KTH Engineering Sciences

Topics in life and disability insurance

BJÖRN LÖFDAHL GRELSSON

Doctoral Thesis
Stockholm, Sweden 2015

TRITA-MAT-A 2015:09
ISRN KTH/MAT/A-15/09-SE
ISBN 978-91-7595-701-2

Department of Mathematics
KTH Royal Institute of Technology
100 44 Stockholm, Sweden

Akademisk avhandling som med tillstånd av Kungl Tekniska Högskolan framlägges till offentlig granskning för avläggande av teknologie doktorsexamen, fredagen den 6 november 2015 klockan 13.15 i sal F3, Lindstedtsvägen 26, Kungl Tekniska Högskolan, Stockholm.

Print: Universitetsservice US AB, Stockholm, 2015

Till Gubben...

...om han finns. Annars är det kört!

- Boulle Boumpa

Abstract

This thesis consists of five papers, presented in Chapters A-E, on topics in life and disability insurance. It is naturally divided into two parts, where papers A and B discuss disability rates estimation based on historical claims data, and papers C-E discuss claims reserving, risk management and insurer solvency.

In Paper A, disability inception and recovery probabilities are modelled in a generalized linear models (GLM) framework. For prediction of future disability rates, it is customary to combine GLMs with time series forecasting techniques into a two-step method involving parameter estimation from historical data and subsequent calibration of a time series model. This approach may in fact lead to both conceptual and numerical problems since any time trend components of the model are incoherently treated as both model parameters and realizations of a stochastic process. In Paper B, we suggest that this general two-step approach can be improved in the following way: First, we assume a stochastic process form for the time trend component. The corresponding transition densities are then incorporated into the likelihood, and the model parameters are estimated using the Expectation-Maximization algorithm.

In Papers C and D, we consider a large portfolio of life or disability annuity policies. The policies are assumed to be independent conditional on an external stochastic process representing the economic-demographic environment. Using the Conditional Law of Large Numbers (CLLN), we establish the connection between claims reserving and risk aggregation for large portfolios. Moreover, we show how statistical multi-factor intensity models can be approximated by one-factor models, which allows for computing reserves and capital requirements efficiently. Paper C focuses on claims reserving and ultimate risk, whereas the focus of Paper D is on the one-year risks associated with the Solvency II directive.

In Paper E, we consider claims reserving for life insurance policies with reserve-dependent payments driven by multi-state Markov chains. The associated prospective reserve is formulated as a recursive utility function using the framework of backward stochastic differential equations (BSDE). We show that the prospective reserve satisfies a nonlinear Thiele equation for Markovian BSDEs when the driver is a deterministic function of the reserve and the underlying Markov chain. Aggregation of prospective reserves for large and homogeneous insurance portfolios is considered through mean-field approximations. We show that the corresponding prospective reserve satisfies a BSDE of mean-field type and derive the associated nonlinear Thiele equation.

Sammanfattning

Den här avhandlingen består av fem vetenskapliga artiklar, vilka presenteras i kapitel A-E. Avhandlingen är naturligt uppdelad i två delar, där artiklarna A och B behandlar skattning av populationssjuklighet, och artiklarna C, D och E behandlar reservsättning och riskberäkningar för stora portföljer.

I den första artikeln föreslår vi ett stokastiskt semi-Markovskt ramverk för modellering av sjuklighet i diskret tid. Med hjälp av räkneprocesser och generaliserade linjära modeller beskriver vi de logistiska transformerna av insjuknande- och avvecklingssannolikheterna i termer av basfunktioner och stokastiska riskfaktorer.

Den andra artikeln behandlar modellering av framtida sjuklighet. Ett i litteraturen vedertaget angreppssätt består i att anpassa generaliserade linjära modeller till historiskt data med en efterföljande kalibrering av en tidsseriemodell. Detta angreppssätt kan dock leda till både numeriska och konceptuella problem, då modellens tidstrend i det första skedet behandlas som parametrar, för att i det andra skedet behandlas som realisationer av en stokastisk process. Vi föreslår en metod där tidstrenden antas vara en dold Markovprocess och skattar modellens parametrar med EM-algoritmen.

I den tredje artikeln betraktar vi en stor, homogen portfölj av försäkringskontrakt, och antar oberoende mellan populationens individer betingat på en stokastisk process som representerar omvärldens tillstånd. Med hjälp av betingade stora talens lag etablerar vi relationen mellan reservsättning och riskaggregering för stora portföljer. Vi visar att alla moment för nuvärdet av portföljens kassaflöden kan beräknas genom att lösa en uppsättning partiella differentialekvationer. Vidare visar vi hur statistiska flerfaktormodeller kan approximeras med enfaktormodeller, vilket medför att differentialekvationerna kan lösas mycket effektivt.

I den fjärde artikeln utvidgar vi modellen från den tredje artikeln. Genom att använda betingade stora talens lag visar vi att om en försäkringsportfölj blir tillräckligt stor kan portföljens värde om ett år approximeras med en funktional av omvärldsprocessen. Vi härleder ett approximativt uttryck för portföljens kvantiler under antaganden om en homogen portfölj under en enfaktormodell. Vidare föreslår vi två beräkningsmetoder för riskaggregering när portföljen består av stora, homogena pooler.

I den femte artikeln föreslår vi ett enhetligt angreppssätt för reservsättning av liv- och sjukförsäkringar där de kontraktuella betalningarna tillåts bero av reserven i sig. Reserven formuleras som en rekursiv nyttofunktion inom ramverket för bakåtriktade stokastiska differentialekvationer (BSDE), och vi visar att den löser en icke-linjär Thiele-ekvation i det Markovska fallet. Aggregering av reserver för stora, homogena portföljer formuleras med hjälp av mean-field approximationer. Vi visar att reserven satisfierar en BSDE av mean-field typ och härleder dess tillhörande Thiele-ekvation.

Acknowledgements

I am immensely grateful to my advisor Professor Boualem Djehiche for his compact support and continuous encouragement, as well as for many tough and productive research meetings during my time as a PhD student. Further, I would like to express my gratitude to my co-advisor Professor Henrik Hult for many fruitful discussions and advice, especially regarding the third and fourth papers.

I want to thank my former and current room mates Thorbjörn Gudmundsson, Johan Nykvist and Johan Westerborn for many good laughs, interesting discussions and competitive chess and table tennis sessions.

Furthermore, I want to thank Jimmy Olsson, Johan Westerborn and Jonas Hallgren for helpful discussions regarding the second paper, and Ali Hamdi for comments regarding the introduction.

I gratefully acknowledge financial support from the Filip Lundberg and Eir's 50 Years foundations led by Patrik Dahl at Sirius and Åsa Larsson at Finansinspektionen.

Finally, I wish to thank Gunnar Andersson and his team at Folksam and Kristina Hagström and Lars Carlsson at Skandia Liv for providing me with data and for valuable discussions regarding this project.

