Non-life insurance mathematics

Nils F. Haavardsson, University of Oslo and DNB Skadeforsikring

Last lecture….

- Exame 2011 problem 1, 2 and 3 (1.5-2h)
- Repetition, highlighting of important topics from pensum and advice for exame (0.5-1h)
- Some brief words about the assignment (0.5h)

About the exame 1

- 4th of December, 1430-1830
- Bring an approved calculator
- Bring no books or notes

About the exame 2

- The exame aims to reflect the focus of the course, which has been practical and focused on the application of statistical techniques in general insurance
- However, since this is a course at the Department of Mathematics, there should be some mathematics in the exame
- The exame aims to be comprehensive, i.e., cover as many topics from the pensum as possible
- There will be 3 practical and 1 theoretical task
- The exame aims to test understanding of important concepts from the course

STK 4540 - main issues

- The concept of diversification and risk premium
- How can claim frequency be modelled?
- How can claim size be modelled?
- How can solvency be simulated?
- Pricing in general insurance by regression
- Pricing in general insurance by credibility theory
- Reduction of risk in general insurance using re-insurance

Insurance works because risk can be diversified away through size

- The core idea of insurance is risk spread on many units
- Assume that policy risks $X_1,\ldots,X_J$ are stochastically independent
- Mean and variance for the portfolio total are then

$$E(\chi) = \pi_1 + \ldots + \pi_J \quad \text{and} \quad \text{var}(\chi) = \sigma_1 + \ldots + \sigma_J$$

$$\text{and} \quad \pi_j = E(X_j) \quad \text{and} \quad \sigma_j = sd(X_j). \text{ Introduce}$$

$$\bar{\pi} = \frac{1}{J}(\pi_1 + \ldots + \pi_J) \quad \text{and} \quad \bar{\sigma}^2 = \frac{1}{J}(\sigma_1 + \ldots + \sigma_J)$$

which is average expectation and variance. Then

$$E(\chi) = J\bar{\pi} = \quad \text{and} \quad \text{sd}(\chi) = \sqrt{J}\,\bar{\sigma} \text{ so that } \frac{\text{sd}(\chi)}{\text{E}(\chi)} = \frac{\bar{\sigma}/\bar{\pi}}{\sqrt{J}}$$

- The *coefficient of variation* approaches 0 as J grows large (law of large numbers)
- Insurance risk can be diversified away through size
- Insurance portfolios are still not risk-free because
    - of uncertainty in underlying models
    - risks may be dependent

Risk premium expresses cost per policy and is important in pricing

- Risk premium is defined as P(Event)*Consequence of Event
- More formally

$$\text{Risk premium} = P(\text{event}) * \text{Consequence of event}$$

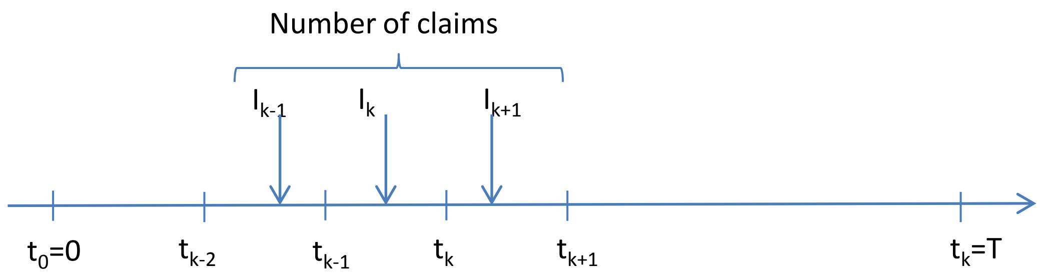
$$= \text{Claim frequency} * \text{Claim severity}$$

$$= \frac{\text{Number of claims}}{\text{Number of risk years}} * \frac{\text{Total claim amount}}{\text{Number of claims}}$$

$$= \frac{\text{Total claim amount}}{\text{Number of risk years}}$$

- From above we see that risk premium expresses cost per policy
- Good price models rely on sound understanding of the risk premium
- We start by modelling claim frequency

The world of Poisson (Chapter 8.2)



- What is rare can be described mathematically by cutting a given time period T into K small pieces of equal length h=T/K
- On short intervals the chance of more than one incident is remote
- Assuming no more than 1 event per interval the count for the entire period is

$N=I_1+...+I_K$, where I_j is either 0 or 1 for j=1,...,K

- If $p=\Pr(I_k=1)$ is equal for all k and events are independent, this is an ordinary Bernoulli series

$$\Pr(N=n)=\frac{K!}{n!(K-n)!}p^n(1-p)^{K-n}, \quad \text{for} \quad n=0,1,\ldots,K$$

- Assume that p is proportional to h and set $p=\mu h$ where μ is an intensity which applies per time unit

The world of Poisson

$$\Pr(N=n)=\frac{K!}{n!(K-n)!}p^n(1-p)^{K-n}$$

$$=\frac{K!}{n!(K-n)!}\left(\frac{\mu T}{K}\right)^n\left(1-\frac{\mu T}{K}\right)^{K-n}$$

$$=\frac{(\mu T)^n}{n!}\underbrace{\frac{K(K-1)\cdots(K-n+1)}{K^n}}_{\underset{K\to\infty}{\to}1}\underbrace{\left(1-\frac{\mu T}{K}\right)^K}_{\underset{K\to\infty}{\to}e^{-\mu T}}\underbrace{\frac{1}{\left(1-\frac{\mu T}{K}\right)^n}}_{\underset{K\to\infty}{\to}1}$$

