum_{i=1}^{\tau} (\tilde{b}_i + \tilde{c}_i) \sum_{h=j-1} \int_{t_i}^{t_{i+1}} Y_h(t) dt \\
& - q^2 \sum_{i=1}^{\tau} (t_i - t_{i-1}) \left[\sum_{l=i+1}^{\tau} (l-i) \rho^{l-i-1} \int_{t_l}^{t_{l+1}} Y_1(t) dt \sum_{l=i}^{\tau} \rho^{l-i} \int_{t_l}^{t_{l+1}} Y_1(t) dt \right].
\end{aligned}$$

Wie die gefundenen Schätzer sich verhalten, soll mit Hilfe von Simulationen im nächsten Kapitel überprüft werden.

4. Simulationsstudien

In diesem Kapitel soll überprüft werden, inwiefern die im vorigen Kapitel entwickelten Methoden funktionieren. Dafür werden zuerst einige benötigte numerische Methoden vorgestellt und anschließend gezeigt, was passiert, wenn man die Frailty bei der Schätzung der Intensität ignoriert. Dies passiert jeweils bei betaverteilter Frailty und AR(1)-Frailty. Danach werden ML-Schätzer auf Grund der Likelihoods im vorhergehenden Kapitel berechnet und deren Eigenschaften überprüft.

4.1. Numerische Verfahren

Wie zuvor erwähnt, ist es nicht möglich, eine geschlossene Form der ML-Schätzer zu den Likelihoods aus Kapitel 3 anzugeben. Deswegen sind numerische Verfahren nötig, um Schätzer zu finden. Hier werden kurz zwei verwendete Verfahren vorgestellt.

Newtons Methode oder Newton-Raphson Algorithmus: Die Newton-Raphson-Methode ist eine einfache und weit verbreitete Methode, Nullstellen einer Funktion zu finden. Sucht man also das Maximum einer Funktion, muss die Methode auf die Ableitung der Funktion angewendet werden, wenn diese existiert. Die hier gefundene Nullstelle ist dann ein Kandidat für das Maximum.

Möchte man nun die Nullstelle einer Funktion $f(x) : \mathbb{R}^k \rightarrow \mathbb{R}$, mit x aus $M \subset \mathbb{R}^k$, finden geht man folgendermaßen vor (vgl. z.B. Boyd und Vanderberghe, 2004, S. 484ff):

1. wähle ein geeigneter Startwert x_0 ,
2. setze $x_{n+1} = x_n - p \nabla f(x_n)^{-1} f(x_n)$, $p > 0$, n beschreibt dabei den Iterationsschritt,

3. wiederhole 2. und breche ab, wenn $f(x_{n+1})\nabla f(x_{n+1})^{-1}f(x_{n+1}) < \delta$.

Hierbei bezeichnet $\nabla f(\cdot)$ die erste Ableitung von $f(\cdot)$ und δ eine positive Konstante. Im Normalfall wird p als 1 gewählt. Falls das jedoch zur Folge hat, dass der Schritt den möglichen Schätzer aus dem Definitionsbereich führt, wird p solange halbiert, bis dies nicht mehr der Fall ist. Die für den Algorithmus benötigten ersten und zweiten Ableitungen der log-Likelihood sind in den Abschnitten 3.2 und 3.3 zu finden

Simplexalgorithmus von Nelder und Mead: Bei der Methode nach Nelder und Mead (1965) sucht man nach dem Minimum einer Funktion. Möchte man das Maximum finden, multipliziert man die Funktion einfach mit -1 . Das Besondere ist, dass weder die erste noch die zweite Ableitung hierfür verwendet wird. Es handelt sich hierbei um ein verbessertes Simplexverfahren, das Polyeder mit $k + 1$ Ecken im \mathbb{R}^k bildet. Der Algorithmus sieht folgendermaßen aus:

1. Wähle einen Startsimplex mit den Punkten $P_1^{(0)}, \dots, P_{k+1}^{(0)} \in \mathbb{R}^k$.
2. Definiere $P_h^{(n)}$ und $P_l^{(n)}$ durch:

$$f(P_h^{(n)}) = \max \left[f(P_1^{(n)}), \dots, f(P_{k+1}^{(n)}) \right],$$

$$f(P_l^{(n)}) = \min \left[f(P_1^{(n)}), \dots, f(P_{k+1}^{(n)}) \right],$$

und $P_{k+2}^{(n)}$ als Schwerpunkt des Polyeder, wobei n den Iterationsschritt angibt.

3. Reflexion: Setze

$$P_{k+3}^{(n)} = (1 + \kappa)P_{k+2}^{(n)} + \kappa P_h^{(n)}, \quad \kappa > 0.$$

4. Expansion: Ist $f(P_{k+3}^{(n)}) \leq f(P_l^{(n)})$, setze

$$P_{k+4}^{(n)} = (1 - \gamma)P_{k+2}^{(n)} + \gamma P_{k+3}^{(n)}, \quad \gamma > 1.$$

Gilt nun, dass $f(P_{k+4}^{(n)}) < f(P_l^{(n)})$, so ersetze $P_h^{(n)}$ durch $P_{k+4}^{(n)}$, starte wieder mit Schritt 2 und setze $n = n + 1$, ansonsten ersetze $P_h^{(n)}$ durch $P_{k+3}^{(n)}$, starte wieder mit Schritt 2 und setze $n = n + 1$.

5. Kontraktion: Ist $f(P_{k+3}^{(n)}) > f(P_i^{(n)})$, $\forall i \in \{1, \dots, k+1\}$, $i \neq h$, setze

$$P_{k+5}^{(n)} = (1 - \beta)P_{k+2}^{(n)} + \beta P_{k+3}^{(n)}, \quad 0 < \beta < 1,$$

ersetze $P_h^{(n)}$ durch $P_{k+5}^{(n)}$, starte wieder mit Schritt 2 und setze $n = n + 1$.

6. Stoppe, wenn $\sqrt{\frac{1}{k+1} \sum_{i=1}^{k+1} (P_i^{(n)} - P_{k+2}^{(n)})^2} < \delta$.

Die Koeffizienten κ , γ und β geben den Faktor an, durch den das Volumen des Simplex durch die drei Operationen Reflexion, Expansion und Kontraktion geändert werden. Nelder und Mead empfehlen als Werte $\kappa = 1$, $\gamma = 2$ und $\beta = 0.5$.

In den Abschnitten 4.3 und 4.4 findet hauptsächlich ein selbst programmierter Newton-Raphson-Algorithmus Verwendung. An einigen Stellen kommt jedoch das in R implementierte Verfahren nach Nelder und Mead (vgl. R Development Core Team, 2011) zum Einsatz. Darauf wird an den entsprechenden Stellen hingewiesen. Im folgenden Abschnitt sind zunächst jedoch keine numerischen Verfahren notwendig.

4.2. Schätzen bei Ignorieren der Frailty

Dieser Abschnitt untersucht, welchen Einfluss das Ignorieren der latenten Variablen auf den ML-Schätzer (vgl. 2.3) hat. Hierzu werden jeweils 20 000 Portfolios mit n Ratinghistorien und $k = 3$ Ratingklassen simuliert. Es werden sowohl Schätzer im Fall einer betaverteilten Frailty mit dem Verteilungsparameter α als auch einer AR(1)-Frailty mit den Parametern ρ und σ betrachtet. Für die Stichprobenumfänge $n = 1000$ bzw. $n = 2000$ simuliere ich aus Kapazitätsgründen jeweils nur 1 000 Portfolios.

Bei der einfachen Schätzung der Beta-Frailty fällt auf, dass die Ergebnisse zwar relativ nahe am wahren Wert liegen, jedoch scheint der Schätzer für die Basisintensität leicht nach unten verzerrt zu sein, da die Schätzwerte mit zunehmendem n nie gegen den wahren Parameter konvergieren. Das ist aber nicht verwunderlich, da die Intensitätsprozesse bei unserer Modellierung durch die Frailty sehr viel kleiner werden

Tabelle 4.1.: Simulationsergebnisse der einfachen Schätzung bei betaverteilter Frailty

n	α	q	\bar{q}	$sd(\hat{q})$	n	α	q	\bar{q}	$sd(\hat{q})$
100	2	0.4	0.387	0.027	100	4	0.4	0.395	0.027
200	2	0.4	0.385	0.019	200	4	0.4	0.393	0.019
300	2	0.4	0.384	0.015	300	4	0.4	0.393	0.015
100	2	0.1	0.099	0.011	100	8	0.4	0.400	0.027
200	2	0.1	0.098	0.008	200	8	0.4	0.398	0.019
300	2	0.1	0.098	0.006	300	8	0.4	0.397	0.015
100	2	0.9	0.865	0.051	1000	4	0.4	0.391	0.008
200	2	0.9	0.863	0.035	2000	4	0.4	0.391	0.006
300	2	0.9	0.862	0.029					

können, sich aber nicht so stark vergrößern können. Es fällt auf, dass die Schätzung umso näher am tatsächlichen Wert ist, desto größer α wird. Das ist durch die Form der Betaverteilung zu erklären, die bei einem größeren Parameter immer weniger Wahrscheinlichkeitsmasse an den Rändern hat und so dichter um Null liegt. Weitere Simulationen haben jedoch gezeigt, dass dieser Effekt nach den hier betrachteten Werten für α schnell stagniert.