Stockholm, November 2015

Björn Löfdahl Grelsson

Contents

Introduction

In this introduction, we give a brief overview of selected topics in life and disability insurance, including the estimation of mortality and disability rates, claims reserving and risk management. The section is concluded with a summary of the five appended papers.

1 Background

A health and disability insurance policy entitles to a monthly benefit, in compensation for a reduction in or loss of income due to sickness or accident. The purpose of the policy is to protect the policy holder from the economic loss that arises when the holder is unable to work. Typically, the public social security systems provide a basic level of protection. To receive additional protection, an individual may purchase a policy from an insurance company.

While the economic loss related to prolonged disability may be unbearable for a single individual, it is feasible to distribute the cost over a large group of individuals. The risk is transferred from the individual to the collective, essentially represented by an insurance company, by having each individual pay a *premium* to the company in relation to the expected cost of disability. In the case of illness, the individual receives benefits from the insurer, who sets aside a *reserve* corresponding to the expected cost of disability.

An insurance company is exposed to many different types of risk. *Premium risk* is the risk that the benefits payed by the insurer exceeds the received premiums. *Reserve risk* is the risk that the payed benefits exceeds the insurer's reserves. If the insurer systematically misprices its products it may suffer economic losses, ultimately leading to the insurer facing *ruin*, i.e. not being able to fulfill its obligations to the policy holders.

In order to determine the premiums and reserves associated with a disability insurance policy, the insurer needs predictions of the future rates of disability onset and recovery, as well as an estimate of the delay from claim occurrence to reporting. Predictions can be obtained by analyzing historical data, such as claims histories. If a collective is large enough and if the insurer manages to correctly price its products, the sum of payed benefits should be equal to the sum of all received premiums due to diversification. That is, the *idiosyncratic*, or *diversifiable,* risk, is negligible.

It must be stressed that the insurer still faces *systematic*, or *undiversifiable*, risk, i.e. the risk that the realized future biometric rates, such as the mortality and disability rates, do not equal the predicted rates, which may result in an economic loss for the insurer. Hence, care must be taken to estimate and manage the systematic risk. This is one of the key questions in the upcoming Solvency II regulatory framework. In particular, the new regulations suggest a new mindset regarding the valuation and risk management of insurance products. The directive's focus is on quantifying *one-year risks*, i.e. the risk that the value of the insurer's liabilities exceeds the value of its assets on a one-year time horizon, in which case the insurer would be deemed *insolvent*.

Historically, premiums and reserves are calculated under the assumption that the underlying transition intensities of death, disability onset, recovery and so on are deterministic. While the estimations should be prudent, this still implies that the systematic risk, i.e. the risk arising from uncertainty of the future development of the hazard rates, is not quantified properly. This may have an impact on pricing as well as on capital charges. For a life insurance company, the value of the liabilities depends on interest rates as well as on biometric rates, such as mortality and disability rates, of the insured population. Hence, obtaining reliable estimates of these rates is of the utmost importance.

2 Estimation of mortality and disability rates

Perhaps the most popular model for mortality rates is the Lee-Carter model [18] and its extensions. In its most basic form, the mortality $\mu_{x,t}$ for an x-year old at time t is assumed to be of the form

$$\log \mu_{x,t} = \alpha_x + \beta_x \kappa_t, \tag{1}$$

where α_x represents the average log-mortality over time, β_x represents the patterns of mortality change, and κ_t represents a time trend. The original model is fitted to observed mortality rates using singular value decomposition, so that $\alpha := (\alpha_x)_x$, $\beta := (\beta_x)_x$ and $\kappa := (\kappa_t)_t$ are estimated by

$$(\widehat{\alpha}, \widehat{\beta}, \widehat{\kappa}) = \arg\min_{\alpha,\beta,\kappa} \sum_{x,t} (\log \mu_{x,t} - \alpha_x - \beta_x \kappa_t)^2.$$

Predictions and simulations of future mortality rates are obtained by fitting a time series model to the estimated values of κ.

While the Lee-Carter model is easy to use and generally provides good fits to historical data, it has some drawbacks. In particular, the errors are assumed to be Gaussian, whereas the observed mortality originates from the number of deaths, which is a discrete random variable. Therefore, assuming a Poisson or binomial distribution seems more natural. Brouhns *et al.* [2] fit a model of the form (1) using maximum likelihood under Poisson assumptions, i.e.

$$D_{x,t} \sim \text{Po}(E_{x,t}\mu_{x,t}), \tag{2}$$

where $D_{x,t}$ denotes the number of deaths and $E_{x,t}$ denotes the number of individuals exposed to the risk of death. The parameters α, β and κ are estimated by maximizing the log-likelihood function $l(\alpha, \beta, \kappa; D_{\cdot,\cdot})$ given by

$$l(\alpha, \beta, \kappa; D_{\cdot,\cdot}) = \sum_{x,t} \Big(D_{x,t}(\alpha_x + \beta_x \kappa_t) - E_{x,t} e^{\alpha_x + \beta_x \kappa_t} \Big). \tag{3}$$

Again, predictions of future mortality rates are obtained by fitting a time series model to the estimated values of κ.

For disability insurance, the most important rates to estimate are the disability inception and recovery rates. While the modern literature for mortality modelling is very rich, there are few studies investigating the forecasting of disability rates [6]. Renshaw and Haberman [25] propose to model the recovery rates $\rho_{t,x,z}$ for an x-year old with disability duration z in year t as

$$\rho_{t,x,z} = A_{x,z} + B_{x,z} t, \tag{4}$$

where $A_{x,z}$ represents the recovery rates at the beginning of the observation period, and $B_{x,z}$ represents the patterns of change in recovery rates with respect to time. In contrast to the Lee and Carter model, specific functional forms of $A_{x,z}$ and $B_{x,z}$ are suggested, and the parameters are fitted to data using maximum likelihood under Poisson assumptions.

Christiansen *et al.* [3] consider a three-state disability model with state space $\mathcal{S} = \{a, i, d\}$, where a, i and d are the states 'active', 'invalid' and 'dead', respectively. The log-transition rate $\gamma_{jk}(x;t)$ between states j and k for an x-year old at time t is assumed to be of the form

$$\gamma_{jk}(x;t) = m_{jk}(x) + \phi_{jk}(x)\beta_{jk}(t),$$

where m_{jk} and ϕ_{jk} are deterministic functions of age and the vectors $\beta(t) = (\beta_{jk}(t))_{jk}$ form a multidimensional time series. The model parameters are fitted using the functional data approach of Hyndman and Ullah [14], which can be seen as a natural extension of the Lee and Carter methodology that allows for smoothing and replaces principal components with its functional counterpart.