$$\Rightarrow \Pr(N=n)\underset{K\to\infty}{\to}\frac{(\mu T)^n}{n!}e^{-\mu T}$$

In the limit N is Poisson distributed with parameter $\lambda=\mu T$

The world of Poisson

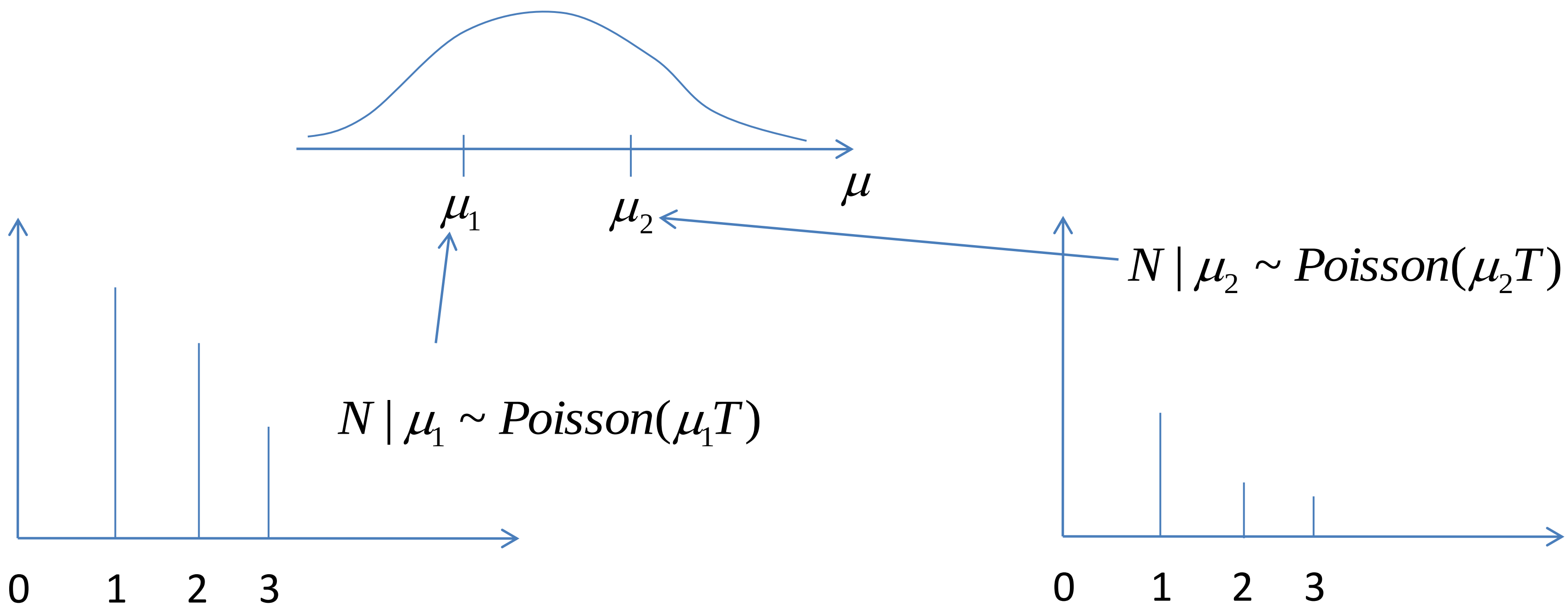
- It follows that the portfolio number of claims **N** is Poisson distributed with parameter

$$\lambda = (\mu_1 + \ldots + \mu_J)T = J\bar{\mu}T, \quad \text{where } \bar{\mu} = (\mu_1 + \ldots + \mu_J)/J$$

- When claim intensities vary over the portfolio, only their average counts

Random intensities (Chapter 8.3)

- How μ varies over the portfolio can partially be described by observables such as age or sex of the individual (treated in Chapter 8.4)
- There are however factors that have impact on the risk which the company can't know much about
  - Driver ability, personal risk averseness,
- This randomeness can be managed by making μ a stochastic variable



Random intensities (Chapter 8.3)

- The models are conditional ones of the form

$$N \mid \mu \sim Poisson(\mu T) \quad \text{and} \quad \mathbf{N} \mid \mu \sim Poisson(J\mu T)$$

Policy level — Portfolio level

- Let $\xi = E(\mu)$ and $\sigma = \mathrm{sd}(\mu)$ and recall that $E(N \mid \mu) = \mathrm{var}(N \mid \mu) = \mu T$

  which by double rules in Section 6.3 imply

$$E(N) = E(\mu T) = \xi T \quad \text{and} \quad \mathrm{var}(N) = E(\mu T) + \mathrm{var}(\mu T) = \xi T + \sigma^2 T^2$$

- Now E(N)<var(N) and N is no longer Poisson distributed

The Poisson regression model (Section 8.4)

- The idea is to attribute variation in μ to variations in a set of observable variables $x_1,\ldots,x_v$. Poisson regressjon makes use of relationships of the form

$$\log(\mu) = b_0 + b_1 x_1 + \ldots + b_v x_v \qquad (1.12)$$

- Why $\log(\mu)$ and not μ itself?
- The expected number of claims is non-negative, where as the predictor on the right of (1.12) can be anything on the real line
- It makes more sense to transform μ so that the left and right side of (1.12) are more in line with each other.

- Historical data are of the following form

Claims	exposure	covariates
n_1	T_1	$x_{11}\ldots x_{1x}$
n_2	T_2	$x_{21}\ldots x_{2x}$
n_n	T_n	$x_{n1}\ldots x_{nv}$

- The coefficients $b_0,\ldots,b_v$ are usually determined by likelihood estimation

The model (Section 8.4)

- In likelihood estimation it is assumed that n_j is Poisson distributed $\lambda_j = \mu_j T_j$ where μ_j is tied to covariates $x_{j1},...,x_{jv}$ as in (1.12). The density function of n_j is then

$$f(n_j) = \frac{(\mu_j T_j)^{n_j}}{n_j!} \exp(-\mu_j T_j)$$

or

$$\log(f(n_j)) = n_j \log(\mu_j) + n_j \log(T_j) - \log(n_j!) - \mu_j T_j$$

- $\log(f(n_j))$ above is to be added over all j for the likehood function $L(b_0,...,b_v)$.
- Skip the middle terms $n_j T_j$ and $\log(n_j!)$ since they are constants in this context.
- Then the likelihood criterion becomes

$$L(b_0,...,b_v) = \sum_{j=1}^{n} \{n_j \log(\mu_j) - \mu_j T_j\} \text{ where } \log(\mu_j) = b_0 + b_1 x_{j1} + ... + b_j x_{jv} \quad (1.13)$$

- Numerical software is used to optimize (1.13).
- McCullagh and Nelder (1989) proved that $L(b_0,...,b_v)$ is a convex surface with a single maximum
- Therefore optimization is straight forward.