In Tabelle 4.2 wird deutlich, dass die einfache Schätzung im Falle einer AR(1)-Frailty noch näher an dem wahren q liegt. Das könnte an den sehr kleinen Varianzen liegen, die zur Simulation der AR(1)-Prozesse verwendet werden. Im Gegensatz zur betaverteilten Frailty konvergieren die Schätzwerte auch mit zunehmendem n gegen den wahren Parameter. Es scheint also noch nicht einmal eine größere Verzerrung vorzuliegen. Interessant ist aber, dass im Gegensatz zur betaverteilten Frailty der Schätzer eher überschätzt, obwohl hier eigentlich derselbe Effekt zu erwarten wäre. Außerdem wird deutlich, dass schon bei einem Stichprobenumfang von $n = 1000$ der

Tabelle 4.2.: Simulationsergebnisse der einfachen Schätzung bei AR(1)-Frailty

n	ρ	σ^2	q	$\bar{\hat{q}}$	$sd(\hat{q})$	n	ρ	σ^2	q	$\bar{\hat{q}}$	$sd(\hat{q})$
100	0.9	0.05	0.4	0.403	0.027	100	0.4	0.05	0.4	0.404	0.027
200	0.9	0.05	0.4	0.402	0.019	200	0.4	0.05	0.4	0.402	0.019
300	0.9	0.05	0.4	0.401	0.015	300	0.4	0.05	0.4	0.401	0.015
100	0.9	0.05	0.1	0.102	0.012	100	0.9	0.1	0.4	*	*
200	0.9	0.05	0.1	0.101	0.008	200	0.9	0.1	0.4	0.401	0.019
300	0.9	0.05	0.1	0.101	0.006	300	0.9	0.1	0.4	0.401	0.015
100	0.9	0.05	0.9	* ¹	*	1 000	0.9	0.05	0.4	0.400	0.008
200	0.9	0.05	0.9	0.903	0.036	2 000	0.9	0.05	0.4	0.401	0.006
300	0.9	0.05	0.9	0.902	0.029						

mittlere Schätzwert sehr exakt wird. Das ist zwar in den Simulationen ein großer Wert, in echten Portfolios dürfte er aber problemlos zu erreichen sein.

In beiden Fällen sind die Standardabweichungen der Schätzer gleich und hängen nur vom Stichprobenumfang und dem tatsächlichen q und nicht von den Parametern der Frailtyverteilung ab. Nachdem nun der einfache Schätzer schon so gute Ergebnisse erzielt, ist es interessant zu sehen, ob der richtige Maximum-Likelihood-Schätzer besser oder schlechter abschneidet.

4.3. Likelihoodschätzer bei betaverteilter Frailty

Um nun herauszufinden, wie gut der Maximum-Likelihood-Schätzer bei einer Laplace Approximation ist, werden Simulationen durchgeführt. Wieder handelt es sich um jeweils 20 000 Portfolios mit n Ratinghistorien und $k = 3$ Ratingstufen. Bei

¹Hier bricht die Datensimulation in einem von ungefähr 1 000 Fällen ab. Da das Problem so selten auftritt, ließ es sich nicht reproduzieren und lösen.

Tabelle 4.3.: Simulationsergebnisse bei bekanntem α

n	α	q	\bar{q}_α^{ML}	$sd(\hat{q}_\alpha^{ML})$	n	α	q	\bar{q}_α^{ML}	$sd(\hat{q}_\alpha^{ML})$
100	2	0.4	0.350	0.024	100	4	0.4	0.391	0.026
200	2	0.4	0.350	0.017	200	4	0.4	0.391	0.018
300	2	0.4	0.350	0.014	300	4	0.4	0.391	0.015
100	2	0.1	0.088	0.010	100	8	0.4	0.398	0.027
200	2	0.1	0.089	0.007	200	8	0.4	0.398	0.019
300	2	0.1	0.089	0.006	300	8	0.4	0.398	0.015
100	2	0.9	*	*	1 000	4	0.4	0.391	0.008
200	2	0.9	0.784	0.032	2 000	4	0.4	0.392	0.006
300	2	0.9	0.784	0.026					

den großen Stichprobenumfängen $n = 1000, 2000$ werden nur 1 000 Portfolios simuliert. Zunächst geht es um betaverteilte Frailties. Der Fall, in dem der Parameter α fest und bekannt ist, wird als erstes betrachtet. Als Startwert für den Newton-Raphson-Algorithmus findet hierbei der naive Schätzer aus (2.3) Verwendung. Die sich ergebenden mittleren Schätzwerte für die Basisintensität q in Tabelle 4.3 liegen teilweise weit von den tatsächlichen Werten entfernt, werden aber exakter, je größer α ist. Das ist wie oben auf die Form der Beta-Verteilung zurückzuführen. Auffällig ist hier außerdem, dass sich die Schätzwerte bei steigendem n nicht wesentlich dem wahren Wert für q annähern. Die Veränderungen werden meistens erst ab der vierten Nachkommastelle sichtbar. Das ändert sich auch bei den großen und realistischeren Stichprobenumfängen nicht wesentlich.

Da der Parameter der Frailty-Verteilung α im Normalfall jedoch nicht bekannt ist, soll dieser Parameter natürlich auch geschätzt werden. Die Simulation wird zunächst aber wieder vereinfacht und der Fall bei bekanntem q wird betrachtet. In Tabelle 4.4 wird offensichtlich, dass die Schätzergebnisse sich für die verschiedenen α nicht

Tabelle 4.4.: Simulationsergebnisse bei bekanntem q

n	α	q	$\bar{\hat{\alpha}}_q^{ML}$	$sd(\hat{\alpha}_q^{ML})$	n	α	q	$\bar{\hat{\alpha}}_q^{ML}$	$sd(\hat{\alpha}_q^{ML})$
100	2	0.4	2.150	0.025	100	4	0.4	2.153	0.025
200	2	0.4	2.154	0.017	200	4	0.4	2.157	0.018
300	2	0.4	2.155	0.014	300	4	0.4	2.158	0.014
100	2	0.1	2.101	0.036	100	8	0.4	2.155	0.025
200	2	0.1	2.110	0.026	200	8	0.4	2.158	0.019
300	2	0.1	2.113	0.021	300	8	0.4	2.159	0.015
100	2	0.9	*	*	1 000	4	0.4	2.159	0.008
200	2	0.9	2.163	0.015	2 000	4	0.4	2.159	0.006
300	2	0.9	2.163	0.012					

deutlich unterscheiden. Insgesamt hat die Wahl von q einen größeren Einfluss auf den Schätzer als die Wahl von α . Da die Standardabweichungen der Schätzer klein ausfallen, lässt sich dieses Ergebnis auch mit großen Schwankungen der Schätzer nicht erklären.

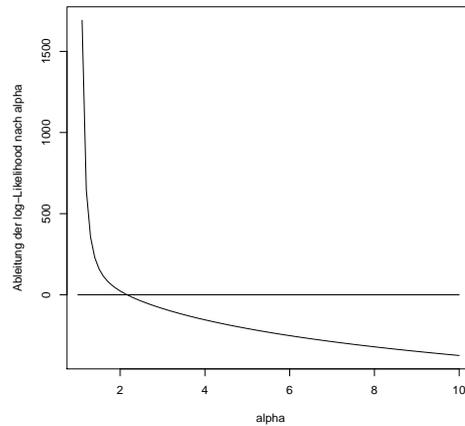
Am Schluss sollen die Parameter nun noch zusammen geschätzt werden. Die Ergebnisse der Simulationen hierfür sind in Tabelle 4.5 zu finden. Sowohl q als auch α werden deutlich unterschätzt, wobei auch hier wieder der Effekt zu erkennen ist, dass für $\alpha > 2$ die mittleren Schätzwerte weit von den tatsächlichen Werten entfernt liegen und eher immer kleiner werden. Die Schätzwerte für die Basisintensität q unterschätzen zwar auch bei großen Stichprobenumfängen deutlich, zeigen aber erwartungsgemäße Unterschiede bei verschiedenen wahren Werten für q . Das ist aber nicht abhängig vom wahren Wert für α , also auch nicht von dem Maß, in dem α unterschätzt wird.

Die Schätzwerte für α schwanken bei der Schätzung mit bekanntem q immer um 2.15 und bei der Schätzung beider Parameter um 2.4, unabhängig vom wahren

Tabelle 4.5.: Simulationsergebnisse für den gemeinsamen Schätzer

n	α	q	\bar{q}^{ML}	$sd(\hat{q}^{ML})$	$\bar{\alpha}^{ML}$	$sd(\hat{\alpha}^{ML})$
100	2	0.4	0.361	0.024	2.239	0.022
200	2	0.4	0.362	0.017	2.242	0.017
300	2	0.4	0.361	0.014	2.243	0.013
100	2	0.1	0.090	0.010	2.185	0.026
200	2	0.1	0.091	0.007	2.194	0.018
300	2	0.1	0.091	0.006	2.198	0.015
100	2	0.9	*	*	*	*
200	2	0.9	0.812	0.032	2.253	0.014
300	2	0.9	0.812	0.027	2.254	0.011
100	4	0.4	0.365	0.025	2.238	0.022
200	4	0.4	0.366	0.017	2.241	0.016
300	4	0.4	0.366	0.014	2.243	0.013
100	8	0.4	0.368	0.025	2.237	0.022
200	8	0.4	0.368	0.017	2.240	0.016
300	8	0.4	0.368	0.014	2.242	0.013
1 000	4	0.4	0.366	0.008	2.244	0.007
2 000	4	0.4	0.366	0.006	2.244	0.005

Wert. Um herauszufinden, ob es sich um ein generelles Problem handelt oder ob diese Phänomen vom Schätzalgorithmus abhängt, wird alternativ der Algorithmus von Nelder und Mead (1965) mit denselben Startwerten wie oben zur Schätzung beider Parameter verwendet. Hier ergeben sich bei $n = 100$, $\alpha = 4$ und $q = 0.4$ die mittleren Schätzer $\bar{q}_2^{ML} = 0.687$ ($sd(\hat{q}_2^{ML}) = 0.048$) und $\bar{\alpha}_2^{ML} = 1$ ($sd(\hat{\alpha}_2^{ML}) < 10^{-10}$). Die Schätzung ist also noch schlechter. Betrachtet man nun beispielhafte Ableitungen von diesem Szenario, ist in Abbildung 4.1 auch deutlich zu erkennen,

Abbildung 4.1.: Ableitung von $l(q, \alpha)$ nach α bei $\alpha = 4$

dass die Nullstelle ungefähr bei 2.15 wie im ersten Fall liegt.