The main drawback of the Renshaw and Haberman model (4) is that the future recovery rates are assumed to be deterministic, i.e. predictions of the future recovery rates for year τ are simply obtained by evaluating $\rho_{\tau,x,z}$, and any uncertainties are completely disregarded. In contrast, the Lee and Carter, Brouhns *et al.* and Christiansen *et al.* models readily supplement all predictions with prediction intervals, which are often used to determine the value of an insurer's liabilities in a 'worst-case' scenario. However, these prediction intervals may have an inherent bias, because the underlying assumptions are inconsistent in the following sense: At first, the yearly values of κ are considered to be parameters which can be estimated from data. In the next step, the assumption is altered so that the estimated yearly values of κ are seen as realizations of a stochastic process. As a result of this inconsistency, the volatility of κ tends to be overestimated, which may have a significant impact on pricing and risk management of insurance products. Qualitatively, this bias stems from the fact that yearly variations in the parameter values are caused by variations in the underlying process (systematic variation) as well as variations in the underlying population (idiosyncratic variation). This incoherence is unfortunately disregarded in most fitting procedures, as was already pointed out by Czado *et al.* [8]. Czado *et al.* propose to avoid this deficiency in the classical approach by integrating both steps into a Bayesian model, where the yearly values of the process corresponding to κ are all treated as random variables with given prior densities. The model parameters are then estimated using the Gibbs sampler and Metropolis-Hastings algorithm.

As an alternative, κ can be treated as an unobservable, or hidden, process. The model parameters can then be estimated in the framework of hidden Markov models using the Expectation-Maximization algorithm introduced by Dempster *et al.* [9]. Assume that κ is a Markov process with transition densities f parameterized by γ. As an example, κ could be a Brownian motion with Gaussian density f parameterized by γ given by the drift and diffusion coefficient. Further, assume that the transition rate is of the form (1) under the Poisson assumption (2), so that the set of free parameters to be estimated is $\theta = (\alpha, \beta, \gamma)$. Given observations $D_{x,1:n} := (D_{x,1}, \dots, D_{x,n})$, for x from a given set X of age groups, the algorithm proceeds iteratively in two alternating steps, the Expectation and Maximization steps:

For iteration $(k+1)$ of the EM-algorithm, given a parameter estimate θ^k, the E-step consists of integrating the complete data log-likelihood $l(\theta; D_{\cdot,1:n}, \kappa_{1:n})$ with respect to the distribution of $\kappa_{1:n} := (\kappa_1, \dots, \kappa_n)$ conditional on the observations

$D_{\cdot,1:n}$. That is, let

$$Q(\theta|\theta^k) = E^{\theta^k}[l(\theta; D_{\cdot,1:n}, \kappa_{1:n})|D_{\cdot,1:n}], \tag{5}$$

where

$$l(\theta; D_{\cdot,1:n}, \kappa_{1:n}) = l(\alpha, \beta, \kappa; D_{\cdot,1:n}) + \sum_{t=1}^{n} \log f_{\kappa_t|\kappa_{t-1}}(\gamma),$$

where $f_{\kappa_t|\kappa_{t-1}}$ denotes the density of κ_t given κ_{t-1} and $l(\alpha, \beta, \kappa; D_{\cdot,1:n})$ is given by (3). For many practical applications, the initial guess θ^0 has a significant impact on the estimation, and care must be taken to choose it in a suitable way. Further, direct integration of the complete data likelihood can be difficult, and (5) must often be approximated by simulation methods.

In the M-step, maximize Q w.r.t. θ to obtain

$$\theta^{k+1} = \arg\max_{\theta} Q(\theta|\theta^k).$$

It can be shown that the EM-algorithm yields successively improved parameter estimates in the sense that

$$Q(\theta^{k+1}|\theta^k) \geq Q(\theta^k|\theta^{k-1}), \quad k \geq 1.$$

If the likelihood function is bounded, this implies convergence of $(Q)_k$ to some Q^*. The convergence properties of the EM-algorithm were investigated by Wu [26]. It is difficult, in general, to say something about convergence of the sequence $(\theta^k)_k$ without introducing further assumptions.

While the hidden Markov approach solves the incoherence problem of the classical approach, it unfortunately introduces new problems. First, the difficulty with showing convergence to a global maximum is problematic, but this can be somewhat mitigated by choosing the initial guess carefully. For example, it could be chosen as the estimate from the classical approach. Second, the fact that the E-step often has to be performed using simulations increases the computational complexity of the estimation problem. However, biometric rates do not generally change by any measurable amount from one day to another, which means that re-estimation on a daily basis is not necessary for valuing insurance liabilities. Hence, increasing the computational cost of estimation is not a problem in practice.

3 Claims reserving and liability valuation

Consider an insurance policy with accumulated benefits less premiums over $(0, t]$ given by A_t. That is, A_t is the net sum of payments from the insurance company to the policy holder up to time t. These payments typically depend on the state of the policy holder, such as annuity payments in the case of illness or unemployment, a lump sum payment in the case of death, or a pension annuity until the death of the policy holder. The *prospective reserve* Y_t of this policy is often defined as the conditional expectation of the sum of all discounted future payments [21, 22], i.e.

$$Y_t = E\Big[\int_t^T e^{-\int_t^s \delta_u du}\, dA_s | \mathcal{F}_t\Big], \tag{6}$$

where δ_u denotes the discount rate, dA_s represents the incremental payments at time t, and $\mathcal{F}_t$ is a σ-algebra representing the information available up to time t.

Let X_t denote the state of the policy holder at time t, and let the state space $\mathcal{S}$ denote the set of all possible states of X. A common assumption is to let X be a Markov process on $\mathcal{S} = \{1, 2, \ldots, m\}$ with transition rates $\mu_{ij},\ i, j \in \mathcal{S}$. The incremental payments dA_t for a simple policy can often be decomposed into annuities $\alpha_i,\ i \in \mathcal{S}$, payable as long as $X_t = i$, and transition lump sum payments $\alpha_{ij},\ i, j \in \mathcal{S}$, to be paid immediately if X jumps from state i to state j. That is, dA_t can be written

$$dA_t = \sum_i I_i(t)\alpha_i(t)dt + \sum_i \sum_{j,\, j\neq i} \alpha_{ij}(t)\, dN_{ij}(t), \tag{7}$$

where I_i is the indicator process taking the value 1 if $X_t = i$, and 0 otherwise, and N_{ij} is a *counting process* that counts the number of transitions from state i to state j. A counting process $N = (N(t))_{t\geq 0}$ is a stochastic process starting from zero, with paths which are piecewise constant and non-decreasing, having jumps of size $+1$ only. The canonical example of a counting process is the Poisson process. The process N is said to have $\mathcal{F}$-*compensator* $\Lambda = (\Lambda(t))_{t\geq 0}$ if Λ is an $\mathcal{F}$-predictable process such that the *compensated process* $M = (M(t))_{t\geq 0}$ defined by

$$M(t) = N(t) - \Lambda(t)$$

is a martingale, that is, for $t \geq s$, we have

$$E[M(t)|\mathcal{F}_s] = M(s).$$

We say that N has $\mathcal{F}$-*intensity* $\lambda = (\lambda(t))_{t\geq 0}$ if λ is an $\mathcal{F}$-predictable process satisfying