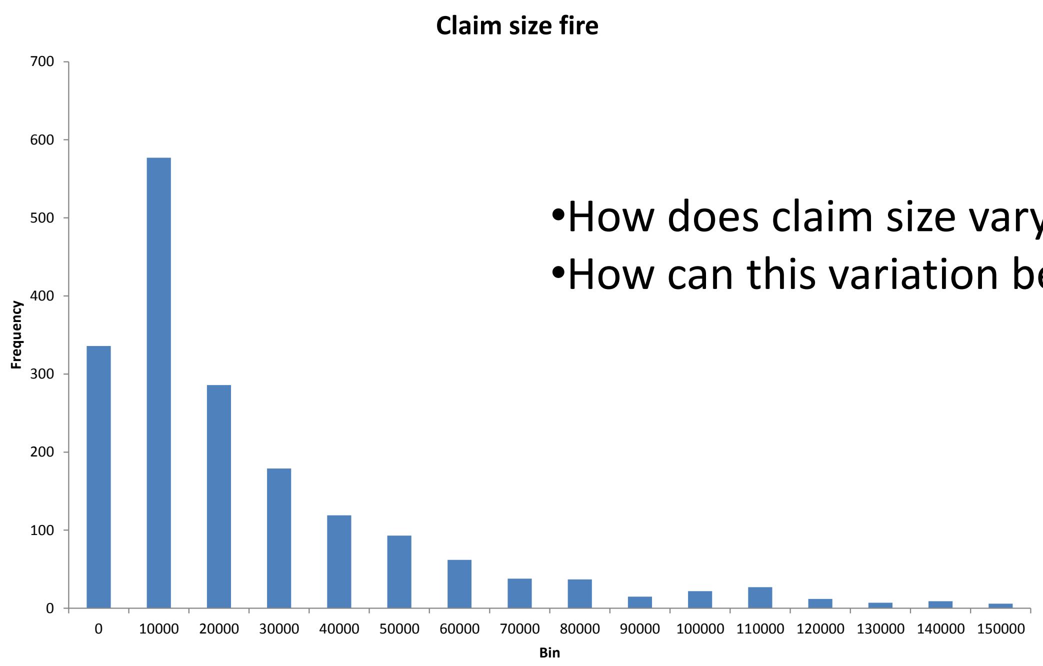
Repetition claim size

- The concept
- Non parametric modelling
- Scale families of distributions
- Fitting a scale family
- Shifted distributions
- Skewness
- Non parametric estimation
- Parametric estimation: the log normal family
- Parametric estimation: the gamma family
- Parametric estimation: fitting the gamma

Claim severity modelling is about describing the variation in claim size

The concept

- The graph below shows how claim size varies for fire claims for houses
- The graph shows data up to the 88th percentile



- How does claim size vary?
- How can this variation be modelled?

- Truncation is necessary (large claims are rare and disturb the picture)
- 0-claims can occur (because of deductibles)
- Two approaches to claim size modelling – non-parametric and parametric

Non-parametric modelling can be useful

- Claim size modelling can be *non-parametric* where each claim z_i of the past is assigned a probability 1/n of re-appearing in the future
- A new claim is then envisaged as a random variable $\hat{Z}$ for which

$$\Pr(\hat{Z} = z_i) = \frac{1}{n}, \quad \mathrm{i} = 1, \ldots, \mathrm{n}$$

- This is an entirely proper probability distribution
- It is known as *the empirical distribution* and will be useful in Section 9.5.

Non-parametric modelling can be useful

- All sensible parametric models for claim size are of the form

$$Z = \beta Z_0, \quad \text{where } \beta > 0 \text{ is a parameter}$$

- and Z0 is a standardized random variable corresponding to $\hat{Z}$. $\beta = 1$
- The large the scale parameter, the more spread out the distribution

$$Z = \beta Z_0, Z_0 \sim N(1,1)$$

$$\beta = 1 \Rightarrow Z \sim N(1,1)$$

$$\beta = 2 \Rightarrow Z \sim N(2,2)$$

$$\beta = 3 \Rightarrow Z \sim N(3,3)$$

Fitting a scale family

- Models for scale families satisfy

$$\Pr(Z \le z) = \Pr(Z_0 \le z/\beta) \quad \text{or} \quad F(z|\beta) = F_0(z/\beta)$$

where $F(z|\beta)$ and $F_0(z/\beta)$ are the distribution functions of Z and Z_0.

- Differentiating with respect to z yields the family of density functions

$$f(z|\beta) = \frac{1}{\beta} f_0\left(\frac{z}{\beta}\right), \; z > 0 \quad \text{where} \quad f_0(z|\beta) = \frac{dF_0(z)}{dz}$$

- The standard way of fitting such models is through likelihood estimation. If $z_1,\ldots,z_n$ are the historical claims, the criterion becomes

$$L(\beta, f_0) = -n\log(\beta) + \sum_{i=1}^{n} \log\{f_0(z_i/\beta)\},$$

which is to be maximized with respect to β and other parameters.

- A useful extension covers situations with *censoring*.

