INSURANCE MATHEMATICS

Bjørn Sundt
Vital Forsikring ASA, Norway

Keywords : insurance, life, non-life, reinsurance, premium, reserve, credibility, ruin, stop loss

Contents

Summary

This article gives a brief introduction to insurance mathematics.

Section 1 gives a general introduction without mathematical formulae. The concepts of insurance and insurance mathematics are explained and motivated and some areas of insurance mathematics indicated. The main difference between life and non-life insurance is pointed out. The relation to some other disciplines is indicated.

The topic of Section 2 is non-life insurance. Some concepts and properties of premium principles are indicated in subsection 2.1 whereas two frameworks of experience rating are discussed in subsections 2.2 (credibility) and 2.3 (bonus systems). An introduction to collective risk theory with emphasis on Lundberg's Inequality is given in subsection 2.4. Subsection 2.5 explains various sorts of reserves. The topic of subsection 2.6 is aggregate claims distributions with emphasis on recursive methods and the normal power approximation. The distinction between individual and collective models is explained. Subsection 2.7 introduces stop loss and excess of loss reinsurance, and the stop loss transform and stop loss ordering are discussed.

Section 3 is devoted to life insurance. In subsection 3.1 a general Markov chain framework is introduced and basic concepts like reserves, premiums and the equivalence principle explained. Thiele's differential equation is deduced. A retrospective expression for reserves is deduced within a less general model in subsection 3.2. As a special case of the general model, spouse pension is discussed in subsection 3.3. Finally, second order basis, safety loading and bonus are discussed in subsection 3.4.

1. Introduction

Insurance is protection against economic losses caused by uncertain events. When you buy an insurance policy, neither you nor the insurance company would know how much it will cost the insurance company to deliver the service that you have paid for. Thus, whereas in most other lines of business, uncertainty would be an unintended and undesired side-effect of the business activities, in insurance, uncertainty is the foundation of the business; insurance would be meaningless without uncertainty.

What would then be a reasonable price for an insurance policy? For most other products, a reasonable price would be the cost of producing and/or delivering the product, plus a reasonable profit. As stated above, in insurance these costs would not be known when the *premium* (price) is set, so that we would have to put a price on uncertain events. A natural framework for modeling such uncertain events is probability theory and mathematical statistics, and insurance mathematics is application of probability theory and mathematical statistics to modeling insurance risk.

We have already indicated that pricing is one area of insurance mathematics. Let us indicate some other areas:

- Reserving. Through its policies the insurance company has taken on the responsibility to cover future costs that are not yet known. How should one quantify the provisions for these liabilities in the balance of the company? In insurance such provisions are usually called *reserves*. However, in insurance mathematics, the term reserves is also often applied for the assets built up to cover these liabilities
- Choice of deductibles and limits.
- Choice of reinsurance programs.

By *reinsurance* we mean that an insurance company insures (reinsures, cedes) with another insurer a part of the risks that it has taken on through its insurance policies. The claim payments the ceding company receives from the reinsurer, are called *recoveries*. Through reinsurance an insurance company can increase its capacity. The reinsurance program of an insurance company is usually a complicated affair involving several reinsurance covers from several reinsurers. A reinsurer would also cede (retrocede) some of its business with other reinsurers.

In life and pension insurance, the policies usually also have an element of saving. For instance, in an old-age pension insurance the premium is typically paid until the insured person retires, and at retirement the pension starts running, that is, a reserve is built up during the premium period and after retirement this reserve is used to pay the pension. With such long-term contracts with a strong element of saving, the return on investments also becomes important. In insurance mathematics, one has traditionally considered the rate of return as non-random. However, the emphasis on stochastic modeling of the return is gradually increasing, enforcing the connection between insurance mathematics and financial mathematics. Traditionally, insurance mathematics has been concerned with risk related to liabilities of the company whereas financial mathematics has focused on asset risk. An interaction between these two disciplines seems natural:

- Both of them are concerned with modeling of risk.
- The insurance company will be affected by financial risk and insurance risk simultaneously.
- New products are developed where the interaction between insurance risk and finance risk is more pronounced, e.g. unit link products and catastrophe bonds.

