Stochastic modelling in disability insurance

Björn Löfdahl

Sammanfattning

Den här avhandlingen består av två artiklar som berör stokastisk modellering av sjukförsäkring. 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. Insjuknandemodellen inkluderar även prediktion av IBNR, det vill säga skador som uppkommit men ännu ej rapporterats. Slutligen anpassar vi modellparametrarna till svenskt sjukförsäkringsdata.

I den andra artikeln betraktar vi en stor, homogen portfölj av sjukförsäkringskontrakt. Vi antar oberoende mellan populationens individer betingat på en stokastisk process som representerar omvärldens tillstånd. Genom att använda 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. Vi ger ett numeriskt exempel baserat på svenskt sjukförsäkringsdata.

Abstract

This thesis consists of two papers related to the stochastic modelling of disability insurance. In the first paper, we propose a stochastic semi-Markovian framework for disability modelling in a multi-period discrete-time setting. The logistic transforms of disability inception and recovery probabilities are modelled by means of stochastic risk factors and basis functions, using counting processes and generalized linear models. The model for disability inception also takes IBNR claims into consideration. We fit various versions of the models into Swedish disability claims data.

In the second 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. Using a conditional law of large numbers, we establish the connection between risk aggregation and claims reserving for large portfolios. Further, we derive a partial differential equation for moments of present values. Moreover, we show how statistical multi-factor intensity models can be approximated by one-factor models, which allows for solving the PDEs very efficiently. Finally, we give a numerical example where moments of present values of disability annuities are computed using finite difference methods.

Acknowledgements

I am immensely grateful to my advisor Boualem Djehiche for his constant support and encouragement. Without him, this thesis would not have been written. Furthermore, I wish to thank my colleagues in the Division of Mathematical Statistics, and especially my fellow PhD candidates for the support they have given me and for the joys we have shared. The financial support from the Filip Lundberg and Eir's 50 Years foundations is gratefully acknowledged. Finally, I wish to thank Gunnar Andersson at Folksam and Kristina Hagström at Skandia Liv for providing me with data and for valuable discussions regarding this project.

Introduction and summary of the papers

A long-term 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 prolongued disability may be unbearable for a single individual, it is feasible to distribute the cost over a large group of individuals. The risk is transfered 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 company. The company sets aside a *reserve* corresponding to the expected cost of disability of the collective. If the collective is large enough, the sum of payed benefits should be close to the sum of all premiums due to diversification.

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.

The insurer faces *systematic risk*, i.e. the risk that the realized future rates do not equal the predicted rates. In particular, the benefits payed by the company may exceed the received premiums, ultimately leading to the company facing *ruin*, i.e. not being able to fulfill its obligations to the policy holders. Hence, estimating the systematic risk and taking measures to mitigate it is a matter of the utmost importance. This is one of the key questions in the upcoming Solvency II regulatory framework.

Solvency II

The upcoming Solvency II regulatory framework brings many new challenges to the insurance industry. In particular, the new regulations suggest a new mindset regarding the valuation and risk management of insurance products. 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 taken into account. This may have an impact on pricing as well as on capital charges.

In the Solvency II standard model, capital charges are computed using a scenario based approach, and the capital charge is given as the difference between the present value under best estimate assumptions, and the present value in a certain shock scenario. As an alternative, insurers may adopt an internal model, which should be based on a Value-at-Risk approach.

A nice introduction to insurer solvency is given in Hult *et. al.* [8]. Let A and L denote the values of the assets and liabilities of an insurance company one year from now. We consider the company to be *solvent* if

$$VaR_{0.005}(A - L) \leq 0, \tag{1}$$

that is, on a one-year horizon, with probability 0.995, the value of the liabilities will not exceed the value of the assets. Assume that the assets of the company can be divided into two parts, a buffer capital K and a hedge portfolio H which should replicate the liabilities L as closely as possible. The problem is then to determine the size of the needed buffer capital K such that

$$K \geq VaR_{0.005}(H - L). \tag{2}$$

The buffer capital depends on the distributions of the future values of the assets and liabilities.

A simple framework for modelling disability insurance liabilities is presented in the next section.

General aspects of a disability insurance policy

Consider an insurance policy paying a continuous annuity of dt units in the interval $[t,t+dt)$, if the insured is ill. Let x be the age of the insured, and $S(x+t)$ the random state occupied by the insured at age $x+t$, $t \geq 0$. We consider the simple model where S can visit the states a ('active'), i ('ill') and d ('dead').