Fitting a scale family

- Full value insurance:
  - The insurance company is liable that the object at all times is insured at its true value
- First loss insurance
  - The object is insured up to a pre-specified sum.
  - The insurance company will cover the claim if the claim size does not exceed the pre-specified sum

- The chance of a claim Z exceeding b is $1-F_0(b/\beta)$ and for nb such events with lower bounds b1,…,bnb the analogous joint probability becomes

$$\{1-F_0(b_1/\beta)\}x\ldots x\{1-F_0(b_{n_b}/\beta)\}.$$

Take the logarithm of this product and add it to the log likelihood of the fully observed claims $z_1,\ldots,z_n$. The criterion then becomes

$$L(\beta, f_0) = -n\log(\beta) + \sum_{i=1}^{n}\log\{f_0(z_i/\beta)\} + \sum_{i=1}^{n_b}\log\{1-F_0(z_i/\beta)\},$$

complete information (for objects fully insured)

censoring to the right (for first loss insured)

Shifted distributions

- The distribution of a claim may start at some treshold b instead of the origin.
- Obvious examples are deductibles and re-insurance contracts.
- Models can be constructed by adding b to variables starting at the origin; i.e. where Z_0 is a standardized variable as before. Now

$$\Pr(Z \le z) = \Pr(b + \beta Z_0 \le z) = \Pr(Z_0 \le \frac{z-b}{\beta})$$

- Example:
  - Re-insurance company will pay if claim exceeds 1 000 000 NOK

$$Z = 1000000 + \beta Z_0$$

The payout of the insurance company

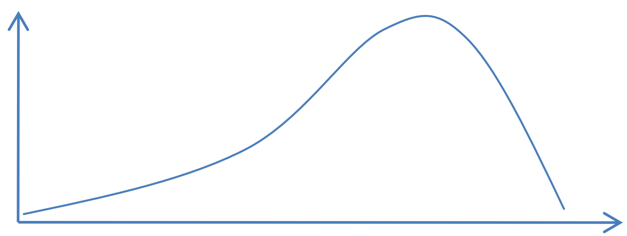
Total claim amount

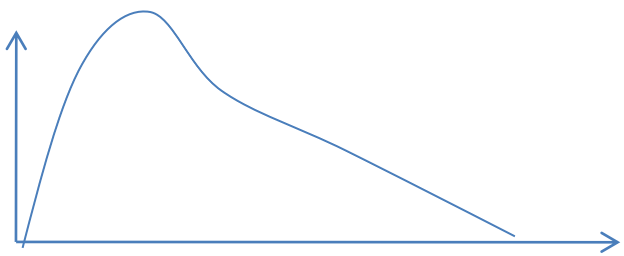
Currency rate for example NOK per EURO, for example 8 NOK per EURO

Skewness as simple description of shape

- A major issue with claim size modelling is asymmetry and the right tail of the distribution. A simple summary is the coefficient of skewness

$$\varsigma = skew(Z) = \frac{\nu^3}{\sigma^3} \quad \text{where } \nu^3 = E(Z-\xi)^3$$





Negative skewness　　　Positive skewness

Negative skewness: the left tail is longer; the mass of the distribution
Is concentrated on the right of the figure. It has relatively few low values

Positive skewness: the right tail is longer; the mass of the distribution
Is concentrated on the left of the figure. It has relatively few high values

Non-parametric estimation

- The random variable $\hat{Z}$ that attaches probabilities 1/n to all claims z_i of the past is a possible model for *future* claims.
- Expectation, standard deviation, skewness and percentiles are all closely related to the ordinary sample versions. For example

$$E(\hat{Z}) = \sum_{i=1}^{n} \Pr(\hat{Z} = z_i) z_i = \sum_{i=1}^{n} \frac{1}{n} z_i = \bar{z}.$$

- Furthermore,

$$\mathrm{var}(\hat{Z}) = E(\hat{Z} - E(\hat{Z}))^2 = \sum_{i=1}^{n} \Pr(\hat{Z} = z_i)(z_i - \bar{z})^2 = \sum_{i=1}^{n} \frac{1}{n}(z_i - \bar{z})^2$$

$$\Rightarrow sd(\hat{Z}) = \sqrt{\frac{n-1}{n}} s, \quad \mathrm{s} = \sqrt{\frac{1}{\mathrm{n-1}} \sum_{i=1}^{n} (z_i - \bar{z})^2}$$

- Third order moment and skewness becomes

$$\hat{\nu}_3(\hat{Z}) = \frac{1}{n} \sum_{i=1}^{n} (z_i - \bar{z})^3 \quad \text{and} \quad \mathrm{skew}(\hat{Z}) = \frac{\hat{\nu}_3(\hat{Z})}{\{sd(\hat{Z})\}^3}$$

The log-normal family

- A convenient definition of the log-normal model in the present context is as $Z = \xi Z_0$ where $Z_0 = e^{-\sigma^2/2+\sigma\varepsilon}$ for $\varepsilon \sim N(0,1)$
- Mean, standard deviation and skewness are

$$E(Z) = \xi, \quad \mathrm{sd}(Z) = \xi\sqrt{e^{\sigma^2}-1}, \quad skew(Z) = (e^{\sigma^2}+2)\sqrt{e^{\sigma^2}-1}$$

see section 2.4.

- Parameter estimation is usually carried out by noting that logarithms are Gaussian. Thus

$$Y = \log(Z) = \log(\xi) - 1/2\sigma^2 + \sigma\varepsilon$$

and when the original log-normal observations $z_1,\ldots,z_n$ are transformed to Gaussian ones through $y_1=\log(z_1),\ldots,y_n=\log(z_n)$ with sample mean and variance $\bar{y}$ and s_y, the estimates of ξ and σ become

$$\log(\hat{\xi}) - 1/2\hat{\sigma}^2 = \bar{y}, \quad \hat{\sigma} = s_y \quad \text{or} \quad \hat{\xi} = e^{s_y^2/2+\bar{y}}, \hat{\sigma} = s_y.$$

The Gamma family

- The Gamma family is an important family for which the density function is

$$f(x) = \frac{(\alpha/\xi)^{\alpha}}{\Gamma(\alpha)} x^{\alpha-1} e^{-\alpha x/\xi}, \quad x > 0, \text{ where } \Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{-x} dx$$