$$\Lambda(t) = \int_0^t \lambda(s)ds, \quad t \geq 0.$$

Heuristically, we can write

$$\lambda(t) = \frac{1}{dt}\mathbb{P}(N_t - N_{t-dt} = 1|\mathcal{F}_{t-dt}),$$

which, roughly speaking, implies that $\lambda(t)dt$ is the probability that the process N jumps in the interval $(t-dt, t]$. A complete treatment of statistical models based on counting processes is given in Andersen *et al.* [12]. The theory of counting processes is an essential tool in the machinery that allows for calculating conditional expectations of the form (6). In particular, since the intensities λ_{ij} of the processes N_{ij} are given by

$$\lambda_{ij}(t) = I_i(t^-)\mu_{ij}(t),$$

the dynamics of the payment process A, defined by (7), can be reformulated as

$$\begin{aligned} dA_t =& \sum_{i\in\mathcal{S}}\Big(I_i(t)\alpha_i(t) + \sum_{j\neq i} I_i(t^-)\alpha_{ij}(t)\mu_{ij}(t)\Big)dt \\ &+ \sum_{i\in\mathcal{S}}\sum_{j\neq i}\alpha_{ij}(t)\,dM_{ij}(t), \end{aligned} \tag{8}$$

given some continuity conditions on μ_{ij}, α_i and α_{ij}. Here, M_{ij} denotes the compensated process corresponding to N_{ij}. Moreover, by the relationship

$$X_t = \sum_{i\in\mathcal{S}} iI_i(t), \quad I_i(t) = I_i(0) + \sum_{j\neq i}\left(N_{ji}(t) - N_{ij}(t)\right),$$

the state process, the indicator processes, and the counting processes carry the same information, which is represented by the natural filtration $\mathcal{F}$ of the process X. Using the martingale property of M_{ij} and the Markov property of X, the prospective reserve becomes

$$Y_t = E\Big[\int_t^T e^{-\int_t^s \delta_u du}\,\beta(s, X_s)ds|X_t\Big], \tag{9}$$

where

$$\beta(t, X_t) = \sum_{i\in\mathcal{S}} I_i(t)\Big(\alpha_i(t) + \sum_{j\neq i}\alpha_{ij}(t)\mu_{ij}(t)\Big).$$

Obtaining a conditional expectation of the form (9) allows us to evaluate the reserve by invoking the Feynman-Kac representation formula, which is a generalization of Thiele's differential equation.

4 Thiele's differential equation and the Feynman-Kac formula

Under continuity and boundedness assumptions on δ, β and $(\mu_{ij})_{ij}$, the function $v : \mathbb{R}^+ \times \mathcal{S} \mapsto \mathbb{R}$ defined by

$$v(t,i) = E\Big[\int_t^T e^{-\int_t^s \delta_u du}\,\beta(s,X_s)ds|X_t = i\Big]$$

satisfies the celebrated Thiele differential equation

$$\begin{cases} \frac{\partial v(t,i)}{\partial t} - \delta(t)v(t,i) + \beta(t,i) + \sum_{j\neq i}\mu_{ij}(t)(v(t,j)-v(t,i)) = 0, \\ v(T,i) = 0, \quad i \in \mathcal{S}. \end{cases} \tag{10}$$

The Thiele equation provides a convenient way of determining the prospective reserves for a wide range of insurance contracts, see e.g. Møller and Steffensen [20] and Norberg [21, 22, 23], although in many cases it must be solved numerically. For payments of the form (8), it is however possible to solve (10) analytically. Indeed, the reserve is given by

$$v(t,i) = \sum_{j\in\mathcal{S}}\int_t^T e^{-\int_t^s \delta_u du}\,\beta(s,j)p_{ij}(s,t)ds,$$

where the probabilities $p_{ij}(s,t) = \mathbb{P}(X_t = j|X_s = i)$ satisfy the Kolmogorov backward equation

$$\begin{cases} \frac{\partial p_{ij}}{\partial s}(s,t) + \sum_k \mu_{ik}(s)(p_{kj}(s,t) - p_{ij}(s,t)) = 0, \quad s < t, \\ p_{ij}(t,t) = \delta_{ij}, \end{cases}$$

where δ_{ij} is the usual Kronecker symbol ($\delta_{ij} = 0$ if $i \neq j$; $\delta_{ij} = 1$ if $i = j$).

The Thiele equation (10) has similarities with the famous Feynman-Kac representation formula for general Markov processes: If X is a Markov process with infinitesimal generator $\mathcal{A}$ on a more general space $\mathcal{S}$, then the function $v : \mathbb{R}^+ \times \mathcal{S} \mapsto \mathbb{R}$ defined by

$$v(t,x) = E\Big[e^{-\int_t^T \delta_u du}f(X_T) + \int_t^T e^{-\int_t^s \delta_u du}\,\beta(s,X_s)ds|X_t = x\Big]$$

satisfies the Feynman-Kac PDE

$$\begin{cases} \frac{\partial v(t,x)}{\partial t} - \delta(t)v(t,x) + \beta(t,x) + \mathcal{A}v(t,x) = 0, \quad t \leq T, \\ v(T,x) = f(x), \end{cases}$$

given some continuity and boundedness assumptions on δ, f and β, see e.g. Friedman [11].

For many insurance products, the contractual payments depend in some way on the prospective reserve itself. A typical example is a pension where the benefits less premiums include a cost of capital fee that is proportional to the reserve. Other examples include life insurance policies with withdrawal options, where the policyholder pays a lump sum fee proportional to the reserve in the case of withdrawal. The case where the payments are proportional to the reserve has received some attention. Norberg [21] lists several examples, including a widow's pension policy in the presence of administration expenses that depend partly of the reserve, and derives explicit solutions to the corresponding Thiele equations.

Christiansen *et al.* [7] expand on the ideas of Norberg and consider reserve-dependent payments for more general contracts. The authors obtain explicit expressions for the prospective reserve for the special case when the payment functions are linear in the reserve. One key observation for making this possible is the following: If the function β in (9) is of the form

$$\beta(t,i) = \beta^1(t,i) + \beta^2(t)v(t,i), \tag{11}$$

for given functions β^1 and β^2, then the Thiele equation (10) for the prospective reserve can be rewritten

$$\begin{cases} \frac{\partial v(t,i)}{\partial t} - (\delta(t) - \beta^2(t))v(t,i) + \beta^1(t,i) + \sum_{j\neq i} \mu_{ij}(t)(v(t,j) - v(t,i)) = 0, \\ v(T,i) = 0, \quad i \in \mathcal{S}. \end{cases} \tag{12}$$

This suggests that adding proportional payments of the form (11) with proportionality factor $\beta^2(t)$ is equivalent to reducing the interest rate by $\beta^2(t)$. Christiansen *et al.* show equivalence of the prospective reserves using a uniqueness result for solutions to (10), and a version of Cantelli's Theorem, which states that the solutions to (10) and (12) are identical. Hence, the prospective reserve corresponding to payments of the form (11) is given by

$$v(t,i) = \sum_{j\in\mathcal{S}} \int_t^T e^{-\int_t^s (\delta(u)-\beta^2(u))du}\, \beta^1(s,j)p_{ij}(s,t)ds.$$

Payments given by (11) correspond to the case where the annuity payments are proportional to the reserve. A similar argument as the one outlined above extends the result to the case where the transition payments are also proportional to the reserve.

