Non-life insurance mathematics

Nils F. Haavardsson, University of Oslo and DNB Skadeforsikring

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

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

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

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

Then we have the important identities

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

The rule of double variance

Portfolio risk in general insurance

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

Let $E(\mathrm{Z}_1) = \xi_Z$ where $\mathrm{var}(\chi \mid \mathrm{N}) = \mathrm{N}\sigma_Z^2$

Elementary rules for random sums imply

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

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

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

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

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

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

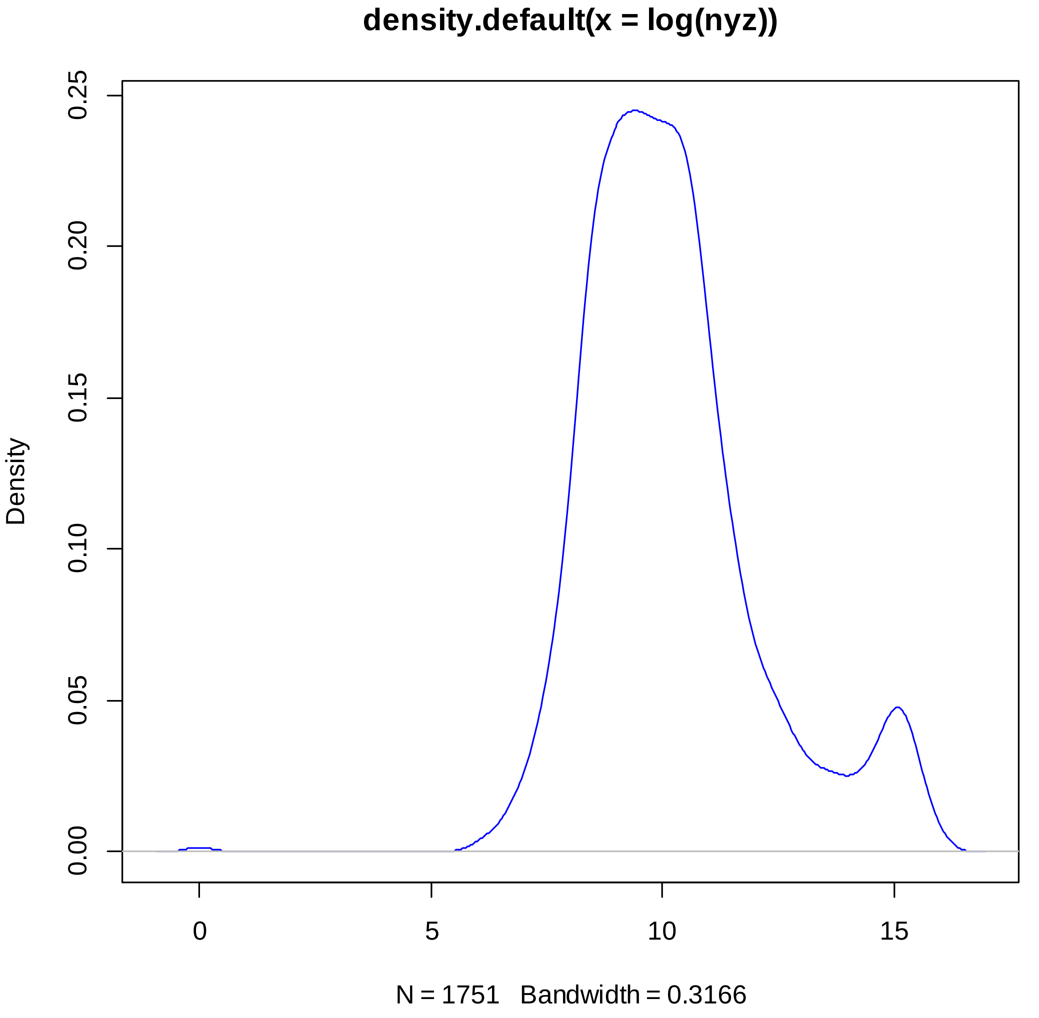
The rule of double variance

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



Normal approximations in R

```
z=scan("C:/Users/wenche_adm/Desktop/Nils/uio/Exercises/Branntest.txt");
# removes negatives
nyz=ifelse(z>1,z,1.001);

mu=0.0065;
T=1;
ksiZ=mean(nyz);
sigmaZ=sd(nyz);
a0 = mu*T*ksiZ;
a1 = sqrt(mu*T)*sqrt(sigmaZ^2+ksiZ^2);
J=5000;
qepsNO95=a0*J+a1*qnorm(.95)*sqrt(J);
qepsNO99=a0*J+a1*qnorm(.99)*sqrt(J);
qepsNO9997=a0*J+a1*qnorm(.9997)*sqrt(J);
c(qepsNO95,qepsNO99,qepsNO9997);
```

The normal power approximation

```
ny3hat  = 0;
n=length(nyz);
for (i in 1:n)
{
		ny3hat = ny3hat + (nyz[i]-mean(nyz))**3
}
ny3hat = ny3hat/n;
LargeKsihat=ny3hat/(sigmaZ**3);
a2 = (LargeKsihat*sigmaZ**3+3*ksiZ*sigmaZ**2+ksiZ**3)/(sigmaZ^2+ksiZ^2);
qepsNP95=a0*J+a1*qnorm(.95)*sqrt(J)+a2*(qnorm(.95)**2-1)/6;
qepsNP99=a0*J+a1*qnorm(.99)*sqrt(J)+a2*(qnorm(.99)**2-1)/6;
qepsNP9997=a0*J+a1*qnorm(.9997)*sqrt(J)+a2*(qnorm(.9997)**2-1)/6;
c(qepsNP95,qepsNP99,qepsNP9997);
```

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

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)