Wenn man die Standardabweichungen in den zwei verschiedenen Szenarien für den Schätzer von q betrachtet und sie auch mit denen des einfachen Schätzers aus dem vorherigen Abschnitt vergleicht, wird deutlich, dass sie hier im Gegensatz zum einfachen Schätzer auch von den Parametern der Frailtyverteilung abhängen. Im Falle von bekannten α sind die Standardabweichungen auch sehr ähnlich wie beim einfachen Schätzer. Erstaunlich ist, dass bei der gemeinsamen Schätzung die Standardabweichung gleich bleibt, obwohl der Schätzer wesentlich schlechtere Ergebnisse erzielt. Die Standardabweichungen des Schätzers für α werden auch bei der gemeinsamen Schätzung kleiner. Sie hängen ebenfalls von den Originalwerten von q ab.

Insgesamt ist eine gute Schätzung von q möglich, doch α ist nicht zufriedenstellend schätzbar. Die Laplace-Approximation nähert die Likelihood also nicht genau genug an. Dies könnte unter anderem daran liegen, dass es sich hierbei um jeweils sehr hochdimensionale Integrale handelt. Betrachtet man jedoch in Abbildung 4.2 den Vergleich der angenäherten log-Likelihood mit der, für die die Frailty bekannt ist, wird deutlich, dass vor allem ein Ausschlag im Bereich für q zwischen 0.2 und 0.4 und für α um die 2 für die Ergebnisse der Schätzung verantwortlich ist. Betrachtet man

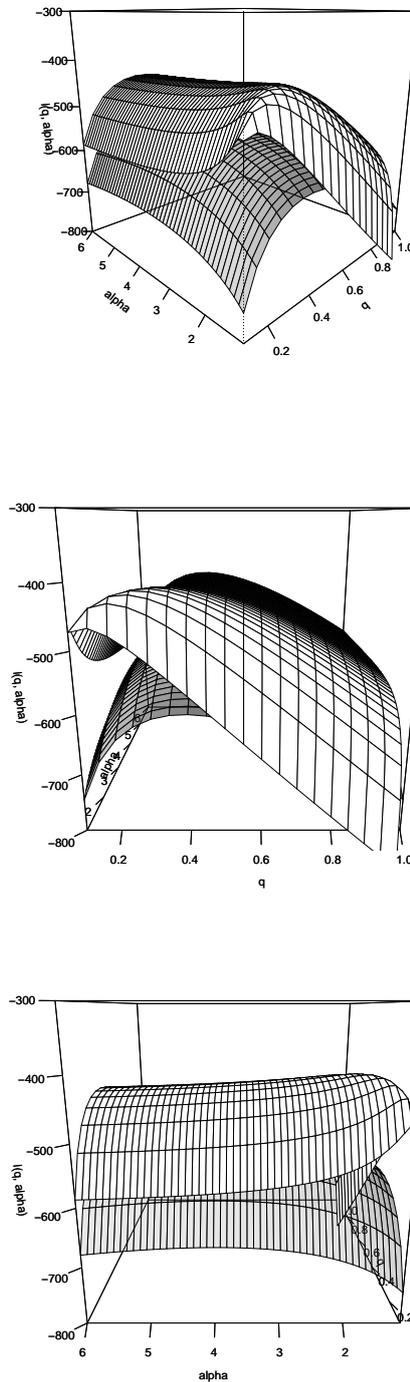
Abbildung 4.2.: $l(q, \alpha)$ (weiß) und $\log(L(q, \alpha | \psi))$ (grau) für $q = 0.4$, $\alpha = 4$ und $n = 100$

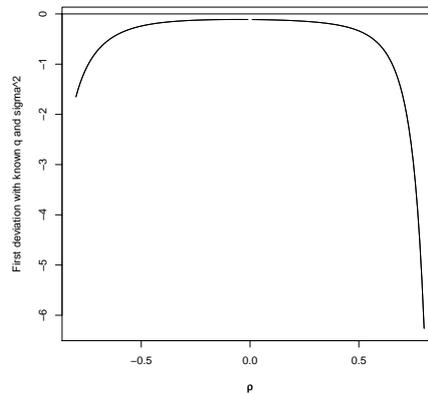
Tabelle 4.6.: Simulationsergebnisse bei bekanntem ρ und σ^2

n	ρ	σ^2	q	$\bar{q}_{\rho,\sigma^2}^{ML}$	$sd(\hat{q}_{\rho,\sigma^2}^{ML})$	n	ρ	σ^2	q	$\bar{q}_{\rho,\sigma^2}^{ML}$	$sd(\hat{q}_{\rho,\sigma^2}^{ML})$
100	0.9	0.05	0.4	0.402	0.027	100	0.4	0.05	0.4	0.404	0.270
200	0.9	0.05	0.4	0.401	0.019	200	0.4	0.05	0.4	0.402	0.019
300	0.9	0.05	0.4	0.401	0.015	300	0.4	0.05	0.4	0.402	0.015
100	0.9	0.05	0.1	0.101	0.012	100	0.9	0.1	0.4	0.397	0.028
200	0.9	0.05	0.1	0.101	0.008	200	0.9	0.1	0.4	0.400	0.019
300	0.9	0.05	0.1	0.101	0.006	300	0.9	0.1	0.4	0.401	0.015
100	0.9	0.05	0.9	*	*	1 000	0.9	0.05	0.4	0.400	0.008
200	0.9	0.05	0.9	0.903	0.036	2 000	0.9	0.05	0.4	0.401	0.006
300	0.9	0.05	0.9	0.902	0.029						

die log-Likelihoodfunktion genauer, wird deutlich, falls man $-(\tau + 1)\log(\sqrt{\alpha - 1})$ aus dem Korrekturterm der Laplace-Approximation weglässt, so verschwindet dieser „Peak“ von den Bildern. Der Term hängt nur über τ von den Daten ab und erklärt so auch, warum der Schätzer kaum einen Zusammenhang zum wahren α zeigt. Ohne diesen Term, bzw. ohne den Korrekturterm zeigt die Funktion im betrachteten Wertebereich jedoch gar kein Maximum für α . Hierdurch wird auch gezeigt, dass die gefundenen Schätzer für α überhaupt nur durch die hohe Dimension von τ möglich sind. Im Falle einer kleineren Dimension hätte der Term kaum Einfluss und es gäbe kein Maximum. Inwiefern die Schätzer bei AR(1)-Frailties besser ausfallen, wird im nächsten Abschnitt betrachtet.

4.4. Likelihoodschätzer bei einer AR(1)-Frailty

Nun soll auch das Verhalten der Schätzer im Falle der AR(1)-Frailty betrachtet werden. Hier werden für den Parameter der Basisintensität q und dem der Varianz

Abbildung 4.3.: Ableitung von $l(q, \sigma^2, \rho)$ nach ρ

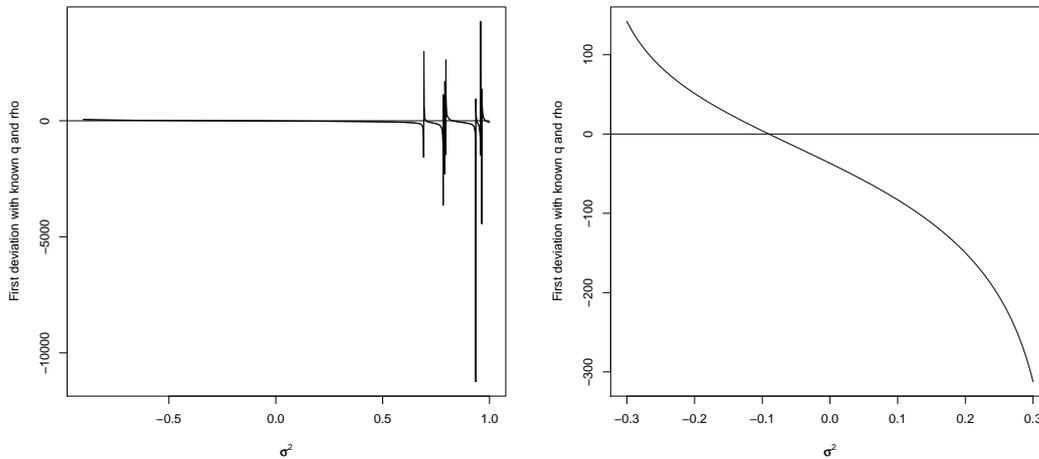
σ^2 Werte zwischen 0 und 1 verwendet. Die Varianz wird möglichst klein gewählt, um nicht Gefahr zu laufen, eine negative Frailty zu bekommen.

Für den Autoregressionsparameter ρ kann zwar gelten $-1 \leq \rho \leq 1$, wir betrachten aber nur Fälle $0 < \rho < 1$, da man bei der Konjunktur von einer positiven Korrelation ausgehen kann.

Zunächst gelten die Parameter σ^2 und ρ als bekannt. Der naive Schätzer dient wieder als Startwert für den Algorithmus. Die Ergebnisse der Simulationen sind in Tabelle 4.6 zu finden.

Es wird deutlich, dass die Schätzwerte auch schon für kleine Stichprobenumfänge wie $n = 200$ im Mittel nahe dem tatsächlichen Parameter liegen. Bei realistischen Portfoliogrößen mit mehreren tausend Schuldner ist das Verhalten des Schätzers natürlich noch besser. Auch die Standardabweichung wird mit zunehmenden Stichprobenumfang genügend klein. Es wird aber deutlich, dass der Schätzer q eher leicht über- als unterschätzt. Jedoch ist auch sofort ersichtlich, dass der als Startwerte verwendete einfache Schätzer in Tabelle 4.2 genauso gut ist. Auch die Standardabweichungen unterscheiden sich nicht. Sie sind hier ebenfalls nicht abhängig von den Parametern der Frailtyverteilung.