Until this point, we have only discussed valuation of insurance liabilities. For risk management purposes, calculation of the prospective reserve is typically insufficient. This is due to the general fact that the conditional expectation of a random variable does not characterize its full distribution. An extreme outcome of biometric or interest rates may cause the payed benefits to exceed an insurance company's reserves, which may well see the company facing ruin. The Solvency II directive is an attempt to set a standard for risk management for European insurance companies by enforcing guidelines for calculating capital requirements.

5 Insurer solvency and capital requirements

Among other things, the Solvency II directive regulates the amount of capital that insurance companies must hold to reduce the risk of insolvency. In the directive's standard model, capital charges are calculated using a scenario based approach, and the Solvency Capital Requirement (SCR) is given as the difference between the present value under best estimate assumptions, which corresponds to the expected value, and the present value in a certain stress scenario. The SCR is computed on a risk-by-risk level, and then aggregated using predetermined correlation matrices. For disability insurance, perhaps the most important risk is *recovery risk*, the risk that the policy holder receives the payments for longer than anticipated, i.e. that claim termination rates are lower than anticipated.

As an alternative to the standard stress scenario, insurers may adopt an internal model, which should be based on a Value-at-Risk approach, such that, on a one-year horizon, with probability 0.995, the value of the insurers liabilities will not exceed the value of their assets. As an example, consider a portfolio of n insurance policies, and let (δ, μ) denote the collection of all interest and biometric rates necessary to value the cash flows from the portfolio. Further, let $\mathcal{F}$ denote the filtration generated by the insured individuals and (δ, μ). Further, let $B^k_{0,T},\ k = 1, \dots, n$, denote the random present value of the benefits less premiums for policy k payable up to the contractual end time T. That is, $B^k_{0,T}$ is essentially given by the sum of all future contractual payments, discounted back to time $t = 0$. The value of the portfolio liabilities in one year, henceforth denoted by $L(\delta, \mu)$, is given by

$$L(\delta, \mu) = E\Big[\sum_{k=1}^{n} B^k_{0,T} | \mathcal{F}_1\Big] = \sum_{k=1}^{n} \Big(B^k_{0,1} + e^{-\int_0^1 \delta_t dt} E[B^k_{1,T} | \mathcal{F}_1]\Big),$$

where $B^k_{0,1}$ represents benefits payed during the first year, and $E[B^k_{1,T}|\mathcal{F}_1]$ represents the value of the remaining liabilities at $t = 1$. The SCR is given by the difference between the best estimate and the 99.5%-quantile of the value of the liabilities in one year, including benefits payed during the first year, i.e.

$$\mathrm{SCR_p} = F_L^{-1}(p) - E[L(\delta, \mu)],$$

where $p = 0.995$. Determining the SCR requires knowledge of the joint distribution of the insured individuals and all interest and biometric rates. Typically, this is a high-dimensional problem that requires simulation techniques. Levantesi and Menzietti [19] estimate SCR under stochastic disability and mortality using simulation methods. The approach covers both systematic and idiosyncratic risk, and is suitable for small portfolios.

In order to reduce the computational complexity of the problem, scenario analysis and approximation methods may be used. Christiansen *et al.* [3] suggest an internal model for Solvency II based on linearization and Gaussian approximations. Christiansen and Steffensen [5] and Christiansen *et al.* [4] develop a safe-side scenario approach for the estimation of SCR using dynamic programming techniques. The main idea of the scenario approach is to find deterministic interest and transition rates $(\tilde{\delta}, \tilde{\mu})$ such that

$$\mathbb{P}\Big(L(\tilde{\delta}, \tilde{\mu}) \geq L(\delta, \mu)\Big) \geq p.$$

The collection $(\tilde{\delta}, \tilde{\mu})$ thus represents an extreme scenario with respect to the future liability value L. To formalize this concept, consider a *confidence set* M consisting of outcomes of (δ, μ) such that

$$\mathbb{P}\Big((\delta, \mu) \in M\Big) \geq p.$$

Christiansen and Steffensen [5] show that by choosing the safe-side scenario $(\tilde{\delta}, \tilde{\mu})$ as

$$(\tilde{\delta}, \tilde{\mu}) = \underset{(\bar{\delta}, \bar{\mu}) \in M}{\arg\max}\, L(\bar{\delta}, \bar{\mu}),$$

the maximum value of L over M is a conservative estimate of the p-quantile of L, i.e.

$$F_L^{-1}(p) \leq L(\tilde{\delta}, \tilde{\mu}) = \sup_{(\bar{\delta}, \bar{\mu}) \in M} L(\bar{\delta}, \bar{\mu}).$$

Hence, an upper bound for the SCR can indeed be obtained using this scenario approach, provided that the confidence set M can be chosen in a suitable way. Unfortunately, determining M is very difficult in most practical applications.

6 Summary of the appended papers

This section contains a brief summary of each of the five appended papers.

Paper A: Stochastic modelling of disability insurance in a multi-period framework

Scandinavian Actuarial Journal, Vol. 2015 No. 1 (2015), pp. 88–106.

In the first paper, we present a framework for modelling disability inception and claim termination in discrete time, extending the mortality model of Aro and Pennanen [1]. Let $E_{x,d,t}$ be the number of individuals with disability inception ages in $[x, x+1)$ and disability duration d at some point in the time period $[t, t+1)$. Further, let $R_{x,d,t}$ denote the number of individuals among $E_{x,d,t}$ with claim termination during $[t, t+1)$ and $[d, d+\Delta d)$. We assume that the conditional distribution of $R_{x,d,t}$ given $E_{x,d,t}$ is binomial:

$$R_{x,d,t} \sim \mathrm{Bin}(E_{x,d,t}, p_{x,d,t}), \tag{13}$$

where $p_{x,d,t}$ denotes the probability that the disability of an individual, with disability inception age in $[x, x+1)$ and disability duration d at some point in the time period $[t, t+1)$, is terminated before duration $d + \Delta d$. We propose the following logistic regression model:

$$\operatorname{logit} p_{x,d,t} = \sum_{i=1}^{n} \phi^i(x) \sum_{j=1}^{k} \nu_t^{ij} \psi^j(d), \tag{14}$$

where ϕ^i and ψ^j are age and duration dependent basis functions, respectively, and ν_t^{ij} are model parameters for year t. Using (13) and (14), the log-likelihood can be written

$$\begin{aligned} l(\nu_t; R_{\cdot,\cdot,t}) = \sum_{\substack{x \in X \\ d \in D}} \Big[& R_{x,d,t} \sum_{i=1}^{n} \phi^i(x) \sum_{j=1}^{k} \nu_t^{ij} \psi^j(d) \\ & - E_{x,d,t} \log\big(1 + \exp\big\{\sum_{i=1}^{n} \phi^i(x) \sum_{j=1}^{k} \nu_t^{ij} \psi^j(d)\big\}\big)\Big]. \end{aligned} \tag{15}$$

The maximum likelihood function is shown to be strictly concave under the assumption that $(\phi^i)_i$ are linearly independent and $(\psi^j)_j$ are linearly independent.