Insurance mathematics is concerned with measuring risk, and within this framework premiums and reserves appear as measures of risk. Considering premiums in relation to market mechanisms would be a topic of insurance economics.

Obviously, many tools developed within a general framework of statistics and probability theory are also applicable within the framework of insurance mathematics. In the following, we shall look at some of the problems and tools that have been developed within insurance mathematics itself.

2. Non-life Insurance

2.2. Premium Principles

Let X denote an *insurance risk*, that is, the aggregate amount of claims to be covered by an insurance policy. Determining a premium for this policy is setting a price for covering this risk. We consider X as a random variable. The *pure premium* of X is the expected value of X. Normally one adds a risk loading to the pure premium and then gets the *net premium*. When adding administration costs to the net premium, we obtain the *gross premium*.

By a *premium principle* we mean a rule that to a risk X within a certain class assigns a non-random net premium $H(X)$ determined by the distribution of X.

Around 1975-85 an extensive theory on premium principles grew up. Several properties that might be desirable for a premium principle were set up and examined by checking which of these properties were satisfied by various premium principles. Such properties include:

- Non-negative risk loading; $H(X) \geq \mathrm{E}\,X$.
- When there is no uncertainty, the risk loading should be equal to zero.
- No rip-off; $\Pr\,(X \geq H(X)) > 0$.
- Additivity; $H(X + Y) = H(X) + H(Y)$ for independent risks X and Y.

The premium principles most common in practice are the expected value principle, the variance principle and the standard deviation principles, where the risk loading $H(X) - EX$ is proportional to, respectively, the expected value, the variance and the standard deviation of X. We observe that under the expected value principle, the premium is the same for all risks with the same mean, in particular risks without uncertainty. Intuitively this seems unreasonable as one would want the risk loading to reflect the uncertainty of the risk. This is achieved by the two other principles where the risk loading increases with the variance of the risk. In particular, we see that when there is no uncertainty, the risk loading is equal to zero under these principles.

2.2. Credibility Theory

When setting the premium for a motor insurance policy one would apply objective rating criteria like district, mileage, make of car, etc. However, there could still be individual differences left that are not captured by the rating criteria that we apply. It is then tempting to utilize the claim history of a policy when rating that policy; we apply experience rating. Credibility theory originated in 1918 within the framework of experience rating.

Let us consider a portfolio of policies that have the same values of all the objective rating criteria. Then they should in principle be equal. However, we want to take into account that there are still individual differences that are not captured by the objective rating criteria. In our model, we can take care of such differences by assuming that the individual risk characteristics of a policy are represented by an unknown random risk parameter, and that the risk parameters of different policies are independent and identically distributed.

Let us consider one of these policies. We denote by X_i the claim amount from the ith year the policy is in force. We want the (pure) premium $\dot{X}_{n+1}$ for year n+1 to be as close to X_{n+1} as possible, according to expected quadratic loss, that is, we want to minimize $\mathrm{E}\,(\dot{X}_{n+1} - X_{n+1})^2$. The *credibility estimator* is the best linear estimator $\tilde{X}_{n+1}$ based on the observed claim amounts, and it is determined by the normal equations

$$\mathrm{Cov}\left(\tilde{X}_{n+1} - X_{n+1}, X_i\right) = 0 \quad (i = 1, \ldots, n)$$

$$\mathrm{E}\left(\tilde{X}_{n+1} - X_{n+1}\right) = 0.$$

Let ${}_n\mathbf{X} = \left(X_1, ..., X_n\right)'$. If $\mathrm{Cov}\,{}_n\mathbf{X}$ is invertible, then

$$\tilde{X}_{n+1} = \mathrm{E}\;X_{n+1} + \mathrm{Cov}\;(X_{n+1},\; {}_n\mathbf{X}')\,(\mathrm{Cov}\;{}_n\mathbf{X})^{-1}\;({}_n\mathbf{X} - \mathrm{E}\;{}_n\mathbf{X}). \tag{1}$$