Assume $S(x) = k$, $k \in \{a, i, d\}$. Then, the present value L_k of the benefits paid by the insurer is given by

$$L_k = \int_0^\infty I_{\{S(x+t)=i|S(x)=k\}} \nu^t dt \tag{3}$$

where ν^t is the discount factor. The single premium payed by an active individual is given by

$$\mathrm{E}[L_a] = \int_0^\infty P(S(x+t) = i|S(x) = a)\nu^t dt. \tag{4}$$

If the insured falls ill at age x, the insurance company sets aside the reserve

$$\mathrm{E}[L_i] = \int_0^\infty P(S(x+t) = i|S(x) = i)\nu^t dt. \tag{5}$$

To compute the premium and reserve, a probabilistic model of the process S is required. Popular choices include Markov and semi-Markov models as depicted by e.g. Haberman and Pitacco [7], and counting process models as described in Andersen *et al.* [2]. Below, we give a short account on counting processes and a classical survival curve estimator known as the Nelson-Aalen estimator.

Counting processes

A *counting process* $N = \{N_t\}_{t\geq 0}$ is a càdlàg process starting from zero, with paths which are piecewise constant and non-decreasing, having jumps of size $+1$ only. The process N is a local submartingale with non-decreasing predictable *compensator* Λ such that the process $M = \{M_t\}_{t\geq 0}$ defined by

$$M_t = N_t - \Lambda_t \tag{6}$$

is a *martingale* with respect to $\mathcal{F}^N = \{\mathcal{F}_t^N\}_{t\geq 0}$, the natural filtration of N. That is, for $t \geq s$, we have

$$E[M_t|\mathcal{F}_s^N] = M_s. \tag{7}$$

We say that N has *intensity process* λ if λ is a predictable process satisfying

$$\Lambda_t = \int_0^t \lambda_s ds, \ \ t \geq 0. \tag{8}$$

Under further conditions, it can be shown that

$$\lim_{h\to 0} \frac{1}{h} P(N_{t+h} - N_t = 1|\mathcal{F}_t^N) = \lambda(t^+) \quad a.s. \tag{9}$$

A complete treatment of statistical models based on counting processes is given in Andersen *et al.* [2].

The multiplicative intensity model and the Nelson-Aalen estimator

Consider a multivariate counting process $N = \{N^1, ..., N^n\}$ starting from zero with natural filtration $\mathcal{F}^N$. Aalen [1] introduced the multiplicative intensity model for counting processes, where it is assumed that the intensity processes $\lambda^k, \ k = 1, \ldots, n$, are of the form

$$\lambda_t^k = \alpha_t(1 - N_{t^-}^k), \quad k = 1, ..., n, \ \ t \geq 0,$$

where α_t is a non-negative deterministic function. Since its introduction in 1978, the Nelson-Aalen estimator has been widely applied in the field of survival analysis, and by scientists and practitioners alike. It is frequently used by Swedish insurance companies to estimate *survival functions*, e.g. termination functions in disability insurance,

The quantity of interest is the *survival function* $S(t),\ t \geq 0$, defined by

$$S(t) = e^{-A_t}, \tag{10}$$

where the *cumulative hazard* A_t is given by

$$A_t = \int_0^t \alpha_s ds. \tag{11}$$

Define the processes $\bar{N}$, Y and J by

$$\begin{aligned} \bar{N}_t &= \sum_{k=1}^n N_t^k \\ Y_t &= \sum_{k=1}^n (1 - N_{t-}^k) \\ J_t &= I\{Y_t > 0\}, \end{aligned} \tag{12}$$

and let

$$A_t^* = \int_0^t J_s \alpha_s ds. \tag{13}$$

Heuristically, when $P(J_t = 0)$ is small, A_t^* is almost equal to A_t. Further, since

$$\int_0^t d\bar{N}_s - \int_0^t \alpha_s Y_s ds \tag{14}$$

is a martingale, so is

$$\int_0^t J_s Y_s^{-1} d\bar{N}_s - \int_0^t J_s \alpha_s ds, \tag{15}$$

since J and Y are predictable. It follows that

$$E[A_t^*] = E[\int_0^t J_s \alpha_s ds] = E[\int_0^t J_s Y_s^{-1} d\bar{N}_s], \tag{16}$$

so that $\hat{A}_t$, the Nelson-Aalen estimator defined by

$$\widehat{A}_t = \int_0^t J_s Y_s^{-1} d\bar{N}_s, \tag{17}$$

is an approximately unbiased non-parametric estimator for A_t. Moreover, a biased estimator of the survival function $S(t)$ is given by

$$\widehat{S}(t) = e^{-\widehat{A}_t} \tag{18}$$

Under additional assumptions, estimators of the form (18) can be used to compute premiums and reserves of the form (4)-(5).