Hence, maximizing $l(\nu_t)$ over $\nu_t \in \mathbb{R}^{n\times k}$ using numerical methods gives a unique estimate for the vector ν_t of risk factors for each time period $[t, t+1)$. A model for disability inception can be obtained in a similar fashion by considering binomially distributed disability inception counts and using slight adaptations of (13)-(15). We fit several models for disability inception and termination probabilities to Swedish disability claims data over the time period 2000-2011.

Paper B: A hidden Markov approach to disability insurance

The second paper discusses prediction of future disability inception and termination rates. It is customary to combine generalized linear models (GLMs) of the form (14) with time series forecasting techniques into the following two-step method: In the first step, the parameters of the GLM are estimated from historical data. In the second step, a time trend component ν_t is assumed to follow a time series model, where a popular choice is the random walk with drift, and the parameters of the model are fitted to the estimated values of $(\nu_t)_t$. Prediction or simulation of future transition rates are obtained by prediction or simulation from the time series model for ν_t. This two-step approach provides an easy way of fitting the model to data and simulating future outcomes, and it has been employed by e.g. Brouhns *et al.* [2], Christiansen *et al.* [3], Djehiche and Löfdahl [10] and others.

An issue with the two-step approach is that at first, $(\nu_t)_t$ are considered parameters to be estimated. After estimating them, the assumption is altered so that ν is treated as a stochastic process. This may lead to both conceptual and numerical problems. In particular, the volatility of ν tends to be overestimated, which may have significant impact on pricing and risk management of insurance products. Qualitatively, this stems from the fact that yearly variations in the parameter values are caused by variations in the underlying process (systematic variation) as well as variations in the underlying population (idiosyncratic variation). The two-step approach makes no distinction between idiosyncratic and systematic variations.

We propose instead to treat ν as a hidden Markov process with transition densities parameterized by θ, say, and estimate θ using the Expectation-Maximization algorithm as follows: Assume that we observe $R_{x,d,1:n} := (R_{x,d,1}, \dots, R_{x,d,n})$, for x and d from given sets X and D of ages and disability durations, respectively. Given a parameter estimate θ^k, integrate the complete data log-likelihood $l(\theta; R_{\cdot,\cdot,1:n}, \nu_{1:n})$ with respect to the distribution of $\nu_{1:n} := (\nu_1, \dots, \nu_n)$ conditional on the observations $R_{\cdot,\cdot,1:n}$, e.g. let

$$Q(\theta|\theta^k) = E^{\theta^k}[l(\theta; R_{\cdot,\cdot,1:n}, \nu_{1:n})|R_{\cdot,\cdot,1:n}],$$

where

$$l(\theta; R_{\cdot,\cdot,1:n}, \nu_{1:n}) = \sum_{t=1}^{n} \Big[l(\nu_t; R_{\cdot,\cdot,t}) + \log f_{\nu_t|\nu_{t-1}}(\theta) \Big],$$

where f denotes the density of ν_t given ν_{t-1} and $l(\nu_t; R_{\cdot,\cdot,t})$ is the one-year log-likelihood given by (15). In many applications, it is difficult to perform the E-step analytically, and we are forced to use particle methods to obtain an estimate of Q.

In the maximization step, we maximize Q w.r.t. θ to obtain

$$\theta^{k+1} = \arg\max_{\theta} Q(\theta|\theta^k).$$

Explicit expressions for θ^{k+1} are obtained when ν is a multivariate Brownian motion with drift. This is possible due to the fact that the first term of the complete data likelihood does not depend on θ, which means that it can be completely disregarded when maximizing Q. We fit several models for disability inception and termination probabilities to Swedish disability claims data over the time period 2000-2011, and compare the resulting estimates of θ with those obtained from the two-step approach. We find that while the drift estimates are similar, the volatility estimates are significantly reduced compared to the two-step approach, as expected.

Paper C: Risk aggregation and stochastic claims reserving in disability insurance

Insurance: Mathematics and Economics, Vol. 59 (2014), pp. 100–108.

In the third paper, we consider a large, homogeneous portfolio of life or disability annuity policies. The policies are assumed to be independent conditional on an external stochastic process representing the economic environment. More precisely, let $N_t^k,\ \ k \geq 1$, be counting processes starting from zero with intensities given by

$$\lambda_t^k = q(t, Z_t)(1 - N_t^k), \quad t \geq 0. \tag{16}$$

Here, N_t^k represents the state (disabled or recovered) of an insured individual at time t, and Z_t represents the state of the economic-demographic environment at time t. Further, assume that $N_t^k,\ \ k \geq 1$, are independent conditional on $\mathcal{F}_t^Z$, the natural filtration of Z.

The rationale behind the model (16) is that every individual is affected by her environment. To take an example from disability insurance, interpret λ_t^k as the claims termination intensity, and let Z_t denote the state of the social security system. The Swedish government launched major reforms of the national sickness insurance system in 2008, changing the rules for obtaining benefits from the Social Insurance Agency. This reform has been of major importance to the reduction in sickness absence, which implies an increase in termination intensities. The relation between the environment process and termination intensity is expressed through the function q in (16).

Consider annuity contracts paying $g(t, Z_t)$ monetary unit continuously as long as $N_t^k = 0$, until a fixed future time T. The random present value L_t of the portfolio consisting of these contracts can be written as

$$L_t = \sum_{k=1}^{n} \int_t^T (1 - N_s^k) e^{-\int_t^s r(u)du} g(s, Z_s) ds,$$

where r is the interest rate process, here assumed adapted to $\mathcal{F}^Z$. Using a conditional version of the Law of Large Numbers due to Prakasa Rao [24], we show that in order to determine the ultimate reserve risk of the portfolio, it suffices to consider the random variable V_t defined by

$$V_t = \int_t^T e^{-\int_t^s q(u,Z_u)du} e^{-\int_t^s r(u)du} g(s, Z_s) ds.$$

This result establishes the connection between claims reserving and risk aggregation for large portfolios. Further, we derive a partial differential equation for moments of V_t. For $n \geq 1$, let $v_n(t, z) = E^{t,z}[V_t^n]$. We show that if Z is a Markov process with infinitesimal generator $\mathcal{A}$, $v_n(t, z)$ satisfies the Feynman-Kac type PDE

$$\begin{cases} -\frac{\partial v_n}{\partial s} + n(q(s, z) + r(s))v_n = \mathcal{A}v_n + ng(s, z)v_{n-1}, & t < s < T \\ v_n(T, z) = 0, \end{cases} \tag{17}$$

where, naturally, $v_0(t, z) = E^{t,z}[V_t^0] = 1$.