- It was defined in Section 2.5 as $Z = \xi G$ where $G \sim \text{Gamma}(\alpha)$ is the standard Gamma with mean one and shape alpha. The density of the standard Gamma simplifies to

$$f(x) = \frac{\alpha^{\alpha}}{\Gamma(\alpha)} x^{\alpha-1} e^{-\alpha x}, \quad x > 0, \text{ where } \Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{-x} dx$$

Mean, standard deviation and skewness are

$$E(Z) = \xi, \quad \text{sd}(Z) = \xi/\sqrt{\alpha}, \quad \text{skew}(Z) = 2/\sqrt{\alpha}$$

and there is a convolution property. Suppose $G_1,\dots,G_n$ are independent with $G_i \sim Gamma(\alpha_i)$. Then

$$\overline{G} \sim Gamma(\alpha_1 + \dots + \alpha_n) \quad \text{if} \quad \overline{G} = \frac{\alpha_1 G_1 + \dots + \alpha_n G_n}{\alpha_1 + \dots + \alpha_n}$$

The Gamma family

- The Gamma family is an important family for which the density function is

$$f(x) = \frac{(\alpha/\xi)^{\alpha}}{\Gamma(\alpha)} x^{\alpha-1} e^{-\alpha x/\xi}, \quad x > 0, \text{ where } \Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{-x} dx$$

- It was defined in Section 2.5 as $Z = \xi G \quad \text{where } G \sim \text{Gamma}(\alpha)$ is the standard Gamma with mean one and shape alpha. The density of the standard Gamma simplifies to

$$f(x) = \frac{\alpha^{\alpha}}{\Gamma(\alpha)} x^{\alpha-1} e^{-\alpha x}, \quad x > 0, \text{ where } \Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{-x} dx$$

The Gamma family

$$\log(f_0(z)) = \alpha\log(\alpha) - \log\Gamma(\alpha) + (\alpha-1)\log(z) - \alpha z$$

$$Z = \xi G$$

$$L(\xi,\alpha) = -n\log(\xi) + \sum_{i=1}^{n}\log(f_0(z/\xi))$$

$$= -n\log(\xi) + \sum_{i=1}^{n}\alpha\log(\alpha) - \log\Gamma(\alpha) + (\alpha-1)\log(z_i/\xi) - \alpha z/\xi$$

$$= -n\log(\xi) + n\alpha\log(\alpha) - n\log\Gamma(\alpha) + (\alpha-1)\sum_{i=1}^{n}\log(z_i/\xi) - \alpha/\xi\sum_{i=1}^{n}z_i$$

$$= -n\log(\xi) + n\alpha\log(\alpha) - n\log\Gamma(\alpha) + (\alpha-1)\sum_{i=1}^{n}\log(z_i)$$

$$+(\alpha-1)(-n\log(\xi)) - \alpha/\xi\sum_{i=1}^{n}z_i$$

$$= n\alpha\log(\alpha/\xi) - n\log\Gamma(\alpha) + (\alpha-1)\sum_{i=1}^{n}\log(z_i) - \alpha/\xi\sum_{i=1}^{n}z_i$$

Solvency

- Financial control of liabilities under nearly worst-case scenarios
- Target: the *reserve*
  – which is the upper percentile of the portfolio liability
- Modelling has been covered (Risk premium calculations)
- The issue now is computation
  – Monte Carlo is the general tool
  – Some problems can be handled by simpler, Gaussian approximations

10.2 Portfolio liabilities by simple approximation

- The portfolio loss for independent risks become Gaussian as J tends to infinity.
- Assume that policy risks $X_1,\ldots,X_J$ are stochastically independent
- Mean and variance for the portfolio total are then

$$E(\chi) = \pi_1 + \ldots + \pi_J \quad \text{and} \quad \text{var}(\chi) = \sigma_1 + \ldots + \sigma_J$$

$$\text{and} \quad \pi_j = E(X_j) \quad \text{and} \quad \sigma_j = sd(X_j). \text{ Introduce}$$

$$\bar{\pi} = \frac{1}{J}(\pi_1 + \ldots + \pi_J) \quad \text{and} \quad \bar{\sigma}^2 = \frac{1}{J}(\sigma_1 + \ldots + \sigma_J)$$

which is average expectation and variance. Then

$$\frac{1}{J}\sum_{i=1}^{J} X_i \xrightarrow{d} N(\bar{\pi}, \bar{\sigma}^2)$$

as J tends to infinfity

- Note that risk is underestimated for small portfolios and in branches with large claims

Normal approximations

Let μ be claim intensity and ξ_z and σ_z mean and standard deviation of the individual losses. If they are the same for all policy holders, the mean and standard deviation of $\mathcal{X}$ over a period of length T become

$$E(\mathcal{X}) = a_0 J, \quad sd(\mathcal{X}) = a_1 \sqrt{J}$$

where

$$a_0 = \mu T \xi_z \quad \text{and} \quad a_1 = \sqrt{\mu T}\sqrt{\sigma_z^2 + \xi_z^2}$$

The rule of double variance

Let X and Y be arbitrary random variables for which

$$\xi(x) = E(Y \mid x) \quad \text{and} \quad \sigma^2 = \mathrm{var}(Y \mid x)$$

Then we have the important identities

$$\underbrace{\xi = E(Y) = E\{\xi(X)\}}_{\text{Rule of double expectation}} \quad \text{and} \quad \underbrace{\mathrm{var(Y)} = \mathrm{E}\{\sigma^2(X)\} + \mathrm{var}\{\xi(X)\}}_{\text{Rule of double variance}}$$

The rule of double variance

Portfolio risk in general insurance

$$\chi = Z_1 + Z_2 + \ldots + Z_\mathrm{N} \quad \text{where} \quad \mathrm{N}, Z_1, Z_2, \ldots \text{ are stochastically independent.}$$

$$\text{Let } E(Z_1) = \xi_Z \quad \text{where } \mathrm{var}(\chi \mid \mathrm{N}) = \mathrm{N}\sigma_Z^2$$