We see that the coefficients of $\tilde{X}_{n+1}$ depend on the joint distribution of the claim amounts only through moments of first and second order. These moments can be estimated from claims data from the policies within the portfolio. Thus, we are within the framework of empirical Bayes philosophy. We first develop the credibility estimator, a linearized Bayes estimator of a random quantity. Then we estimate parameters of this estimator from collateral data and insert these estimators in the expression for the credibility estimator to obtain an *empirical credibility estimator*.

In this set-up we have assumed that there exists a portfolio of policies with independent and identically distributed risk parameters, that is, the distribution of the risk parameters has a frequentist interpretation. This distribution represents the risk structure of the portfolio and is called the *structure distribution*. However, we can also consider credibility estimators within a pure Bayesian framework. Then we consider only one policy. In this setting the distribution of the risk parameter represents the subjective knowledge of the statistician.

The most well-known credibility model is when the claim amounts are conditionally independent and identically distributed given the unknown random risk parameter Θ. We then obtain

$$\tilde{X}_{n+1} = \frac{n}{n+\kappa}\bar{X}_n + \frac{\kappa}{n+\kappa}\mu$$

with

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i; \quad \kappa = \frac{\varphi}{\lambda}$$

$$\mu = \mathrm{E}\, X_1; \quad \varphi = \mathrm{E}\ \mathrm{Var}[X_1 | \Theta]; \quad \lambda = \mathrm{Var}\ \mathrm{E}[X_1 | \Theta].$$

For parameter estimation within this model, we assume that we have N independent and identical policies observed over n years with n, $N > 1$. We let X_{ij} denote the claim amount of policy i in year j and introduce

$$X_{i\cdot} = \frac{1}{n}\sum_{j=1}^{n} X_{ij}; \quad X_{\cdot\cdot} = \frac{1}{N}\sum_{i=1}^{N} X_{i\cdot}.$$

Then

$$\mu^* = X_{\cdot\cdot}; \quad \varphi^* = \frac{1}{N(n-1)}\sum_{i=1}^{N}\sum_{j=1}^{n}(X_{ij} - X_{i\cdot})^2$$

$$\hat{\lambda} = \frac{1}{N-1}\sum_{i=1}^{N}(X_{i\cdot} - X_{\cdot\cdot})^2 - \frac{\varphi^*}{n}$$

are unbiased estimators of μ, φ and λ respectively. However, as $\hat{\lambda}$ can take negative values whereas λ is always non-negative, we suggest to replace it with $\lambda^* = \max\ (\hat{\lambda},\ 0)$.

Issues handled within more elaborate credibility models include time-dependent risk volumes, regression, hierarchical structures, random risk parameters that develop over time, truncation, etc.

3.3. Bonus Systems

In motor insurance, experience rating is often performed with *bonus systems*. In a bonus system there is normally a limited number K of *bonus classes*.

The premium in class i is $\pi(i)$. We assume that all policies enter the bonus system in the *initial class k*. After a year in class i with j claims, the policy is transferred to class $T(i,\ j)$. The classes are ordered in the sense that $T(i,\ j) \geq T(i,\ j+1)$ and $T(i,\ j) \leq T(i+1,\ j)$, that is, one wants the better drivers in the higher classes, and consequently we also want the bonus function π to be decreasing.

A bonus system is often called a bonus-malus system to emphasize that in low classes the premium can be higher than in the initial class; malus is simply negative bonus. Bonus systems seem to have originated in Britain before 1930.

Let us consider a policy and denote the number of claims in year i by M_i, the total amount of these claims by X_i and the bonus class by Z_i. We obviously have

$$Z_i = \begin{cases} k & (i=1) \\ T(Z_{i-1}, M_{i-1}) & (i = 2, 3, \ldots) \end{cases}$$

It is assumed that the pairs (M_i, X_i) are conditionally independent and identically distributed given an unknown random risk parameter Θ that represents unknown risk characteristics of the policy.