Next, we consider a continuous time approximation of a statistical generalized linear model for the transition probability of the form (14), where the transition rate q is defined by

$$q(t, \nu_t) = c \log(1 + \exp(a(t)^T \nu_t)),$$

where $a : \mathbb{R}^+ \mapsto \mathbb{R}^n$ is a time-dependent vector of covariates and $c \in \mathbb{R}$ represents a modelling choice. In principle we could directly solve the PDE (17) using the infinitesimal generator of ν to calculate moments of V_t. However, if n is large, this procedure will have a high computational cost. It is tempting to try to reduce the dimensionality of the problem by defining the stochastic process Z by

$$Z_t = a(t)^T \nu_t,$$

and solve the PDE using the generator of Z. Unfortunately, even in the simple case when ν is an n-dimensional Brownian motion, the one-dimensional process Z is not Markovian. In pursuit of this goal, we make several attempts to construct a process $\widehat{Z}$ sharing some key characteristics with Z.

As it turns out, the simple Markovian projection introduced by Krylov [15] and extended by Gyöngy [13], Kurtz and Stockbrigde [16, 17] and several other authors, numerically yields a very good approximation of moments of present values of disability annuities.

Paper D: Aggregation of one-year risks in life and disability insurance

In the fourth paper, we consider an extension of the conditional independence model proposed by Djehiche and Löfdahl [10], where the individuals in a large life insurance portfolio are assumed indendent conditional on a stochastic process representing the economic-demographic environment. Let $N_t^k,\ k \geq 1$, denote the states of insured individuals, $\tau^k,\ k \geq 1$, be random event times (e.g. times of death or recovery from disability), q_x denote the transition rate and let Z_t represent the state of the environment. Consider a policy with payment stream dA_t^k given by

$$dA_t^k = g_x(t, Z_t)(1 - N_t^k)dt + h_x(t, Z_{t-})dN_t^k, \quad t \leq T_x,$$

where x is a parameter representing e.g. the age of the insured. The policy pays $g_x(t, Z_t)$ continuously, as long as $N_t^k = 0$, until a fixed future time T_x. In addition, if the event time τ^k is reached before T_x, the policy immediately pays a lump sum of $h_x(\tau^k, Z_{\tau^k})$. This type of policy allows for payments from the contract to depend on time as well as on the state of the economic-demographic environment and the age of the insured. For example, the contract could be inflation-linked and contain a deferred period.

For a portfolio consisting of n policies, let $L_{t+1}^{(n)}$ denote the value of the portfolio liabilities in one year from now. Further, let $\bar{q}_x = q_x + r$ denote the sum of transition and interest rates, and assume that Z is a Markov process with infinitesimal generator $\mathcal{A}$. Using the Conditional Law of Large Numbers (CLLN), we show that if n becomes large enough, the future liability value can be approximated by

$$L_{t+1}^{(n)} \approx \sum_{x,k} (1 - N_t^k) V^x,$$

where

$$V^x = \int_t^{t+1} \tilde{g}_x(s, Z_s) e^{-\int_t^s \bar{q}_x(u,Z_u)du} ds \\ + e^{-\int_t^{t+1} \bar{q}(u,Z_u)du} v_x(t+1, Z_{t+1}),$$

and v_x is a function satisfying the Feynman-Kac PDE

$$\begin{cases} -\frac{\partial v_x}{\partial s} + \bar{q}_x(s,z)v_x = \mathcal{A}v_x + \tilde{g}_x(s,z), & t+1 \le s < T_x, \\ v_x(T_x, z) = 0, \end{cases}$$

where $\tilde{g}$ is defined by

$$\tilde{g}_x(s,z) = g_x(s,z) + h_x(s,z)q_x(s,z).$$

This result indicates that the idiosyncratic risks are diversified away, and that only the systematic risk, i.e. the risk that the environment changes, remains.

Based on the CLLN, we derive a semi-analytical approximation of the systematic risk quantiles of the future liability value for a homogeneous portfolio under a one-factor diffusion model. Further, we propose two different risk aggregation techniques for a portfolio consisting of large, homogeneous pools under multi-factor models based on the variance-covariance method and convex optimization techniques, respectively. Finally, we present numerical results based on disability claims data from the Swedish insurance company Folksam, and compare the resulting capital charges with the Solvency II standard method.

Paper E: Nonlinear reserving in life insurance: aggregation and mean-field approximation

In the fifth paper, we consider claims reserving for life insurance policies with reserve-dependent payments. A typical example is a pension where the benefits less premiums include a cost of capital fee that is proportional to the reserve. Other examples include life insurance policies with withdrawal options, where the policyholder pays a lump sum fee proportional to the reserve in the case of withdrawal.

Let X_t denote the state of the policy holder at time t, and assume that X is a continuous time Markov chain with state space $\mathcal{S} = \{1, 2, \ldots, m\}$, transition matrix $G(t) = (\mu_{ij}(t))_{ij}$, and natural filtration $\mathbb{F} = (\mathcal{F}_t)_{t\ge 0}$. To X we associate the indicator processes $I_i(t) = I\{X_t = i\}$, whose value is 1 if the policy is at state i at time t and 0 otherwise, and the counting processes $N_{ij}(t),\ i \neq j$, such that

$$N_{ij}(t) = \#\{\tau \in (0,t] : X_{\tau^-} = i, X_\tau = j\},$$

which count the number of jumps from state i to state j during $(0,t]$.