Elementary rules for random sums imply

$$E(\chi \mid \mathrm{N} = N) = N\xi_z \text{ and } \mathrm{var}(\chi \mid \mathrm{N} = N) = N\sigma_z^2$$

$$\text{Let } Y = \chi \quad \text{and} \quad \mathrm{x} = \mathrm{N} \text{ in the formulas on the previous slide}$$

$$\mathrm{var}(\chi) = \mathrm{var}\{E(\chi \mid \mathrm{N} = N)\} + E\{sd(\chi \mid \mathrm{N} = N)\}^2$$

$$= \mathrm{var}(\mathrm{N}\xi_z) + E(\mathrm{N}\sigma_z^2)$$

$$= \xi_z^2 \,\mathrm{var}(\mathrm{N}) + \sigma_z^2 E(\mathrm{N})$$

$$= J\mu T(\xi_z^2 + \sigma_z^2)$$

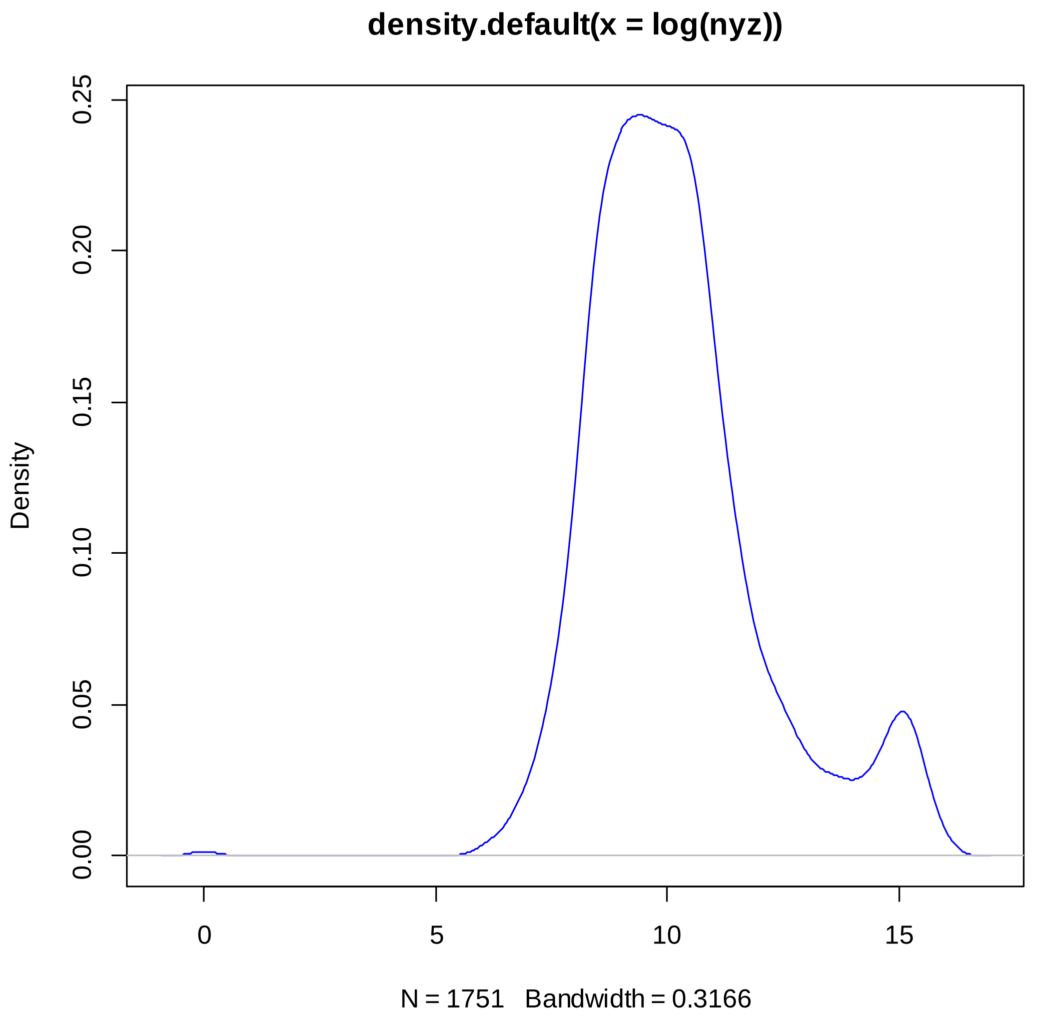
The rule of double variance

This leads to the true percentile qepsilon being approximated by

$$q_{\varepsilon}^{NO} = a_0 J + a_1 \phi_{\varepsilon} \sqrt{J}$$

Where phi epsilon is the upper epsilon percentile of the standard normal distribution

Fire data from DNB



Portfolio liabilities by simulation

- Monte Carlo simulation
- Advantages
  - More general (no restriction on use)
  - More versatile (easy to adapt to changing circumstances)
  - Better suited for longer time horizons
- Disadvantages
  - Slow computationally?
  - Depending on claim size distribution?