For each n, we want the premium $\pi(Z_n)$ to be as close as possible to the claim amount X_n, and once more we use expected quadratic loss as measure of distance. A problem when minimizing $Q_n(\pi, T, k) = \mathrm{E}\left(X_n - \pi(Z_n)\right)^2$ is then that the bonus function π will depend on n. This problem can be solved by considering the age of the policy as a random variable N independent of the (M_i, X_i)s and Θ. Then we determine π so that on the average within the portfolio, the premium of a policy should be as close as possible to the claims this premium is paid to cover; we want to minimize $Q(\pi, T, k) = \mathrm{E}\, Q_N(\pi, T, k)$. For fixed T and k the optimal π is then given by $\pi_{T,k}(i) = \mathrm{E}\left[X_N \mid Z_N = i\right]$ for i = 1,…, K. The bonus function is often restricted to the class of linear functions, and from (1) we then obtain the optimal function

$$\pi_{T,k}(i) = \mathrm{E}\ X_N + \frac{\mathrm{Cov}\ (X_N,\ Z_N)}{\mathrm{Var}\ Z_N}\ (Z_N - \mathrm{E}\, Z_N). \qquad (i = 1, \ldots,\ K)$$

To find reasonable k and T one should apply trial and error.

A complication when trying to determine optimal experience rating schemes, is that the experience rating scheme may affect the claim pattern. This can happen in two ways:

1. Moral hazard. The policyholder takes more care (e.g. drives more carefully) to obtain lower premium.
2. Bonus hunger. To avoid higher premium, the policyholder does not report small claims.

3.4. Collective Risk Theory

In collective risk theory one considers the risk process of the portfolio as a whole without relating the claims to the individual policies. One is interested in the behavior of processes like $\{N(t)\}_{t\geq 0}$, $\{X(t)\}_{t\geq 0}$ and $\{U(t)\}_{t\geq 0}$ where $N(t)$, $X(t)$ and $U(t)$ denote, respectively, the number of claims occurred in the time interval $(0, t]$, the aggregate amount of these claims and the reserve of the portfolio at time t; by reserve in this connection, one means the initial reserve at time zero plus the premium income minus occurred claims, possibly with return. Particular emphasis has been put on calculation of the *ruin probability*, the probability that the reserve at some time becomes negative.

The following result is central in ruin theory.

Theorem 1: (*Lundberg's inequality*). *Let* V_1, $V_2,\ldots$ *be independent and identically distributed random variables with* $\mathrm{E}V_1 < 0$ *and* $\Pr\,(V_1 > 0) > 0$*, and assume that there exists an* $R > 0$ *such that* $\mathrm{E}\ e^{R V_1} = 1$. *Then*

$$\Pr\ \left(\bigcup\nolimits_{n=0}^{\infty} \left(\sum\nolimits_{i=1}^{n} V_i > u \right) \right) \leq e^{-Ru}. \quad (u \geq 0)$$

The constant R is called the *adjustment coefficient.*

Let Y_n denote the amount of the nth claim and W_n the interoccurrence time between the $(n-1)$th and the nth claims. We assume that the pairs (Y_1, W_1), $(Y_2, W_2),\ldots$ are independent and identically distributed. The premium is paid continuously at constant rate P, and we do not take return into account. Under these assumptions the risk process is called a *Sparre Ander*sen *process*. The occurrence time for the nth claim is $T_n = \sum_{i=1}^{n} W_i$, and we obtain $N(t) = \max\{n : T_n \leq t\}$ and $X(t) = \sum_{i=1}^{N(t)} Y_i$. When not taking return into account, the reserve at time t becomes $U(t) = u + Pt - X(t)$ with u denoting the initial reserve at time zero. As ruin can occur only at the time of occurrence of a claim, we obtain by application of Lundberg's inequality with $V_i = Y_i - PW_i$