Further, let A_t denote the payment process of accumulated policy benefits less premiums during $(0,t]$, and let δ_t denote the discount rate, here assumed progressively measurable and bounded. Following Norberg [21, 22], we recall the conditional expectation formulation of the prospective reserve associated with A, X, G and δ as

$$Y_t = E\Big[\int_t^T e^{-\int_t^s \delta_u du}\, dA_s|\mathcal{F}_t\Big].$$

If the incremental payments dA_t are bounded functions depending on the sojourns and transitions of X_t, then the prospective reserve can be written

$$Y_t = E\Big[\int_t^T (\beta(s,X_s) - \delta_s Y_s)\, ds|\mathcal{F}_t\Big], \quad t \in [0,T],$$

for some function β. This representation suggests to consider more general nonlinear reserving schemes through making the payments streams general adapted processes which are reserve-dependent. Given some integrability and measurability conditions, Y_t defined by

$$Y_t = E\Big[\int_t^T g(s,\omega,Y_s)\, ds|\mathcal{F}_t\Big], \quad t \in [0,T],$$

satisfies the BSDE

$$-dY_t = g(t,\omega,Y_t)dt - Z_t dM_t, \quad Y_T = 0,$$

where M is the accompanying martingale of X and Z is an $\mathbb{F}$-adapted process. We show that the prospective reserves of some life insurance contracts fit within the following more general class of Markovian BSDEs defined by

$$-dY_t = g(t,X_t,Y_t,Z_t)dt - Z_t dM(t), \quad Y_T = \phi(X_T). \tag{18}$$

It is easily seen that if (Y,Z) solves (18), then it admits the representation

$$Y_t = E\Big[\phi(X_T) + \int_t^T g(s,X_s,Y_s,Z_s)\, ds|X_t\Big], \quad t \in [0,T].$$

Further, we have $Y_t = V(t,X_t)$, where the function $V : [0,T] \times \mathcal{S} \to \mathbb{R}$ satisfies the following Feynman-Kac formula, which we call the nonlinear Thiele's equation:

$$\begin{cases} \frac{\partial V}{\partial t}(t,i) + g(t,i,V(t,i),(V(t,j)-V(t,k))_{jk}) \\ +\sum_{j,\,j\neq i} \mu_{ij}(t)\,(V(t,j)-V(t,i)) = 0, \\ V(T,i) = \phi(i), \qquad i \in \mathcal{S}. \end{cases}$$

Another important consideration for insurance companies is that of diversification and aggregation of insurance portfolios. Consider, for example, policies where, at the death of one policyholder, (a proportion of) the accumulated reserve of that policy is distributed back to the collective. The payments will then depend on the performance of the reserve of each policy relative to that of the 'average' policy, or the mean performance of the reserves of all policies, in the case of a very large and homogeneous portfolio. This example motivates for the concept of aggregation and mean-field approximation for life insurance contracts and for the need to formulate a prospective reserve (of mean-field type) that reflects this averaging. Further examples including reserving for gender-neutral life insurance policies are considered. We formulate a prospective reserve of mean-field type and show how to obtain it using aggregation and mean-field approximation of the prospective reserves of large and homogeneous insurance portfolios.

7 References

[1] ARO, H., AND PENNANEN, T. A user-friendly approach to stochastic mortality modelling. *European Actuarial Journal 1*, 2 (2011), 151–167.

[2] BROUHNS, N., DENUIT, M., AND VERMUNT, J. K. A poisson log-bilinear regression approach to the construction of projected lifetables. *Insurance: Mathematics and Economics 31*, 3 (2002), 373–393.

[3] CHRISTIANSEN, M., DENUIT, M., AND LAZAR, D. The Solvency II square-root formula for systematic biometric risk. *Insurance: Mathematics and Economics 50* (2012), 257–265.

[4] CHRISTIANSEN, M., HENRIKSEN, L., SCHOMACKER, K., AND STEFFENSEN, M. Stress scenario generation for solvency and risk management. *Scandinavian Actuarial Journal*, ahead-of-print (2014), 1–28.

[5] CHRISTIANSEN, M., AND STEFFENSEN, M. Safe-side scenarios for financial and biometrical risk. *Astin Bulletin 43*, 03 (2013), 323–357.

[6] CHRISTIANSEN, M. C. Multistate models in health insurance. *AStA Advances in Statistical Analysis 96*, 2 (2012), 155–186.

[7] CHRISTIANSEN, M. C., DENUIT, M. M., AND DHAENE, J. Reserve-dependent benefits and costs in life and health insurance contracts. *Insurance: Mathematics and Economics 57* (2014), 132–137.

[8] CZADO, C., DELWARDE, A., AND DENUIT, M. Bayesian Poisson log-bilinear mortality projections. *Insurance: Mathematics and Economics 36*, 3 (2005), 260–284.

[9] DEMPSTER, A., LAIRD, N., AND RUBIN, D. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)* (1977), 1–38.

[10] DJEHICHE, B., AND LÖFDAHL, B. Risk aggregation and stochastic claims reserving in disability insurance. *Insurance: Mathematics and Economics 59* (2014), 100–108.

[11] FRIEDMAN, A. *Stochastic differential equations and applications*. Courier Dover Publications, 2012.

[12] GILL, R., KEIDING, N., AND ANDERSEN, P. *Statistical models based on counting processes*. Springer Science & Business Media, 1997.

[13] GYÖNGY, I. Mimicking the one-dimensional marginal distributions of processes having an Itô differential. *Probability Theory and Related Fields 71*, 4 (1986), 501–516.

[14] HYNDMAN, R., AND ULLAH, M. S. Robust forecasting of mortality and fertility rates: A functional data approach. *Computational Statistics and Data Analysis 51* (2007), 4942–4956.

[15] KRYLOV, N. Once more about the connection between elliptic operators and Itô's stochastic equations. In *Statistics and control of stochastic processes, Steklov Seminar* (1984), pp. 214–229.

[16] KURTZ, T. G., AND STOCKBRIDGE, R. H. Existence of Markov controls and characterization of optimal Markov controls. *SIAM Journal on Control and Optimization 36* (1998), 609–653.

[17] KURTZ, T. G., AND STOCKBRIDGE, R. H. Stationary solutions and forward equations for controlled and singular martingale problems. *Electronic Journal of Probability 6*, 17 (2001), 1–52.

[18] LEE, R. D., AND CARTER, L. R. Modeling and forecasting US mortality. *Journal of the American statistical association 87*, 419 (1992), 659–671.

[19] LEVANTESI, S., AND MENZIETTI, M. Managing longevity and disability risks in life annuities with long term care. *Insurance: Mathematics and Economics 50* (2012), 391–401.

[20] MØLLER, T., AND STEFFENSEN, M. *Market-valuation methods in life and pension insurance*. Cambridge University Press, 2007.

[21] NORBERG, R. Reserves in life and pension insurance. *Scandinavian Actuarial Journal 1991*, 1 (1991), 3–24.

[22] NORBERG, R. Hattendorff's theorem and Thiele's differential equation generalized. *Scandinavian Actuarial Journal 1992*, 1 (1992), 2–14.

[23] NORBERG, R. Basic life insurance mathematics. *Lecture notes, Laboratory of Actuarial Mathematics, University of Copenhagen* (2002).

[24] PRAKASA RAO, B. L. S. Conditional independence, conditional mixing and conditional association. *Annals of the Institute of Statistical Mathematics 61*, 2 (2009), 441–460.

[25] RENSHAW, A., AND HABERMAN, S. Modelling recent time trends in uk permanent health insurance recovery, mortality and claim inception transition intensities. *Insurance: Mathematics and Economics 27* (2000), 365–396.

[26] WU, C. J. On the convergence properties of the EM algorithm. *The Annals of statistics* (1983), 95–103.