An algorithm for liabilities simulation

- Assume claim intensities $\mu_1, \ldots, \mu_J$ for J policies are stored on file
- Assume J different claim size distributions and payment functions $H_1(z), \ldots, H_J(z)$ are stored
- The program can be organised as follows (Algorithm 10.1)

$$\text{Input}: \lambda_j = \mu_j T \, (j = 1, .., J), \quad \text{claim size models}, \quad H_1(z), \ldots, H_J(z)$$

$$X^* \leftarrow 0$$

$$\text{For } j = 1, \ldots, J \quad \text{do}$$

$$\quad \text{Draw } U^* \sim \mathit{Uniform} \quad \text{and} \quad S^* \leftarrow -\log(U^*)$$

$$\quad \text{Repeat while } S^* < \lambda_j$$

$$\quad\quad \text{Draw claim size } Z^*$$

$$\quad\quad X^* \leftarrow X^* + H_j(z)$$

$$\quad\quad \text{Draw } U^* \sim \mathit{Uniform} \quad \text{and} \quad S^* \leftarrow S^* - \log(U^*)$$

$$\text{Return } X^*$$

Experiments in R

1. Log normal distribution
2. Gamma on log scale
3. Pareto
4. Weibull
5. Mixed distribution 1
6. Monte Carlo algorithm for portfolio liabilities
7. Mixed distribution 2

Comparison of results

Percentile	95 %	99 %	99.97%
Normal approximations	19 025 039	22 962 238	29 347 696
Normal power approximations	20 408 130	26 540 012	38 086 350
Monte Carlo algorithm log normal claims	12 650 847	24 915 297	102 100 605
Monte Carlo algorithm gamma model for log claims	88 445 252	401 270 401	6 327 665 905
Monte Carlo algorithm mixed empirical and Weibull	20 238 159	24 017 747	30 940 560
Monte Carlo algorithm empirical distribution	19 233 569	24 364 595	32 387 938

Monte Carlo theory

Suppose X_1, X_2,… are independent and exponentially distributed with mean 1. It can then be proved

$$\Pr(X_1 + \ldots + X_n < \lambda \leq X_1 + \ldots + X_{n+1}) = \frac{\lambda^n}{n!} e^{-\lambda} \qquad (1)$$

for all n >= 0 and all lambda > 0.

- From (1) we see that the exponential distribution is the distribution that describes time between events in a Poisson process.
- In Section 9.3 we learnt that the distribution of $X_1+\ldots+X_n$ is gamma distributed with mean n and shape n
- The Poisson process is a process in which events occur continuously and independently at a constant average rate
- The Poisson probabilities on the right define the density function

$$\Pr(N = n) = \frac{\lambda^n}{n!} e^{-\lambda}, n = 0,1,2,\ldots$$

which is the central model for claim numbers in property insurance.
Mean and standard deviation are E(N)=lambda and sd(N)=sqrt(lambda)

Monte Carlo theory

It is then utilized that X_j=-log(U_j) is exponential if U_j is uniform, and the sum X_1+X_2+… is monitored until it exceeds lambda, in other words

Algorithm 2.14 Poisson generator

0 Input: λ

1 $Y^* \leftarrow 0$

2 For $n = 1,2,\ldots$ do

3 Draw $\mathrm{U}^* \sim \mathit{Uniform}$ and $Y^* \leftarrow Y^* - \log(\mathrm{U}^*)$

4 *I*f $\mathrm{Y}^* \geq \lambda$ then

5 stop and return $\mathrm{N}^* \leftarrow n-1$

Monte Carlo theory

Proof of Algorithm 2.14 Let $X_1,...,X_{n+1}$ be stochastically independent with common density function $f(x) = e^{-x}$ for $x > 0$ and let $S_n = X_1 + ... + X_n$.
The Poisson generator of Algorithm 2.14 is based on the probability

$$p_n(\lambda) = \Pr(S_n < \lambda \le S_{n+1}) \qquad (1)$$

which can be evaluated by conditioning on S_n. If its density function is $f_n(s)$, then

$$p_n(\lambda) = \int_0^\lambda \Pr(\lambda \le s + X_{n+1} \mid S_n = s) f_n(s)ds = \int_0^\lambda e^{-(\lambda - s)} f_n(s)ds.$$

But S_n is Gamma distributed with mean $\xi = \mathrm{n}$ and shape $\alpha = \mathrm{n}$ (section 9.3). This means that $f_n(s) = s^{n-1}e^{-s}/(n-1)!$ and

$$p_n(\lambda) = \int_0^\lambda e^{-(\lambda - s)} s^{n-1} e^{-s}/(n-1)!ds = \frac{e^{-\lambda}}{(n-1)!}\int_0^\lambda s^{n-1}ds = \frac{\lambda^n}{n!}e^{-\lambda}$$

as was to be proved.