$$\psi(u) = \Pr\left(\bigcup_{t>0}^{\infty} (U(t) < 0) \right) =$$
$$\Pr\left(\bigcup_{n=1}^{\infty} (U(T_n) < 0) \right) \leq e^{-Ru}, \ (u \geq 0)$$

provided that the adjustment coefficient exists. The condition $\mathrm{E}V_1 < 0$ now becomes $P > \frac{\mathrm{E}Y_1}{\mathrm{E}W_1,}$ that is, we must have a positive risk loading.

We see that Lundberg's upper bound decreases when R increases. Thus the adjustment coefficient can be considered as a measure of risk; the lower R, the higher risk. In some situations maximization of R can be applied to optimize the choice of deductibles and reinsurance arrangements.

If

$$\Pr\ (W_1 \leq w) = 1 - e^{-\lambda w}, \qquad (w \geq 0;\ \ \lambda\ >\ 0)$$

then the claim number process is a Poisson process. In this case the adjustment coefficient is determined by

$$1 + R\frac{P}{\lambda} - \mathrm{E}\ e^{RY_1} = 0\,.$$

In this case we have $\psi(0) = \frac{\lambda\,\mathrm{E}\,Y_1}{P}$ and *Cramér's asymptotic result*

$$\lim_{u\uparrow\infty} \psi(u)\, e^{Ru} = \frac{P - \lambda\,\mathrm{E}\,Y_1}{\lambda\,\mathrm{E}\,Y_1\ \ e^{RY_1} - P},$$

which is exact when the claim amounts are exponentially distributed.

There exists a vast literature on ruin theory. The situation usually becomes much more complicated when one leaves the Poisson assumption or introduces return on investments.

-
-
-

Bibliography

Bowers, N.L., Gerber, H.U., Hickman, J.C., Jones, D.A. & Nesbitt, C.J. (1997). *Actuarial Mathematics*. (2. ed.), 753 pp. Itasca, USA: Society of Actuaries. [Standard reference on life insurance mathematics.]

De Vylder, F. (1996). *Advanced risk theory*: *A self-contained introduction*. 969 pp. Brussels, Belgium: Editions de l'Université de Bruxelles.

de Wit, G.W. (1986). Special issue on credibility. *Insurance: Abstract and Reviews* **2-3**. 95 pp. [Bibliography on credibility with abstracts.]

Kaas, R., van Heerwaarden, A.E. & Goovaerts, M.J. (1994). *Ordering of actuarial risks*. 144 pp. Brussels, Belgium: Caire Education Series 1, Caire. [Discusses stop loss ordering and other orderings of risks.]

Rolski, T., Schmiedli, H., Schmidt, V. & Teugels, J. (1999). *Stochastic processes for insurance and finance*. 680 pp. New York, USA: Wiley.

Sundt, B. (1999). *An introduction to non-life insurance mathematics*. (4.ed.), 222 pp. Karlsruhe, Germany. Verlag Versicherungswirtschaft GmbH. [Textbook on non-life insurance mathematics similar in style to the present article.]

Taylor, G. (2000). *Loss Reserving. An actuarial perspective*. 408 pp. Boston, USA: Kluwer Academic Publishers. [Monograph on claims reserving.]

Wolthuis, H. (1994). *Life insurance mathematics (the Markovian model)*. 255 pp. Brussels, Belgium: Caire.

Biographical Sketch

Bjørn Sundt was born in Oslo in 1951. In 1978 he got the degree cand.real. in statistics at the University of Oslo and in 1981 dr.sc.math. at the Swiss Federal Institute of Technology in Zurich. He has held positions at universities and in the insurance industry, actuarial consulting and insurance supervision. At present he works for the Norwegian life insurance company Vital. The main areas of his research are aggregate claims distributions and credibility theory. He has written a text-book on non-life insurance mathematics.