Möchte man dasselbe nun für die Parameter des AR(1)-Prozesses ρ und σ^2 durch-

Abbildung 4.4.: Ableitung von $l(q, \sigma^2, \rho)$ nach σ^2

führen, aber dabei kommt es zu Problemen. Der Newton-Raphson Algorithmus konvergiert auch bei Verwendung der wahren Werte als Startwerte nicht. Betrachtet man nun exemplarisch simulierte Ableitungen nach dem jeweiligen Parameter, wenn die beiden übrigen als bekannt angenommen werden, können Erklärungen gefunden werden. Es gilt jeweils $q = 0.4$, $\sigma^2 = 0.05$ und $\rho = 0.9$. Für den Parameter ρ wird in Abbildung 4.3 deutlich, dass keine Nullstelle existiert. Für σ^2 existieren dagegen gleich mehrere. In Abbildung 4.4 ist zu erkennen, dass es mehrere lokale Maxima geben muss. Möchte man so eine Art globales Maximum finden, wäre dies wahrscheinlich die Nullstelle, die im Ausschnitt zu erkennen ist. Leider ist diese negativ und demnach für eine Varianz ein ungünstiger Schätzer. Dasselbe Verhalten ist auch bei anderen Parameter-Konstellationen zu beobachten. Als Problemquelle ist auszumachen, dass negative Werte für die Intensitäten möglich sind, was nicht nur Probleme im Modell, sondern auch bei der Laplace-Approximation erzeugen kann. Zusammenfassend kann die Basisintensität zwar bei einer AR(1)-Frailty wesentlich besser geschätzt werden als für eine betaverteilte Frailty, jedoch ist das auch nur bei bekannten Parametern des AR(1)-Prozesses möglich. Diese können überhaupt nicht geschätzt werden.

5. Anwendung auf ein Portfolio der WestLB AG

Nun folgt die Anwendung der neuen Schätzer auf ein echtes Portfolio. Die verschiedenen Modelle werden miteinander verglichen und der Einfluss verschiedener Parameter auf den Schätzer untersucht.

5.1. Daten

Als Daten stehen interne Ratings der WestLB AG Düsseldorf zur Verfügung. Der Datensatz basiert auf einer Stichprobe von 2 278 Kunden aus dem Geschäftskundenportfolio. Das Teilportfolio ist aus Datenschutzgründen jedoch nicht repräsentativ. Die Ratingmigrationen wurden über einen Zeitraum von sechseinhalb Jahren von Mitte 1999 bis Ende 2005 beobachtet. Das Portfolio beinhaltet also einen kompletten Kreditzyklus. Da die Ratinghistorien hier tagesgenau abgefasst werden, kann man sie somit als zeitstetig ansehen.

Die Ratings wurden mit Hilfe des bankinternen Ratingverfahrens für Geschäftskunden erstellt, das die 20 nicht-Ausfall Ratingklassen 1-20 sowie eine absorbierenden Default-Status „21“ umfasst. In einigen Fällen (cure cases), in denen ein Schuldner einige Zeit nach dem Ausfall wieder Zahlungsverpflichtungen nachkommt, wird die Beobachtungshistorie zum Zeitpunkt des Ausfalls zensiert. Die Historie wird ebenfalls zensiert, falls die Geschäftsbeziehung zwischen der Bank und einem ihrer Kunden regulär endet. Dieser Zeitpunkt wird auch erfasst.

Die Untersuchung umfasst insgesamt 4 393 Migrationen zwischen den 21 Ratingklassen binnen ca. 8 710 Jahren an kumulierten Beobachtungshistorien. Davon sind mehr als die Hälfte, nämlich 2 540 Migrationen, zwischen zwei benachbarten Rating-

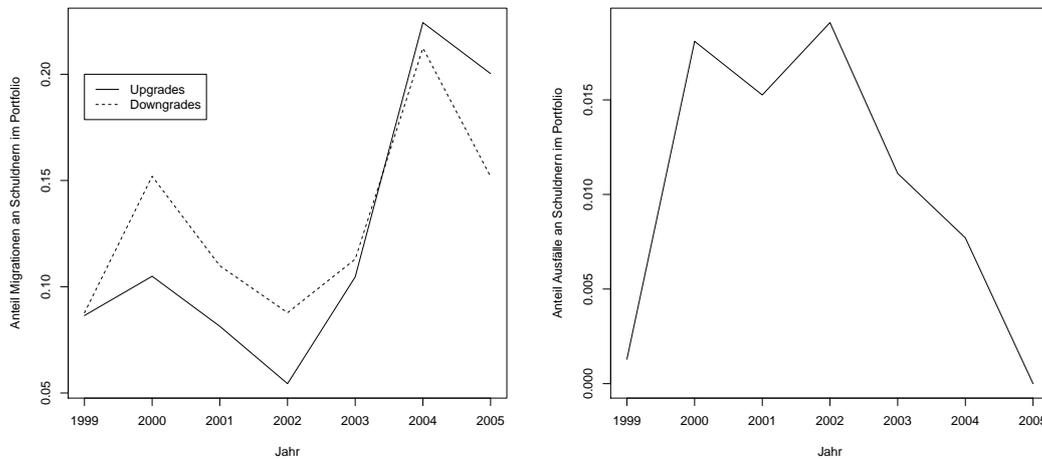


Abbildung 5.1.: Relative Migrationen und Ausfälle pro Jahr

stufen. Da im ersten Portfoliojahr wird meistens keine Änderung an der Einstufung eines Schuldners vorgenommen, so entsteht eine Verzerrung des Risikos, da die Migrationswahrscheinlichkeiten unterschätzt werden. Deswegen wird dieses Jahr für alle Schuldner aus den Daten entfernt. Es bleiben noch 3 426 Migrationen und 7 048 Jahre an kumulierten Beobachtungshistorien übrig.

Für die Schätzung bei betaverteilten Fraitleties ist es eine Voraussetzung, dass in der ersten Klasse in jedem Zeitabschnitt mindestens ein Schuldner unter Risiko stehen muss. Bei den vorliegenden Daten ist dies leider nicht der Fall. Es sind überhaupt nur 2 Schuldner jemals in der besten Ratingkategorie, ähnliches gilt für die zweit- bis fünftbeste Ratingklasse (8, 8, 24 und 54 Schuldner). Darum werden die ersten fünf Klassen zusammengelegt. Danach liegen nur noch 3 414 Migrationen vor. Die Länge der kumulierten Beobachtungshistorien bleibt natürlich gleich.

In Abbildung 5.1 ist zu sehen, dass in den Jahren 2000-2002, also als die Neuen Märkte zusammenbrachen, der Anteil der Downgrades wesentlich höher ist. In den späteren Jahren der Erholung ändert sich dieses Verhältnis. Ab 2004 überwiegen dann die Upgrades. Das bestätigt die Annahme, dass die Wahrscheinlichkeit für ein Upgrade in wirtschaftlich schlechten Zeiten kleiner ist als für ein Downgrade und dass

es in wirtschaftlich guten Zeiten genau umgekehrt ist. Trotz dieser Beobachtungen wird aber auch deutlich, dass es dazu vor allem noch einen gemeinsamen Trend von Upgrades und Downgrades gibt.

Es wird auch deutlich, dass in den Jahren 2000-2002 wesentlich mehr Schuldner ausfallen als in den Jahren danach. Die Abhängigkeit des Ausfalls von der Konjunktur wird also noch einmal untermauert. Im Jahr 1999 gibt es nur einen Ausfall, im Jahr 2005 gar keinen. Der erste Fall lässt sich dadurch erklären, dass sich zu diesem Zeitpunkt vor allem „junge“ Kredite im Portfolio befinden. Der Wert für 2005 lässt sich so nicht erklären, auch nicht durch eine geringe Anzahl an Schuldnern. Es befinden sich noch 1 383 Beobachtungen im Portfolio, von denen die meisten am Ende des Jahres zensiert werden. Offensichtlich führt die gute konjunkturelle Lage dazu, dass tatsächlich keine Firma ausfällt oder später wieder in der Lage ist zu bezahlen und nur zensiert wird.

5.2. Ergebnisse

Zunächst wird angenommen, es lägen betaverteilte Frailties vor. Die Annahmen an die Intensitäten entsprechen denen in 3.2. Zunächst wird wie in den Simulationen der einfache Schätzer berechnet, der auch als Startwert für die Schätzalgorithmen fungiert. Er ergibt sich hier zu 0.177. Die für verschiedene α berechneten ML-Schätzer weichen hiervon kaum ab.

Tabelle 5.1.: Schätzer für verschiedene α

α	2	4	8	20
\hat{q}_{ML}	0.177	0.177	0.177	0.177

Der Unterschied wird erst in der vierten oder fünften Nachkommastelle ersichtlich, auch zwischen den verschiedenen verwendeten α (vgl. Tabelle 5.1). Zu bemerken ist

also, dass die Abweichung zwischen einfachem und neuem Schätzer nicht so groß ist, wie die in den Simulationen gefundene.

Als nächstes soll α bei als bekannt angenommenen q geschätzt werden. Hier werden wieder verschiedene Werte ausprobiert. Als Startwert für α wird 2 verwendet, da eine Änderung hier ohne Effekt auf die Schätzung blieb.

Tabelle 5.2.: Schätzer für verschiedene q

q	0.1	0.2	0.5	1
$\hat{\alpha}_{ML}$	1.005	1.011	1.028	1.062

Nach den Ergebnissen aus Tabelle 5.2 zeigt die Wahl des Parameters q einen großen Einfluss auf die Schätzung von α . Trotzdem liegen alle Werte nur knapp über über dem Randwert 1. Das bedeutet, dass hier fast von einer Rechteckverteilung auf $(-1, 1)$ als Frailty ausgegangen werden kann, was auch auf eine relativ große Streuung der Frailty hinweist.