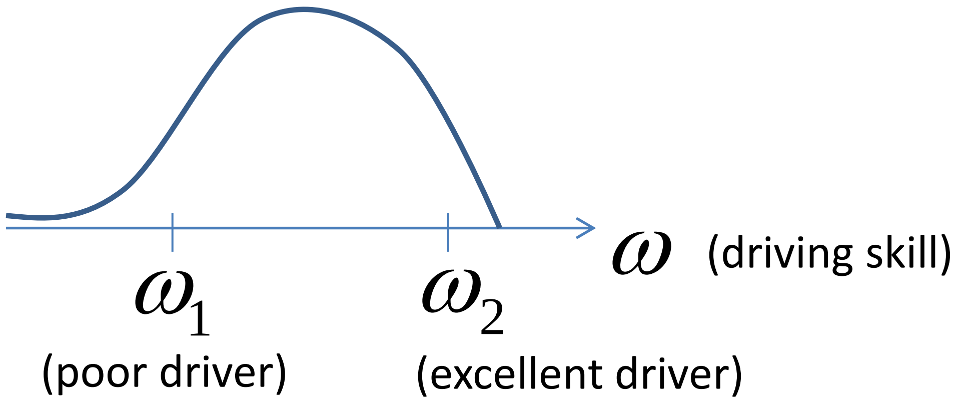
The credibility approach

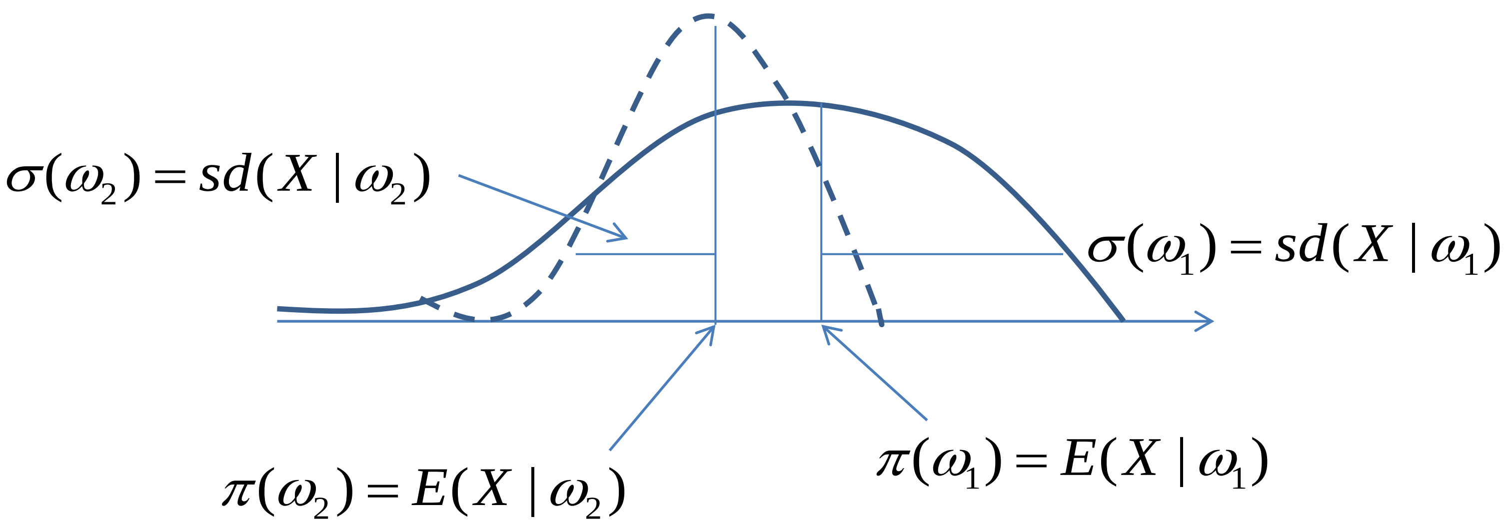
- The basic assumption is that policy holders carry a list of attributes with ω impact on risk
- The parameter ω could be how the car is used by the customer (degree of recklessness) or for example driving skill
- It is assumed that ω exists and has been drawn randomly for each individual
- X is the sum of claims during a certain period of time (say a year) and introduce

$$\pi(\omega) = E(X \mid \omega) \quad \text{and} \quad \sigma(\omega) = \text{sd}(X \mid \omega)$$

- We seek $\pi = \pi(\omega)$ the conditional pure premium of the policy holder as basis for pricing
- On group level there is a common ω that applies to all risks jointly
- We will focus on the individual level here, as the difference between individual and group is minor from a mathematical point of view

Omega is random and has been drawn for each policy holder





The most accurate estimate

- Let $X_1,\ldots,X_K$ (policy level) be realizations of X dating K years back
- The most accurate estimate of pi from such records is (section 6.4) the conditional mean

$$\hat{\pi}_K = E(X \mid x_1,\ldots,x_K)$$

where $x_1,\ldots,x_K$ are the actual values.

- A natural framework is the common factor model of Section 6.3 where $X, X_1,\ldots,X_K$ are identically and independently distributed given omega
- This won't be true when underlying conditions change systematically
- A problem with the estimate above is that it requires a joint model for $X, X_1,\ldots,X_K$ and omega.
- A more natural framework is to break X down on claim number N and losses per incident Z.
- First linear credibility is considered

Linear credibility

- The standard method in credibility is the linear one with estimates of pi of the form

$$\hat{\pi}_K = b_0 + b_1 X_1 + \ldots + b_K X_K$$

where $b_0,\ldots,b_K$ are coefficients so that the mean squared error $E(\hat{\pi}_K - \pi)^2$ is as small as possible.

- The fact that $X, X_1,\ldots,X_K$ are conditionally independent with the same distribution forces $b_1=\ldots=b_K$, and if w/K is their common value, the estimate becomes

$$\hat{\pi}_K = b_0 + w\overline{X}_K \quad \text{where} \quad \overline{X}_K = (X_1 + \ldots + X_K)/K$$

- To proceed we need the so-called *structural parameters*

$$\overline{\pi} = E\{\pi(\omega)\}, \quad \nu^2 = \mathrm{var}\{\pi(\omega)\}, \quad \tau^2 = E\{\sigma^2(\omega)\}$$

where $\overline{\pi}$ is the average pure premium for the entire population.

- It is also the expectation for individuals since by the rule of double expectation

$$E(X) = E\{E\{X \mid \omega\}\} = E\{\pi(\omega)\} = \overline{\pi}$$

Linear credibility

- Both ν and τ represent variation. The former is caused by diversity between individuals and the latter by the physical processes behind the incidents.Their impact on var(X) can be understood through the rule of double variance, i.e.,

$$\text{var}(X) = E\{\text{var}(X \mid \omega)\} + \text{var}\{E(X \mid \omega)\}$$
$$= E\{\sigma^2(\omega)\} + \text{var}\{\pi(\omega)\} = \tau^2 + \nu^2$$

and ν^2 and τ^2 represent uncertainties of different origin that add to var(X)

- The optimal linear credibility estimate now becomes

$$\hat{\pi}_K = (1-w)\overline{\pi} + w\overline{X}_K, \quad \text{where} \quad w = \frac{\nu^2}{\nu^2 + \tau^2 / K}$$

which is proved in Section 10.7, where it is also established that

$$E(\hat{\pi}_K - \pi) = 0 \quad \text{and that} \quad \text{sd}(\hat{\pi}_K - \pi) = \frac{\nu}{\sqrt{1 + K\nu^2/\tau^2}}.$$

- The estimate is ubiased and its standard deviation decreases with K

Linear credibility

- The weight w defines a compromise between the average pure premium pi bar of the population and the track record of the policy holder
- Note that w=0 if K=0; i.e. without historical information the best estimate is the population average