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

$$\text{X}^* \leftarrow 0$$

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

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

$$\qquad \text{Repeat while } S^* < \lambda_{\text{j}}$$

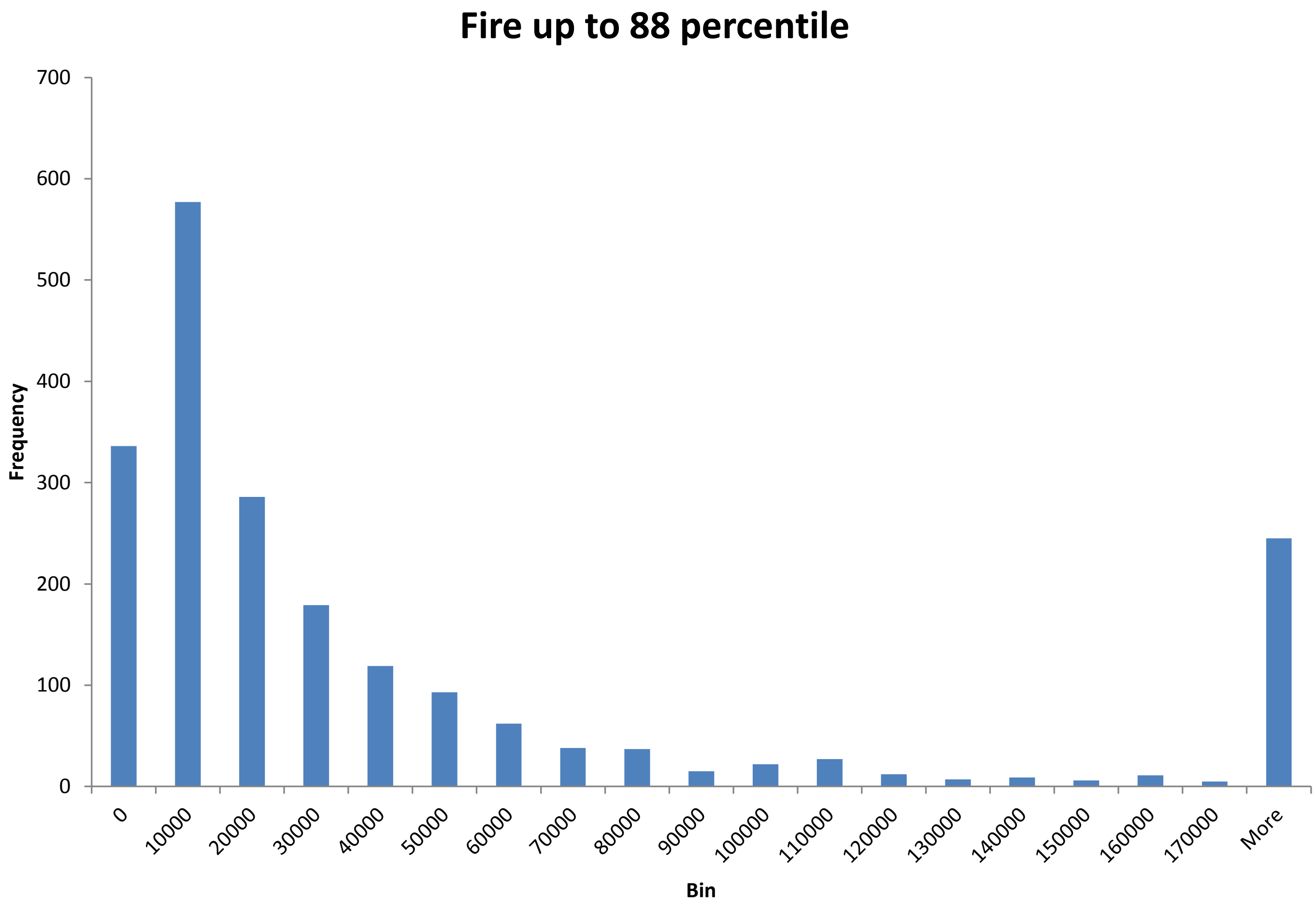
$$\qquad\qquad \text{Draw claim size Z*}$$