Schätzt man nun die Parameter q und α gemeinsam, ergeben sich die Schätzer zu $\hat{q}^{ML} = 0.148$ und $\hat{\alpha}_{ML} = 1.012$, wobei für die Basisintensität q der übliche Startwert verwendet wurde und für α wiederum 2. Aber auch hier hat die Wahl des Startwerts keinen Einfluss auf den bzw. die Schätzer. Wie schon bei den Simulationen ist der Schätzwert für die Basisintensität um einiges kleiner als im Fall eines bekannten α . Gleiches gilt für $\hat{\alpha}_{ML}$ verglichen mit dem Schätzer bei bekanntem q . Das Ergebnis weist wieder auf eine quasi rechteckverteilte Frailty hin. Insgesamt zeigen die verschiedenen Schätzer auf jeden Fall ein ähnliches Verhalten wie in den Simulationen. Als nächstes wird q bei Annahme einer AR(1)-verteilten Frailty geschätzt. Da die Schätzer für die Parameter ρ und σ^2 sich nicht ermitteln lassen, sie aber dennoch unbekannt sind, werden auch hier verschiedene Parameterwerte ausprobiert.

Die Ergebnisse in Tabelle 5.3 sind für diese verschiedenen Werte nicht exakt gleich, unterscheiden sich aber erst ab der vierten, zum Großteil sogar erst ab der fünften

Tabelle 5.3.: Schätzer für verschiedene ρ und σ^2

ρ	0.90	0.90	0.90	0.99	0.40
σ^2	0.05	0.01	5.00	0.05	0.05
\hat{q}_{ML}	0.177	0.177	0.177	0.177	0.177

Nachkommastelle. Die Wahl der unbekannt Parameter ρ und σ^2 hat also wenig Einfluss auf das Ergebnis. Die gefundenen Schätzer entsprechen dem Ergebnis des einfachen Schätzers, aber auch des Schätzers bei betaverteilter Frailty. Auch damit zeigen sie erneut ein ähnliches Verhalten wie in den Simulationen. Dass die Abweichungen noch geringer sind, kann vor allem auf den großen Stichprobenumfang zurückgeführt werden. Ähnliches ließ sich schon bei den größeren n in den Simulationen beobachten.

6. Zusammenfassung und Diskussion

Ziel dieser Arbeit ist es, neue Schätzer für Zählprozesse mit vorliegender dynamischer Frailty am konkreten Problem von Ratingmigrationen zu entwickeln. Wenn man diese in Kalenderzeit betrachtet, wird davon ausgegangen, dass sie über die Konjunktur korreliert sind. Eine latente Variable oder Frailty modelliert dann diese Konjunktur. Bisherige Arbeiten haben in solchen Modellen bisher nur mit großem numerischen Aufwand Schätzer gefunden.

Zwei verschiedene Modelle werden dafür vorgestellt: eines mit einer betaverteilten Frailty, welche aber nicht autokorreliert ist, und eines mit einer AR(1)-Frailty. Beim zweiten Modell kann jedoch das Problem auftreten, dass die Intensitätsprozesse negativ werden. Möchte man für beide Modelle eine Likelihood entwickeln, stößt man jeweils auf das Problem eines hochdimensionalen Integrales, das hier eine Laplace-Approximation lösen soll.

In den Simulationen zeigt sich, dass bei bekannten Parameter der Frailtyverteilung sich der Parameter der Basisintensität q sehr gut schätzen lässt, vor allem bei großen und so auch realistischen Portfolios. Im Fall der AR(1)-Frailty werden jedoch keine besseren Ergebnisse erzielt als bei einer einfachen Schätzung, die die Frailty ignoriert. Wenn hingegen eine betaverteilte Frailty vorliegt, sind die Schätzer bei deren Beachtung zwar weiter vom wahren q entfernt als die einfache Schätzung, dort ist der Schätzer aber eindeutig verzerrt. Die Schätzwerte entwickeln sich bei steigendem n vom wahren Wert des Parameters weg.

Bei der Schätzung der Frailty-Verteilungsparameter bei bekannter Basisintensität und der gemeinsamen Schätzung aller Parameter kommt es aber zu größeren Problemen. Im Falle der AR(1)-Frailty ist eine Schätzung der Parameter σ^2 und ρ gar nicht möglich. Der Parameter α der betaverteilten Frailty lässt sich zwar schätzen, der Schätzer ist aber teilweise erheblich verzerrt, vor allem nicht abhängig vom wahren

ren Wert des Parameters. Dies liegt vor allem an einem Korrekturterm der Laplace-Approximation, der nicht von den Daten abhängt. Bei der gemeinsamen Schätzung wird außerdem q stark unterschätzt, was im Falle von Ratingmigrationen eine sehr schlechte Eigenschaft ist.

Für beide Modelle werden Schätzer aus einem Datensatz berechnet. Der Schätzer hängt nur marginal von der Wahl der Frailty-Verteilungsparameter ab und zeigt auch nur geringfügige Unterschiede zu dem einfachen Schätzer. Nur bei der Schätzung beider Parameter im betaverteilten Fall gibt es, ähnlich wie bei den Simulationen, einen bemerkenswerten Unterschied.

Eine Schätzung des Parameter der Basisintensität ist folglich sehr gut möglich, eine gemeinsame Schätzung mit den Parametern der Verteilung des latenten Faktors jedoch schwieriger. Somit ist eine Einschätzung des Migrations- und auch des Ausfallrisikos mit Hilfe von q erreichbar. Prognosen sind dagegen bei unbekanntem Frailtyparametern aber eher nicht möglich. Das Problem ist wohl auf die Ungenauigkeit der Laplace-Approximation zurückzuführen. Ursprung könnte unter anderem sein, dass je nach Stichprobenumfang und Beobachtungsdauer des Datensatzes ein sehr hochdimensionales Integral vorliegt, für das eine Laplace Approximation nur noch wenig geeignet ist. Dass bei einer AR(1)-Frailty die Probleme verstärkt auftreten, ist eben dadurch zu erklären, dass die Voraussetzungen für Intensitätsprozesse und damit auch für die Laplace-Approximation nicht eingehalten werden. Zudem existiert aber auch nicht zufällig ein Term, der eine Form der Approximation erzeugt, dass Maxima für alle Parameter existieren, wie es bei der betaverteilten Variante der Frailty der Fall ist.

Zusammenfassend ist die Laplace-Approximation nicht der Lösungsansatz, der im vorgestellten Problem zu guten und handhabbaren Ergebnissen führt. Die Anwendung auf zwei verschiedene Frailty-Modelle macht deutlich, dass eine andere Wahl der Verteilung kaum bessere Ergebnisse erzielen wird. Eine Möglichkeit wäre, die Dimension des Integrals radikal einzuschränken (evtl. Änderung der Frailty nur einmal im Jahr), oder die Second-Order Laplace-Approximation zur Annäherung des hoch-

dimensionalen Integrals zu verwenden. Diese enthält einen zusätzlichen Korrekturterm und ist besser geeignet, wenn die Dimension des Integrals auch von n abhängt, wie in unserem Fall. Für AR(1)-Frailties wäre Sie jedoch nicht gut geeignet, da sie dritte und vierte Ableitungen der Funktion $g(\cdot)$ benötigt, die hier aber verschwinden. Alternativen sind also die Wahl einer anderen Approximation des Integrals wie z. B. mit Hilfe der Simpsonregel oder einer Taylorentwicklung, ein einfacheres numerisches Verfahren als die bisher verwendeten oder ein völlig anderes Vorgehen. Ein Ansatz wäre hier, die Likelihood mit Hilfe des Innovationstheorems von Aalen (1978), wie z.B. in Nielsen et al. (1992) vorgeschlagen, zu berechnen. Hierfür müsste sich aber zunächst ein passendes Modell für den latenten Faktor finden. Die Verwendung des Innovationstheorems hätte vor allem den Vorteil, dass die Eigenschaften der so gefundenen Schätzer mit Hilfe von bekannten Regularitätsbedingungen für Zählprozesse überprüft werden können.

Literaturverzeichnis

- O.O. Aalen. Nonparametric inference for a family of counting processes. *The Annals of Statistics*, 6(4):701–726, 1978.
- E.I. Altman und D.L. Kao. Rating drift of high yield bonds. *Journal of Fixed Income*, 1992.
- P. K. Andersen, Ø. Borgan, R. D. Gill und N. Keiding. *Statistical models based on counting processes*. Springer, New York, 1993.
- A. Bangia, F.X. Diebold, A. Kronimus, C. Schlagen und T. Schuermann. Rating migration and the business cycle, with application to credit portfolio stress testing. *Journal of Banking and Finance*, 26:235–264, 2002.
- O. Barndorff-Nielsen und D.R. Cox. Edgeworth and saddle-point approximations with statistical applications. *Journal of the Royal Statistical Society B*, 41:279–312, 1979.
- C. Bluhm, L. Overbeck und C. Wagner. *An Introduction to Credit Risk Modeling*. Chapman & Hall, 2002.
- S. Boyd und L. Vanderberghe. *Convex Optimization*. Cambridge University Press, Cambridge, 2004.
- R.W. Butler. *Saddlepoint Approximations with Applications*. Cambridge University Press, Cambridge, 2007.
- R.W. Butler und A.T.A. Wood. Laplace approximation for hypergeometric functions of matrix argument. *The Annals of Statistics*, 30:1155–1177, 2002.