Summary of Paper 1: Stochastic modelling of disability insurance in a multi-period framework

In the first paper, we present a stochastic semi-Markovian framework for modelling disability inception and recovery in discrete time. It also allows for the consideration of Incurred But Not Reported (IBNR) claims. Following the approach of Aro and Pennanen [3], we model the logit probabilities for disability inception and termination rates by means of stochastic risk factors and basis functions across cohorts.

To be more specific, let $E_{x,t}$ be the number of healthy individuals aged $[x, x+1)$ years at the beginning of time period t in a given disability insurance scheme. We denote by $D_{x,t}$ the number of individuals falling ill amongst the $E_{x,t}$ insured healthy individuals during time interval $[t, t+1)$. We assume that the conditional distribution of $D_{x,t}$ given $E_{x,t}$ is binomial:

$$D_{x,t} \sim \mathrm{Bin}(E_{x,t}, p_{x,t}), \tag{19}$$

where $p_{x,t}$ is the probability that an x-year-old individual randomly selected at t falls ill during $[t, t+1)$. We model the logistic disability inception probabilites by

$$\operatorname{logit} p_{x,t} = \sum_{i=1}^{n} \nu_t^i \phi^i(x), \tag{20}$$

where ϕ^i are user-defined basis functions and ν_t^i are stochastic risk factors. This structure renders the framework very flexible, allowing for the incorporation of user views into the model through the selection of numbers of basis functions as well as their form and properties. Moreover, suitable choices of basis functions assign the risk factors of the model tangible interpretations, which facilitates both the assessment of the model and the study of the connections between disability inception and economic factors, similarly to [9, 5].

The historical values of the risk factors $\nu_t = (\nu_t^1, \ldots, \nu_t^n)$ can be easily obtained by maximum likelihood estimation as follows. Given a set of basis functions $\{\phi^i\}$ and the historical values of $D_{x,t}$ and $E_{x,t}$, the log-likelihood function for yearly values of ν_t can be written using (19) and (20) as

$$l(\nu_t) = \sum_{x \in X} \Big[D_{x,t} \sum_{i=1}^{n} \nu_t^i \phi^i(x) - E_{x,t} \log\big(1 + \exp\big\{\sum_{i=1}^{n} \nu_t^i \phi^i(x)\big\}\big) \Big] + c_t, \tag{21}$$

where c_t is a constant. The maximum likelihood function is shown to be strictly concave under the assumption that the basis functions $\{\phi^i\}$ are linearly independent. This feature not only means that the risk factors are unique and well-defined, but also enables the use of convex optimization tools in model calibration. Maximizing $l(\nu_t)$ over $\nu_t \in \mathbb{R}^n$ using numerical methods gives an estimate for the vector ν_t of risk factors for each time period $[t, t+1)$.

A similar framework is proposed for modelling disability termination. 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 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:

$$R_{x,d,t} \sim \text{Bin}(E_{x,d,t}, p_{x,d,t}), \tag{22}$$

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:

$$\operatorname{logit} p_{x,d,t} = \sum_{i=1}^{n} \phi^i(x) \sum_{j=1}^{k} \nu_t^{ij} \psi^j(d), \tag{23}$$

where ϕ^i and ψ^j, are age and duration dependent basis functions, respectively, and ν_t^{ij} are stochastic risk factors. Using (22) and (23), the log-likelihood can be written

$$\begin{aligned} l(\nu_t) = \sum_{\substack{x \in X \\ d \in D}} \Big[& R_{x,d,t} \sum_{i=1}^{n} \phi^i(x) \sum_{j=1}^{k} \nu_t^{ij} \psi^j(d) \\ & - E_{x,d,t} \log \big(1 + \exp\big\{ \sum_{i=1}^{n} \phi^i(x) \sum_{j=1}^{k} \nu_t^{ij} \psi^j(d) \big\}\big) \Big] + c_t. \end{aligned} \tag{24}$$

The maximum likelihood function is shown to be strictly concave under the assumption that $\{\phi^i\}$ are linearly independent and $\{\psi^j\}$ are linearly independent. Again, maximizing $l(\nu_t)$ over $\nu_t \in \mathbb{R}^{n \times k}$ using numerical methods gives an estimate for the vector ν_t of risk factors for each time period $[t, t+1)$. Inserting the estimator $\hat{\nu}_t$ into (23), and denoting the estimator of $p_{x,d,t}$ by $p_{\hat{\nu}_t}(x, d)$, we obtain the termination curves

$$\lambda_{\hat{\nu}_t}(x, d) = \prod_{n=0}^{d/\Delta d - 1} (1 - p_{\hat{\nu}_t}(x, n\Delta d)). \tag{25}$$

We fit several models for disability inception and termination probabilities to Swedish disability claims data over the time period 2000-2011 by maximum likelihood.