$$\qquad\qquad \text{X}^* \leftarrow \text{X}^* + H_j(z)$$

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

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

- Non parametric
- Log-normal, Gamma
- The Pareto
- Extreme value
- Searching



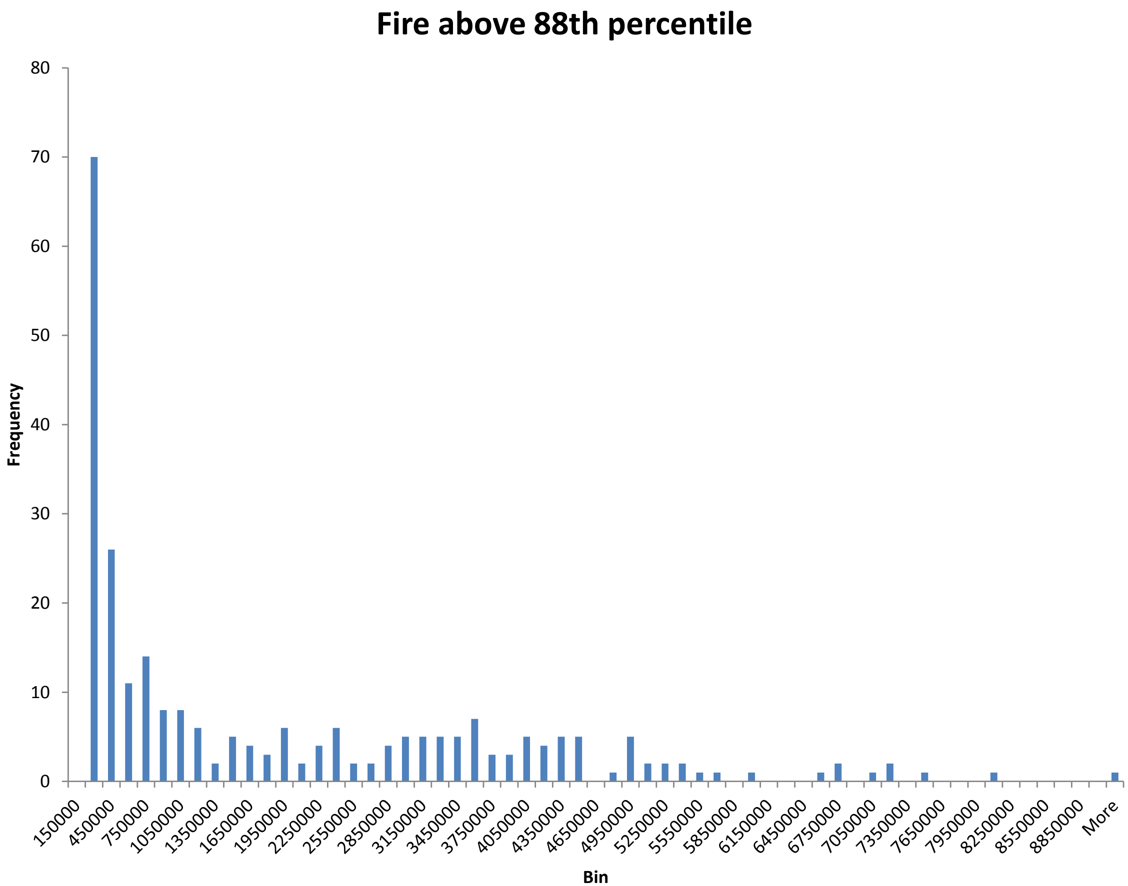
Non parametric

Log-normal, Gamma

The Pareto

Extreme value

Searching



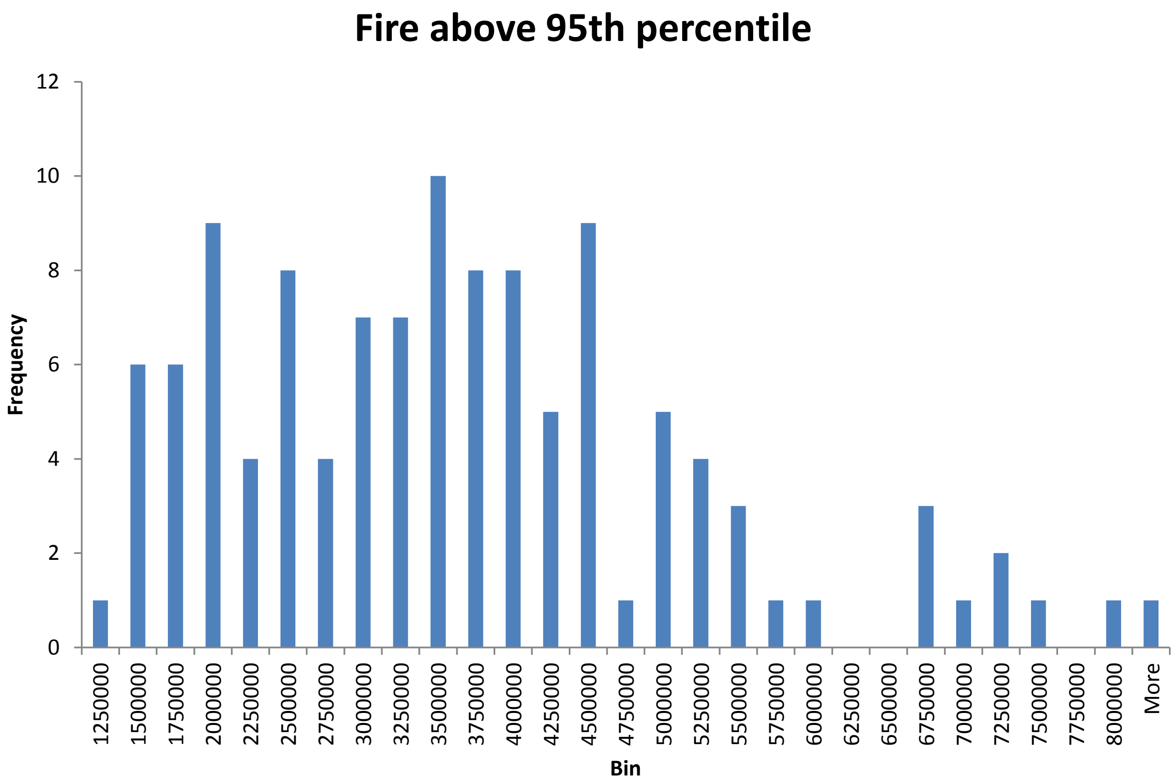
Non parametric

Log-normal, Gamma

The Pareto

Extreme value

Searching



Searching for the model

- How is the final model for claim size selected?
- Traditional tools: QQ plots and criterion comparisons
- Transformations may also be used (see Erik Bølviken’s material)

Descriptive Statistics for Variable Skadeestimat	
Number of Observations	185
Number of Observations Used for Estimation	185
Minimum	331206.17
Maximum	9099311.62
Mean	2473661.54
Standard Deviation	1892916.16