- S. Casjens, F. Kramer, T. Mollenhauer und R. Walter. Die optimale Kombination interner und externer Ratings. In W. von Schimmelmann und G. Franke, editors, *Interne und externe Ratings*, pages 123–162. Frankfurter Allgemeine Buch, Frankfurt, 2007.
- H.E. Daniels. Saddlepoint approximations in statistics. *The Annals of Mathematical Statistics*, 25(4):631–650, 1954.
- M. Delloye, J.-D. Fermanian und M. Sbai. Dynamic frailties and credit portfolio modelling. *Risk*, October:100–105, 2006.
- D. Duffie, A. Eckner, G. Horel und L. Saita. Frailty correlated default. *The Journal of Finance*, 64(5):2089–2123, 2009.
- M. Efstathiou, E. Gutierrez-Pena und A.F.M. Smith. Laplace approximations for natural exponential families with cuts. *Scandinavian Journal of Statistics*, 25(1):77–92, 1998.
- M. Fiocco, H. Putter und J.C. van Houwelingen. A new serially correlated gamma-frailty process for longitudinal count data. *Biostatistics*, 10(2):245–257, 2009.
- P. Gagliardini und C. Gouriéroux. Efficiency in large dynamic panel models with common factor. *Swiss Finance Institute Research Paper Series*, 09(12):1–50, 2008.
- R.D. Gill. Marginal partial likelihood. *Scandinavian Journal of Statistics*, 19(2):133–137, 1992.
- W. J. Glynn. Asymptotic representations of the densities of canonical correlations and latent roots in manova when the population parameters have arbitrary multiplicity. *The Annals of Statistics*, 8(5):958–976, 1980.
- M.C. Harding und J. Hausman. Using a Laplace approximation to estimate the random coefficients logit model by nonlinear least squares. *International Economic Review*, 48(4):1311–1328, 2007.

- M.T. Hasan, B.B. Sutradhar und G. Sneddon. On correlation models for longitudinal failure time data. *The Indian Journal of Statistics*, 69(3):548–580, 2007.
- R. Henderson und S. Shimakura. A serially correlated gamma frailty model for longitudinal count data. *Biometrika*, 90:355–366, 2003.
- P. Hougaard. Frailty models for survival data. *Lifetime Data Analysis*, 1:255–273, 1995.
- L.C. Hsu. A theorem on the asymptotic behavior of a multiple integral. *Duke Mathematical Journal*, pages 623–632, 1948.
- S. J. Koopman, Roman Kräussl, A. Lucas und A. Monteiro. Credit cycles and macro fundamentals. *Journal of Empirical Finance*, 16(1):42–54, 2009.
- S. J. Koopman, A. Lucas und A. Monteiro. The multi-state latent factor intensity model for credit rating transitions. *Journal of Econometrics*, 142:399–424, 2008.
- D. Lando und T. Skødeberg. Analyzing rating transitions and rating drift with continuous observations. *Journal of Banking and Finance*, 26:423–444, 2002.
- S. O. Manda und R. Meyer. Bayesian inference for recurrent events data using time-dependent frailty. *Statistics in Medicine*, 24:1263–1274, 2005.
- K. G. Manton, E. Stallard und J. W. Vaupel. Alternative models for the heterogeneity of mortality risks among the aged. *Journal of the American Statistical Association*, 81:635–644, 1986.
- S. A. Murphy. Asymptotic theory for the frailty model. *The Annals of Statistics*, 23:182–198, 1995.
- J. A. Nelder und R. Mead. A simplex algorithm for function minimization. *The Computer Journal*, 7:308–313, 1965.

- G. G. Nielsen, R. D. Gill, P. K. Andersen und T. I. A. Sorensen. A counting process approach to maximum likelihood estimation in frailty models. *Scandinavian Journal of Statistics*, 19:25–43, 1992.
- E. Parner. Asymptotic theory for the correlated gamma-frailty model. *The Annals of Statistics*, 26(1):183–214, 1998.
- A. Pickles und R. Crouchley. A comparison of frailty models for multivariate survival data. *Statistics in Medicine*, 14:1447–1461, 1995.
- J. C. Pinheiro und D. M. Bates. Approximations to the log-likelihood function in the nonlinear mixed-effects model. *Journal of Computational and Graphical Statistics*, 4(1):12–35, 1995.
- R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2011. URL <http://www.R-project.org>. ISBN 3-900051-07-0.
- P. J. Schönbucher. Information-driven default contagion. *Working Paper*, pages 1–27, 2003.
- C. Stefanescu, R. Tunarub und S. Turnbull. The credit rating process and estimation of transition probabilities: A Bayesian approach. *Journal of Empirical Finance*, 16(2):216–234, 2009.
- L. Tierney und J. Kadane. Accurate approximations for posterior moments and marginal densities. *Journal of the American Statistical*, 81(393):82–86, 1986.
- J.W. Vaupel, K.G. Manton und E. Stallard. The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography*, 16(3):439–454, 1979.
- R. Weißbach und T. Mollenhauer. Modelling rating transitions. *Journal of the Korean Statistical Society*, im Druck, 2011.

-
- R. Weißbach und R. Walter. A likelihood ratio test for stationarity of rating transitions. *Journal of Econometrics*, 155(2), 2010.
- R. Wolfinger. Laplace's approximation for nonlinear mixed models. *Biometrika*, 80(4):791–795, 1993.
- K. K. W. Yau und C. A. McGilchrist. ML and REML estimation in survival analysis with time dependent correlated frailty. *Statistics in Medicine*, 17(11):1201–1213, 1998.
- H. Yue und K. S. Chan. A dynamic frailty model for multivariate survival data. *Biometrics*, 53(3):785–793, 1997.

A. R-Code

A.1. Simulationen

```
#####betaverteilte Frailty#####
##Verallgemeinerte Betaverteilung
rbeta2 <- function(n,p=1,q=1,a=0, b=1){
return((b-a)*rbeta(n,p,q)+a)
}

beta_NR<-function(p=20000, n=100,k=3, q=0.4 , jahr=7, alpha=2){

#Frailtyvektor
xi<-numeric()
xi[1]<-rbeta2(1,alpha,alpha,-1,1)
#Intensitätsmatrix
Q<-matrix(c(0,q*(1-xi),0,q*(1+xi),0,q*(1-xi)), ncol=3, byrow=TRUE)
X<-matrix(numeric(n), ncol=1)           #wie lange ist Historie
klassex<-matrix(numeric(n), ncol=1)    #in welcher Ratingklasse
abstand<-numeric()                     # Zeit zwischen t_i und t_i+1

###ziehen erster Zufallszahlen
m<-0
if(k>2){
for(h in 1:(k-2)){
X[(m+1):(m+floor(n/(k-1))),1] <-rexp(floor(n/(k-1)),sum(Q[h,]))
```

```
###in welcher Klasse
klassex[(m+1):(m+floor(n/(k-1))),1]<-h
m<-m+floor(n/(k-1))
}}
X[(m+1):n,1] <-rexp((n-m),sum(Q[(k-1),]))
klassex[(m+1):n,1] <-k-1

##Wann migriert der erste
t<-min(X)
abstand[1]<-t
c<-which(X==min(X))
h<-klassex[c]
klassex<-cbind(klassex,klassex)
X[-c]<-t

X2<-X          #wie lange in Ratingklasse in diesem Zeitintervall

###weitere Historie
j<-2
repeat{
X2<-cbind(X2,NA)
X<-cbind(X,NA)

neu<-rmultinom(1,size=1,prob=Q[h,]/sum(Q[h,]))
klassex[c,j]<-which(neu==1)

xi[j]<-rbeta2(1,alpha,alpha,-1,1)
Q<-
```

```
matrix(c(0,q*(1-xi[j]),0,q*(1+xi[j]),0,q*(1-xi[j])), ncol=3, byrow=TRUE)
for(h in 1:(k-1)){
X2[which(klassex[,j]==h),j]<-
rexp(length(which(klassex[,j]==h)),sum(Q[h,]))
}
X[,j]<-X[,j-1]+X2[,j]
if(length(which(is.na(X[,j])))>0){
c<-which(X[,j]==min(X[-which(is.na(X[,j])),j]))
}else{
c<-which(X[,j]==min(X[,j]))
}
t<-X2[c,j]
abstand[j]<-t
h<-klassex[c,j]
X2[-c(c,which(is.na(X[,j])),j)]<-t
X[,j]<-X[,j-1]+X2[,j]
klassex<-cbind(klassex,klassex[,j])
if(length(which(X[,j]<jahr))==0)break
j<-j+1
}

####Zählprozess berechnen#####

tau<-j

Nplusm<-numeric(j)
Nminusm<-numeric(j)
Yplus<-numeric(j)
Yminus<-numeric(j)
```

```
for(d in 1:j){
  if(any(is.na(X2[which(klassex[,d]%in%1:k),d]))){
    Yminus[d]<-
    sum(X2[which(klassex[,d]%in%1:k),d]
    [-which(is.na(X2[which(klassex[,d]%in%1:k),d]))])
    Yplus[d]<-
    sum(X2[which(klassex[,d]%in%2:k),d]
    [-which(is.na(X2[which(klassex[,d]%in%2:k),d]))])
  }else{
    Yminus[d]<-sum(X2[which(klassex[,d]%in%1:k),d])
    Yplus[d]<-sum(X2[which(klassex[,d]%in%2:k),d])
  }
}
Y1<-Yminus-Yplus

for(d in 1:(j)){
  Nplusm[d]<-length(which(klassex[,d]==klassex[,d+1]+1))
  Nminusm[d]<-length(which(klassex[,d]==klassex[,d+1]-1))
}
Nplus<-c(0,cumsum(Nplusm))
Nminus<-c(0,cumsum(Nminusm))

#####Schätzen#####

####Suchalgorithmus
ha<-10^-10 #Hilfszahl für zweite Ableitung (numerisch)
```

##1. Ableitung bei bekanntem alpha

```
fuq<-function(x){
a<-(alpha0-1)/(x^2*Y1)
b<-(alpha0-1)^2/(x^3*Y1^2*sqrt((alpha0-1)^2/(x^2*Y1^2)+1))
c<-1+sqrt((alpha0-1)^2/(x^2*Y1^2)+1)-(alpha0-1)/(x*Y1)
d<-1-sqrt((alpha0-1)^2/(x^2*Y1^2)+1)+(alpha0-1)/(x*Y1)
return(sum(Nplusm+Nminusm)/x
+sum((2*(a-b)/c^3-2*(a-b)/d^3)/(2/c^2+2/d^2))
+sum(Nplusm*(a-b)/c-Nminusm*(a-b)/d)
+sum(-x*(a-b)*Yplus-c*Yplus+x*(a-b)*Yminus-d*Yminus)
+(alpha0-1)*sum((a-b)/c-(a-b)/d))}
```