Summary of Paper 2: Risk aggregation and stochastic claims reserving in disability insurance

In the second 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

$$\lambda_t^k = q(t, Z_t)(1 - N_t^k), \quad t \geq 0. \tag{26}$$

Here, N_t^k represents the state 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 (26) is that every individual is affected by her environment. To take an extreme example, interpret λ_t^k as the mortality intensity and let Z_t denote whether the country is at war or not. In times of war, the population mortality is usually higher than in times of peace. This effect is reflected through the function q in (26).

To take a recent example from disability insurance, interpret λ_t^k as the 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. Again, this effect is reflected through the function q in (26).

Now, consider annuity contracts paying 1 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

$$L_t = \sum_{k=1}^{n} L_t^k = \int_t^T \sum_{k=1}^{n} (1 - N_s^k) e^{-\int_t^s r(u)du} ds, \tag{27}$$

where r is the interest rate process, here assumed deterministic. Using a conditional version of the Law of Large Numbers due to Prakasa Rao [13], we show that, given the history of the policies, in order to determine the distribution of the present value of the portfolio, it suffices to consider the random variable V_t defined by

$$V_t = \int_t^T e^{-\int_t^s q(u, Z_u)du} e^{-\int_t^s r(u)du} ds. \tag{28}$$

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

$$\begin{cases} -\frac{\partial v_n}{\partial s} + n(q(s, z) + r(s))v_n = \mathcal{A}v_n + nv_{n-1}, & t < s < T \\ v_n(T, z) = 0, \end{cases} \tag{29}$$

where, naturally, $v_0(t,z) = E^{t,z}[V_t^0] = 1$. When Z is not a Markov process, this claim does not hold, but we might still wish to compute $v_n(t,z)$.

As an example, consider a statistical generalized linear model for the intensity λ_t of the form

$$\lambda_t = q(Z_t), \tag{30}$$

with

$$Z_t = \beta_t^T \psi(t), \tag{31}$$

where $\psi \in \mathbb{R}^n$ is a vector of, possibly time-dependent, covariates, and β_t is the parameter vector corresponding to time t. A popular approach to intensity modelling includes fitting a model of the form (31) over a range of time periods, and fitting a time series model to the time series of parameter estimates. The future development of the intensities is obtained by forecasting or simulation of the time series model. In principle, any dynamic for β_t is possible. The random walk is a natural choice, since it is easy to fit and simulate, and has been the model of choice in e.g. Christiansen et. al. [4].

However, even in the simple case when β is an n-dimensional Brownian motion, the one-dimensional process Z is not Markovian. If it was, and if we could determine its generator, we could solve the PDEs (29) very efficiently due to its low dimension. 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 [10] and extended by Gyöngy [6], Kurtz and Stockbrigde [11, 12] and several other authors, numerically yields a very good approximation of moments of present values of disability annuities.

References

[1] Aalen, O. Nonparametric inference for a family of counting processes. *The Annals of Statistics 6*, 4 (1978), 701–726.

[2] Andersen, P., Borgan, O., Gill, R., and Keiding, N. *Statistical Models Based on Counting Processes*. Springer, 1993.

[3] Aro, H., and Pennanen, T. User-friendly approach to stochastic mortality modelling. *European Actuarial Journal* (2011).

[4] 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.

[5] Donnelly, C., and Wüthrich, M. V. Bayesian prediction of disability insurance frequencies using economic indicators. *Annals of Actuarial Science 1*, 1, 1–20.

[6] Gyöngy, I. Mimicking the One-Dimensional Marginal Distributions of Processes Having an Itô Differential. *Probab. Theory Relat. Fields 71* (1986), 501–516.

[7] Haberman, S., and Pitacco, E. *Actuarial Models for Disability Insurance*. Chapman & Hall/CRC, 1999.

[8] Hult, H., Lindskog, F., Hammarlid, O., and Rehn, C. *Risk and Portfolio Analysis: Principles and Methods*. Springer, 2012.

[9] König, B., Weber, F., and Wüthrich, M. V. Prediction of disability frequencies in life insurance. *Zavarovalniski horizonti 7*, 3 (2011), 5–23.

[10] 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.

[11] Kurtz, T. G., and Stockbridge, R. H. Existence of markov controls and characterization of optimal markov controls. *SIAM J. Control Optim. 36* (1998), 609–653.

[12] 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.

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