Results from top 12% modelling

Weibull is best in top 12% modelling

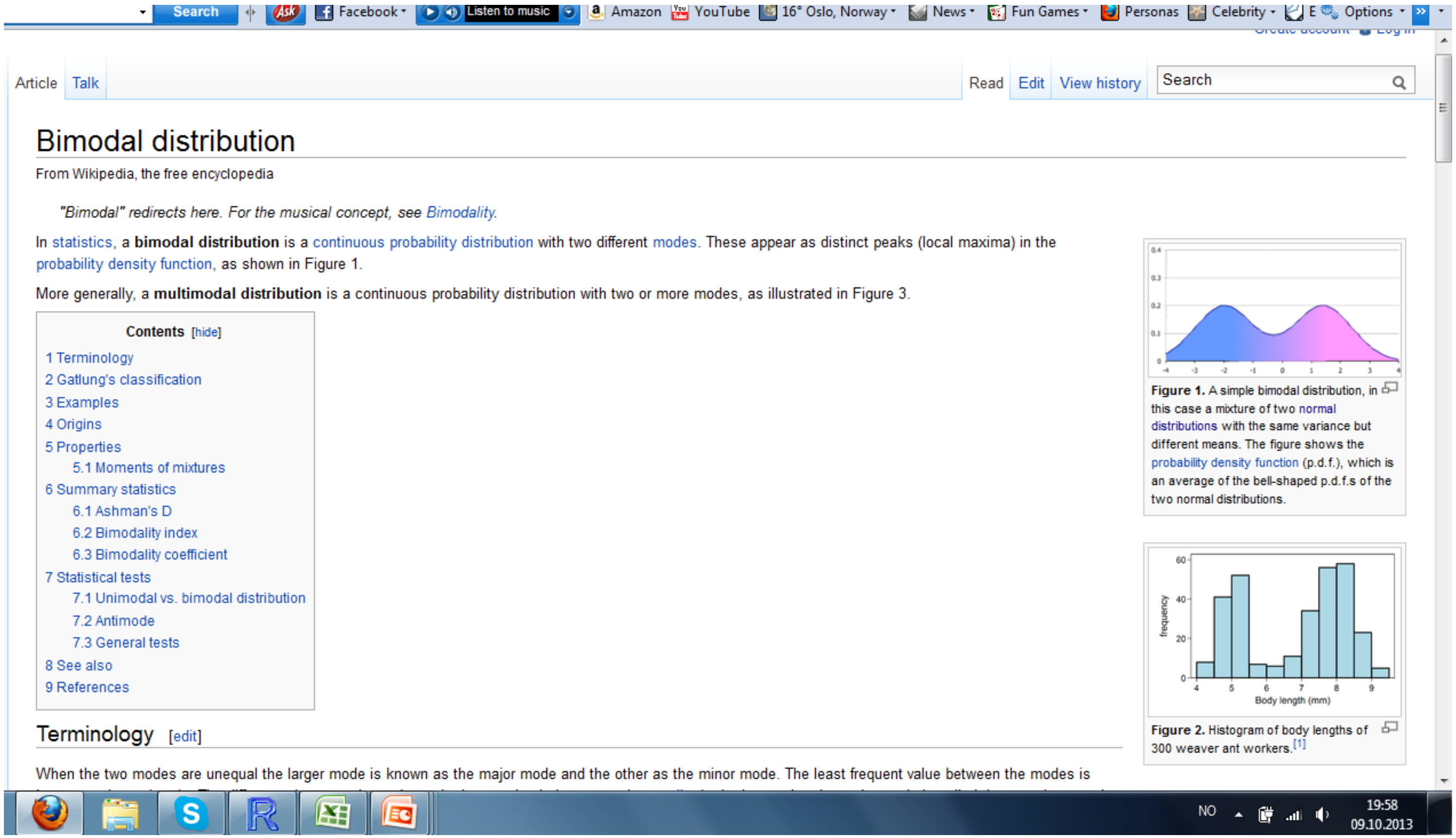
Model Selection Table

Distribution	Converged	-2 Log Likelihood	Selected
Burr	Yes	5808	No
Logn	Yes	5807	No
Exp	Yes	5817	No
Gamma	Yes	5799	No
Igauss	Yes	5804	No
Pareto	Yes	5874	No
Weibull	Yes	5799	Yes

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

Check out bimodal distributions on wikipedia



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

Or…check out mixture distributions on wikipedia

A mixture of two unimodal distributions with differing means is not necessarily bimodal. The combined distribution of heights of men and women in fact the difference in mean heights of men and women is too small relative to their standard deviations to produce bimodality.[4]

Bimodal distributions have the peculiar property that - unlike the unimodal distributions - the mean may be a more robust sample estimator than U shaped like the arcsine distribution. It may not be true when the distribution has one or more long tails.

Moments of mixtures [edit]

Let

$$f(x) = pg_1(x) + (1-p)g_2(x)$$

where g_i is a probability distribution and p is the mixing parameter.

The moments of f(x) are[8]

$$\mu = p\mu_1 + (1-p)\mu_2$$
$$\nu_2 = p[\sigma_1^2 + \delta_1^2] + (1-p)[\sigma_2^2 + \delta_2^2]$$
$$\nu_3 = p[S_1\sigma_1^3 + 3\delta_1\sigma_1^2 + \delta_1^3] + (1-p)[S_2\sigma_2^3 + 3\delta_2\sigma_2^2 + \delta_2^3]$$
$$\nu_4 = p[K_1\sigma_1^4 + 4S_1\delta_1\sigma_1^3 + 6\delta_1^2\sigma_1^2 + \delta_1^4] + (1-p)[K_2\sigma_2^4 + 4S_2\delta_2\sigma_2^3 + 6\delta_2^2\sigma_2^2 + \delta_2^4]$$

where

$$\mu = \int xf(x)dx$$
$$\delta_i = \mu_i - \mu$$
$$\nu_r = \int (x-\mu)^r f(x)dx$$

Solvency – day 2

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

Structure

- Normal approximation
- Monte Carlo Theory
- Monte Carlo Practice – an example with fire data from DNB

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

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

where

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

Normal approximations

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

Monte Carlo theory

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

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

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

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

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

4 If $\mathrm{Y}^* \geq \lambda$ 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

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

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.

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)

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

$$X^* \leftarrow 0$$

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

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

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

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

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

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

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

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