##1. Ableitung bei bekanntem q

```
fua<-function(y, qn=q){
e<-1/(qn*Y1)
f<-(y-1)/(qn^2*Y1^2*sqrt((y-1)^2/(qn^2*Y1^2)+1))
c<-1+sqrt((y-1)^2/(qn^2*Y1^2)+1)-(y-1)/(qn*Y1)
d<-1-sqrt((y-1)^2/(qn^2*Y1^2)+1)+(y-1)/(qn*Y1)
return(-log(2)+tau*digamma(2*y)-2*tau*digamma(y)-tau/(2*(y-1))
+sum((2*(e-f)/d^3-2*(e-f)/c^3)/(2/c^2+2/d^2)
)+sum(-Nplusm*(e-f)/c+Nminusm*(e-f)/d)
+qn*sum((e-f)*Yplus-(e-f)*Yminus)
+sum(log(c)+log(d))+y*sum((e-f)/d-(e-f)/c))}
```

###1. Ableitung für beide Parameter

```
fu<-function(x,y){
e<-1/(x*Y1)
f<-(y-1)/(x^2*Y1^2*sqrt((y-1)^2/(x^2*Y1^2)+1))
a<-(y-1)/(x^2*Y1)
```

```

b<-(y-1)^2/(x^3*Y1^2*sqrt((y-1)^2/(x^2*Y1^2)+1))
c<-1+sqrt((y-1)^2/(x^2*Y1^2)+1)-(y-1)/(x*Y1)
d<-1-sqrt((y-1)^2/(x^2*Y1^2)+1)+(y-1)/(x*Y1)
return(c(sum(Nplusm+Nminusm)/x
+sum((2*(a-b)/c^3-2*(a-b)/d^3)/(2/c^2+2/d^2))
+sum(Nplusm*(a-b)/c-Nminusm*(a-b)/d)
+sum(-x*(a-b)*Yplus-c*Yplus+x*(a-b)*Yminus-d*Yminus)
+(y-1)*sum((a-b)/c-(a-b)/d),-log(2)
+tau*digamma(2*y)-2*tau*digamma(y)-tau/(2*(y-1))
+sum((2*(e-f)/d^3-2*(e-f)/c^3)/(2/c^2+2/d^2))
+sum(-Nplusm*(e-f)/c+Nminusm*(e-f)/d)
+x*sum((e-f)*Yplus-(e-f)*Yminus)+sum(log(c)+log(d))
+y*sum((e-f)/d-(e-f)/c)))}

#Startwerte
q0<-sum(Nplusm+Nminusm)/sum(Yminus+Yplus)
qsch0<-sum(Nplusm+Nminusm)/sum(Yminus+Yplus)
alpha0<-alpha
repeat{
##q: 2. Ableitung numerisch
J<-(fuq(q0+ha)-fuq(q0-ha))/(2*ha)
parameter<-q0-1/J*fuq(q0)
if(abs(parameter-q0)<10^-6)break
q0<-parameter
}
qsch1<-parameter

#Startwerte##alpha: 2. Ableitung numerisch
alpha0<-alpha

```

```

repeat{
  J<-(fua(alpha0+ha)-fua(alpha0-ha))/(2*ha)
  parameter<-alpha0-1/J*fua(alpha0)
  if(abs(parameter-alpha0)<10^-6)break
  if(parameter<=1){t<-1
  repeat{t<-1/2*t
  parameter<-alpha0-t*1/J*fua(alpha0)
  if(parameter>1)break}}
  alpha0<-parameter
}
alphasch1<-parameter

#Startwerte###beide: 2. Ableitung numerisch
q0<-sum(Nplusm+Nminusm)/sum(Yminus+Yplus)
alpha0<-alpha
repeat{
  J<-matrix(c((fu(x=q0+ha,y=alpha0)-fu(x=q0-ha,y=alpha0))[1]/(2*ha),
  (fu(x=q0,y=alpha0+ha)-fu(x=q0,y=alpha0-ha))[1]/(2*ha),
  (fu(x=q0,y=alpha0+ha)-fu(x=q0,y=alpha0-ha))[1]/(2*ha),
  (fu(x=q0,y=alpha0+ha)-fu(x=q0,y=alpha0-ha))[2]/(2*ha)),
  ncol=2,byrow=TRUE)
  parameter<-c(q0,alpha0)-ginv(J)%*%fu(x=q0,y=alpha0)
  if(sum(abs(parameter-c(q0,alpha0)))<10^-6)break
  if(parameter[1]<0 | parameter[2]<=1){
  t<-1
  repeat{
  t<-1/2*t
  parameter<-c(q0,alpha0)-t*ginv(J)%*%fu(x=q0,y=alpha0)

```

```

if(parameter[1]>=0 & parameter[2]>1)break
}}
q0<-parameter[1]
alpha0<-parameter[2]
}
qsch2<-parameter[1]
alphasch2<-parameter[2]

return(c(qsch0, qsch1, qsch2, alphasch1, alphasch2))
}

#####
#####AR(1)-Fraitly#####
normal_NR<-function( n=100,k=3, q=0.4 , jahr=7, sigma=0.05,rho=0.9){

Q<-matrix(c(0,q,0,q,0,q), ncol=3, byrow=TRUE)
xi<-0
X<-matrix(numeric(n), ncol=1)
klassex<-matrix(numeric(n), ncol=1)
abstand<-numeric()
#ziehen erster Zufallszahlen
m<-0
if(k>2){
for(h in 1:(k-2)){
X[(m+1):(m+floor(n/(k-1))),1] <-rexp(floor(n/(k-1)),sum(Q[h,]))
#in welcher Klasse
klassex[(m+1):(m+floor(n/(k-1))),1]<-h
m<-m+floor(n/(k-1))

```

```
}}
X[(m+1):n,1] <-rexp((n-m),sum(Q[(k-1),]))
klassex[(m+1):n,1] <-k-1

t<-min(X)
abstand[1]<-t
c<-which(X==min(X))
h<-klassex[c]
klassex<-cbind(klassex,klassex)
X[-c]<-t

X2<-X
rhoneu<-c(rho,1)
#weitere Historie
j<-2
repeat{
X2<-cbind(X2,NA)
X<-cbind(X,NA)

neu<-rmultinom(1,size=1,prob=Q[h,]/sum(Q[h,]))
klassex[c,j]<-which(neu==1)

xi[j]<-rnorm(1,sd=sqrt(sigma*t))
Q<-matrix(c(0,q*(1-sum(rhoneu*xi)),0,
q*(1+sum(rhoneu*xi)),0,q*(1-sum(rhoneu*xi))),ncol=3, byrow=TRUE)
for(h in 1:(k-1)){
X2[which(klassex[,j]==h),j]<-
rexp(length(which(klassex[,j]==h)),sum(Q[h,]))
```

```
}
X[,j]<-X[,j-1]+X2[,j]
if(length(which(is.na(X[,j])))>0){
c<-which(X[,j]==min(X[-which(is.na(X[,j])),j]))
}else{
c<-which(X[,j]==min(X[,j]))
}
t<-X2[c,j]
abstand[j]<-t
h<-klassex[c,j]
X2[-c(c,which(is.na(X[,j])),j)<-t
X[,j]<-X[,j-1]+X2[,j]
if(length(which(X[,j]<jahr))==0)break
klassex<-cbind(klassex,klassex[,j])
rhoneu<-c(rho^j,rhoneu)
j<-j+1
}

####Zählprozess berechnen#####

tau<-j

Nplusm<-numeric(j-1)
Nminusm<-numeric(j-1)
Yplus<-numeric(j)
Y1<-numeric(j)
Yminus<-numeric(j)

for(d in 1:j){
```

```
Yminus[d]<-
sum(X2[which(klassex[,d]%in%1:k),d]
[-which(is.na(X2[which(klassex[,d]%in%1:k),d]))])
Yplus[d]<-
sum(X2[which(klassex[,d]%in%2:k),d]
[-which(is.na(X2[which(klassex[,d]%in%2:k),d]))])
}
Y1<-Yminus-Yplus
for(d in 1:(j-1)){
Nplusm[d]<-length(which(klassex[,d]==klassex[,d+1]+1))
Nminusm[d]<-length(which(klassex[,d]==klassex[,d+1]-1))
}
Nplus<-c(0,cumsum(Nplusm))
Nminus<-c(0,cumsum(Nminusm))

#####Schätzen#####
#Startwerte
q0<-sum(Nplusm+Nminusm)/sum(Yminus+Yplus)
qsch1<-sum(Nplusm+Nminusm)/sum(Yminus+Yplus)
sigma2<-sigma
rhon<-rho

####Suchalgorithmus

repeat{
yepsilon<-numeric()
yepsilon2<-numeric()
yepsilon3<-numeric()
}
```

```

yepsilon4<-numeric()
yepsilon5<-numeric()
yepsilon6<-numeric()
yepsilon7<-numeric()
yepsilon8<-numeric()
yepsilon9<-numeric()
for(d in 2:tau){
yepsilon[d-1]<-sum(rhon^(0:(tau-d))*(Y1)[d:tau])
yepsilon3[d-1]<-sum(0:(tau-d)*rhon^((0:(tau-d))-1)*(Y1)[d:tau])
yepsilon4[d-1]<-
sum(0:(tau-d)*((0:(tau-d))-1)*rhon^((0:(tau-d))-2)*(Y1)[d:tau])
}

#####
for(d in 2:tau){
yepsilon2[d-1]<-sum(rhon^((d-2):0)*abstand[1:(d-1)]*yepsilon[1:(d-1)])
yepsilon5[d-1]<-sum(rhon^((d-2):0)*abstand[1:(d-1)]*yepsilon3[1:(d-1)])
yepsilon6[d-1]<-sum(rhon^((d-2):0)*abstand[1:(d-1)]*yepsilon4[1:(d-1)])
yepsilon7[d-1]<-
sum(((d-2):0)*rhon^(((d-2):0)-1)*abstand[1:(d-1)]*yepsilon[1:(d-1)])
yepsilon8[d-1]<-
sum(((d-2):0)*rhon^(((d-2):0)-1)*abstand[1:(d-1)]*yepsilon3[1:(d-1)])
yepsilon9[d-1]<-sum(((d-2):0)*(((d-2):0)-1)*rhon^(((d-2):0)-2)
*abstand[1:(d-1)]*yepsilon[1:(d-1)])
}

##q
fu<-sum(Nplusm+Nminusm)/q0-Yplus[1]-Yminus[1]

```

```

-sigma2*q0*sum(abstand[1:(tau-1)]*(yepsilon)^2)
-sum((1+2*sigma2*q0*yepsilon2)*Yplus[-1])
-sum((1-2*sigma2*q0*yepsilon2)*Yminus[-1])
+sum(-((Nminus[2:tau]-Nminus[1:(tau-1)])*sigma2*yepsilon2)
/(1-sigma2*q0*yepsilon2)
+((Nplus[2:tau]-Nplus[1:(tau-1)])*sigma2*yepsilon2)
/(1+sigma2*q0*yepsilon2))
J<--sum(Nplusm+Nminusm)/q0^2
-sigma2*sum((abstand[-tau]*yepsilon)^2)
-sum(2*sigma2*yepsilon2*Yplus[-1])
+sum(2*sigma2*yepsilon2*Yminus[-1])
-sum(((Nminus[2:tau]-Nminus[1:(tau-1)])*(sigma2*yepsilon2)^2)
/(1-sigma2*q0*yepsilon2)^2)
+((Nplus[2:tau]-Nplus[1:(tau-1)])*(sigma2*yepsilon2)^2)
/(1+sigma2*q0*yepsilon2)^2)
parameter<-q0-1/J*fu

if(abs(parameter-q0)<10^-6)break
q0<-parameter

}

qsch2<-parameter

return(c(qsch1,qsch2))
}

```

A.2. Datenauswertung

```
daten1<-
read.csv2("D:\\Benutzer\\Walter\\Diss\\Daten\\WestLB\\GPortfolio.csv",
  header=TRUE)
  #alle bis 30.06.1999 raus
rating22<-as.Date(daten1$Next_Ratingdate_art, format="%d.%m.%Y")
daten<-
daten1[-which(rating22<=as.Date("01.07.1999",format="%d.%m.%Y" )),]
attach(daten)

Kunden<-unique(CustID)           #Welche Kunden
m<-length(Kunden)

rating1<-as.Date(Ratingdate_art, format="%d.%m.%Y") #In Datumsformat
rating1[which(rating1<=as.Date("30.06.1999",format="%d.%m.%Y" ))]<-
as.Date("1999-07-01")

rating2<-as.Date(Next_Ratingdate_art, format="%d.%m.%Y")

###erstes Portfoliojahr streichen

for(k in 1:m){
erstrating<-min(rating1[which(CustID==Kunden[k])])
  ##erstes Rating nach Beginn der Beobachtung?
if(erstrating>=as.Date("02.07.1999",format="%d.%m.%Y" )){
```

```
##keine Veränderung im ersten Jahr
if(
rating2[which(rating1==erstrating&CustID==Kunden[k])]
[length(which(rating1==erstrating&CustID==Kunden[k]))]
>(erstrating+365)){
rating1[which(rating1==erstrating&CustID==Kunden[k])]<-
erstrating+365
}else{      ##Veränderungen im ersten Jahr
x<-which((rating2<=(erstrating+365)) &(daten$CustID==Kunden[k]))
detach(daten)
daten<-daten[-x,]
attach(daten)
rating1<-rating1[-x]
rating2<-rating2[-x]
if(length(which((CustID==Kunden[k])&(rating2>(erstrating+365))))>0){
erstrating2<-
min(rating2[which(CustID==Kunden[k]&rating2>(erstrating+365))])
rating1[which(rating2==erstrating2 & (CustID==Kunden[k]))]<-
erstrating+365
}}
}}
detach(daten)
daten<-daten[-which(rating1==rating2)]
attach(daten)

rating[which(rating%in%1:4)]<-5
next_rating[which(rating%in%1:4)]<-5

n<-length(rating)
```

```
#####
##### Berechnung Zählprozesse #####
o<-order(rating2)
datenneu<-daten[o,]
tage<-unique(sort(rating2[which(next_rating!=999 &
  (rating==(next_rating+1) | rating==(next_rating-1))]))
tage1<-c(min(rating1),tage)
tau<-length(tage)
abstand<-as.numeric(tage[1:tau]-c(min(rating1),tage[1:(tau-1)]))/365
h<-unique(sort(next_rating))

Nplus<-numeric()
Nminus<-numeric()
Yplus<-numeric()
Yminus<-numeric()
Y1<-numeric()

for(i in 1:tau){
  Nplus[i]<-length(which(rating==(next_rating+1)
    & rating2<=tage[i] & next_rating!=999))
  Nminus[i]<-length(which(rating==(next_rating-1)
    & rating2<=tage[i] & next_rating!=999))
}
Nplusm<-Nplus-c(0,Nplus[-tau])
Nminusm<-Nminus-c(0,Nminus[-tau])

for(i in 2:(tau+1)){
```

```

Yminus[i-1]<-length(which(
rating==5 & rating1<=tage1[i-1]&rating2>=tage1[i]))*abstand[i-1]
+ sum(as.numeric(rating2[which(rating==5&rating1<=tage1[i-1]&
rating2>tage1[i-1]&rating2<tage1[i])]-tage1[i-1]))/365
+sum(as.numeric(tage1[i]-rating1[which(rating==5&rating1>tage1[i-1]&
rating1<tage1[i]&rating2>=tage1[i])]))/365
+sum(as.numeric(rating2-rating1)[which(rating==5&rating1>tage1[i-1]&
rating1<tage1[i]&rating2<tage1[i]&rating2>tage1[i-1])]))/365
Y1[i-1]<-length(which(rating==5 & rating1<=tage1[i-1]&
rating2>=tage1[i]))*abstand[i-1]
+ sum(as.numeric(rating2[which(rating==5&rating1<=tage1[i-1]&
rating2>tage1[i-1]&rating2<tage1[i])]-tage1[i-1]))/365+
sum(as.numeric(tage1[i]-rating1[which(rating==5&rating1>tage1[i-1]&
rating1<tage1[i]&rating2>=tage1[i])]))/365
+sum(as.numeric(rating2-rating1)[which(rating==5&rating1>tage1[i-1]&
rating1<tage1[i]&rating2<tage1[i]&rating2>tage1[i-1])]))/365
Yplus[i-1]<-0
for(j in 6:20){
Yplus[i-1]<-
Yplus[i-1]+ length(which(rating==j & rating1<=tage1[i-1]&
rating2>=tage1[i]))*abstand[i-1]
+sum(as.numeric(rating2[which(rating==j&rating1<=tage1[i-1]&
rating2>tage1[i-1]&rating2<tage1[i])]-tage1[i-1]))/365+
sum(as.numeric(tage1[i]-rating1[which(rating==j&rating1>tage1[i-1]&
rating1<tage1[i]&rating2>=tage1[i])]))/365
+sum(as.numeric(rating2-rating1)[which(rating==j&rating1>tage1[i-1]&
rating1<tage1[i]&rating2<tage1[i]&rating2>tage1[i-1])]))/365
Yminus[i-1]<-
Yminus[i-1]+length(

```

```

which(rating==j & rating1<=tage1[i-1]&rating2>=tage1[i]))*abstand[i-1]
+sum(as.numeric(rating2[which(rating==j&rating1<=tage1[i-1]&
rating2>tage1[i-1]&rating2<tage1[i])]-tage1[i-1]))/365+
sum(as.numeric(tage1[i]-rating1[which(rating==j&
rating1>tage1[i-1]&rating1<tage1[i]&rating2>=tage1[i])]))/365
+sum(as.numeric(rating2-rating1)[which(rating==j&rating1>tage1[i-1]&
rating1<tage1[i]&rating2<tage1[i]&rating2>tage1[i-1])]))/365
}}

```

```

### Relative Migrations und Ausfallhäufigkeiten pro Jahr##
jahr<-sort(unique(substring(as.character(c(rating1,rating2)),1,4)))
Nplusneu<-numeric(length(jahr))
Nminusneu<-numeric(length(jahr))
Ausfall<-numeric(length(jahr))
anzahl<-numeric(length(jahr))
for(j in 1:length(jahr)){
a<-which(tage<=as.Date(paste( jahr[j],"-12-31", sep="")))&
tage>=as.Date(paste( jahr[j],"-01-01", sep="")))
b<-which((rating1<=as.Date(paste( jahr[j],"-12-31", sep=""))&
rating1>=as.Date(paste( jahr[j],"-01-01", sep="")))|
(rating2<=as.Date(paste( jahr[j],"-12-31", sep=""))&
rating2>=as.Date(paste( jahr[j],"-01-01", sep="")))|
(rating2>=as.Date(paste( jahr[j],"-12-31", sep=""))&
rating1<=as.Date(paste( jahr[j],"-01-01", sep=""))))
Nplusneu[j]<-sum(Nplusm[a])
Nminusneu[j]<-sum(Nminusm[a])
Ausfall[j]<-length(which(next_rating==21&
rating2<=as.Date(paste( jahr[j],"-12-31", sep=""))

```

```
&rating2>=as.Date(paste( jahr[j], "-01-01", sep=""))))  
anzahl[j]<-length(unique(CustID[b]))  
}
```