

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, \dots, X_J are stochastically independent
- Mean and variance for the portfolio total are then

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

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

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

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 infinity

- 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 X over a period of length T become

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

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

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

Then we have the important identities

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

Rule of double expectation Rule of double variance

The rule of double variance

Portfolio risk in general insurance

$\chi = Z_1 + Z_2 + \dots + Z_N$ where N, Z_1, Z_2, \dots are stochastically independent.

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

Elementary rules for random sums imply

$$E(\chi | N = N) = N\xi_Z \text{ and } \text{var}(\chi | N = N) = N\sigma_Z^2$$

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

$$\begin{aligned} \text{var}(\chi) &= \text{var}\{E(\chi | N = N)\} + E\{sd(\chi | N = N)\}^2 \\ &= \text{var}(N\xi_Z) + E(N\sigma_Z^2) \\ &= \xi_Z^2 \text{var}(N) + \sigma_Z^2 E(N) \\ &= J\mu T(\xi_Z^2 + \sigma_Z^2) \end{aligned}$$

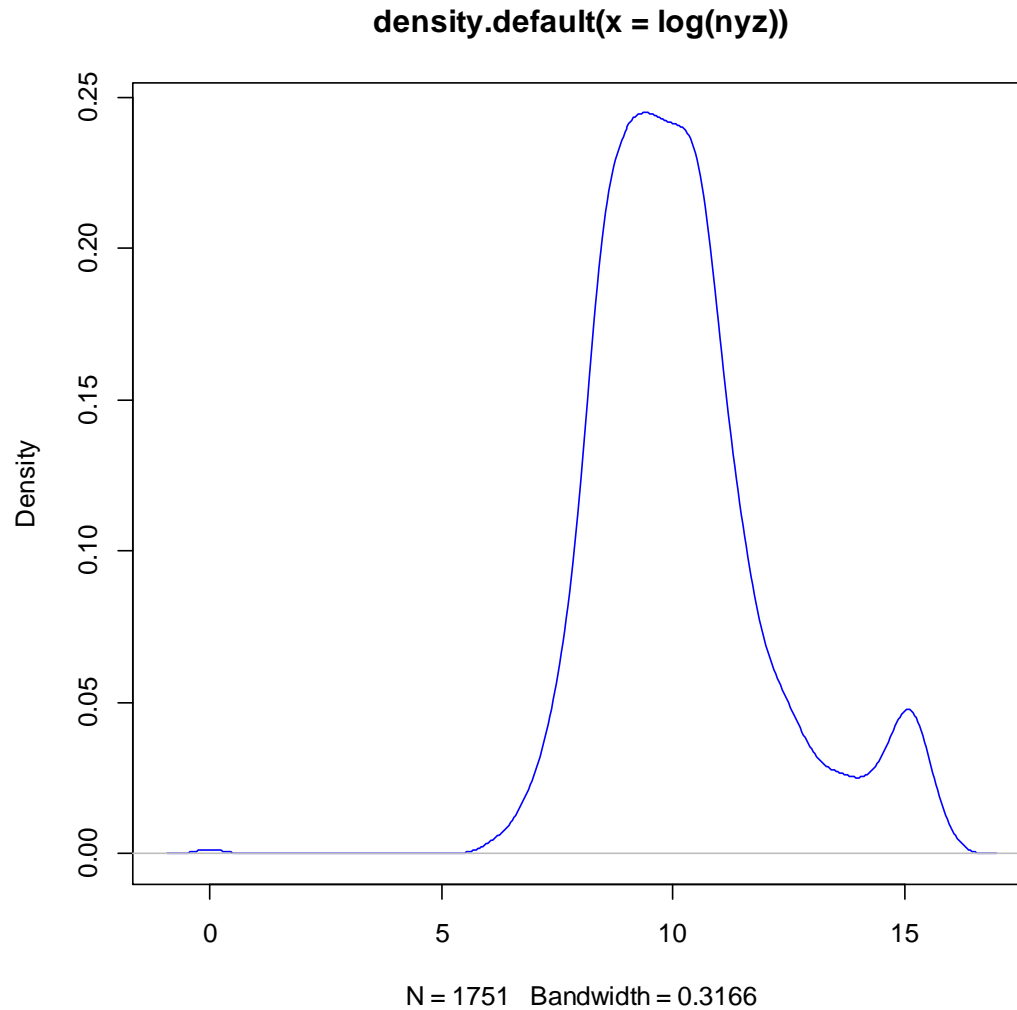
The rule of double variance

This leads to the true percentile q_ϵ being approximated by

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

Where ϕ_ϵ is the upper ϵ percentile of the standard normal distribution

Fire data from DNB



Normal approximations in R

```
z=scan("C:/Users/wenche_admin/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 μ_1, \dots, μ_J for J policies are stored on file
- Assume J different claim size distributions and payment functions $H_1(z), \dots, H_J(z)$ are stored
- The program can be organised as follows (Algorithm 10.1)

0 Input: $\lambda_j = \mu_j T$ ($j = 1, \dots, J$), claim size models, $H_1(z), \dots, H_J(z)$

1 $X^* \leftarrow 0$

2 For $j = 1, \dots, J$ do

3 Draw $U^* \sim Uniform$ and $S^* \leftarrow -\log(U^*)$

4 Repeat while $S^* < \lambda_j$

5 Draw claim size Z^*

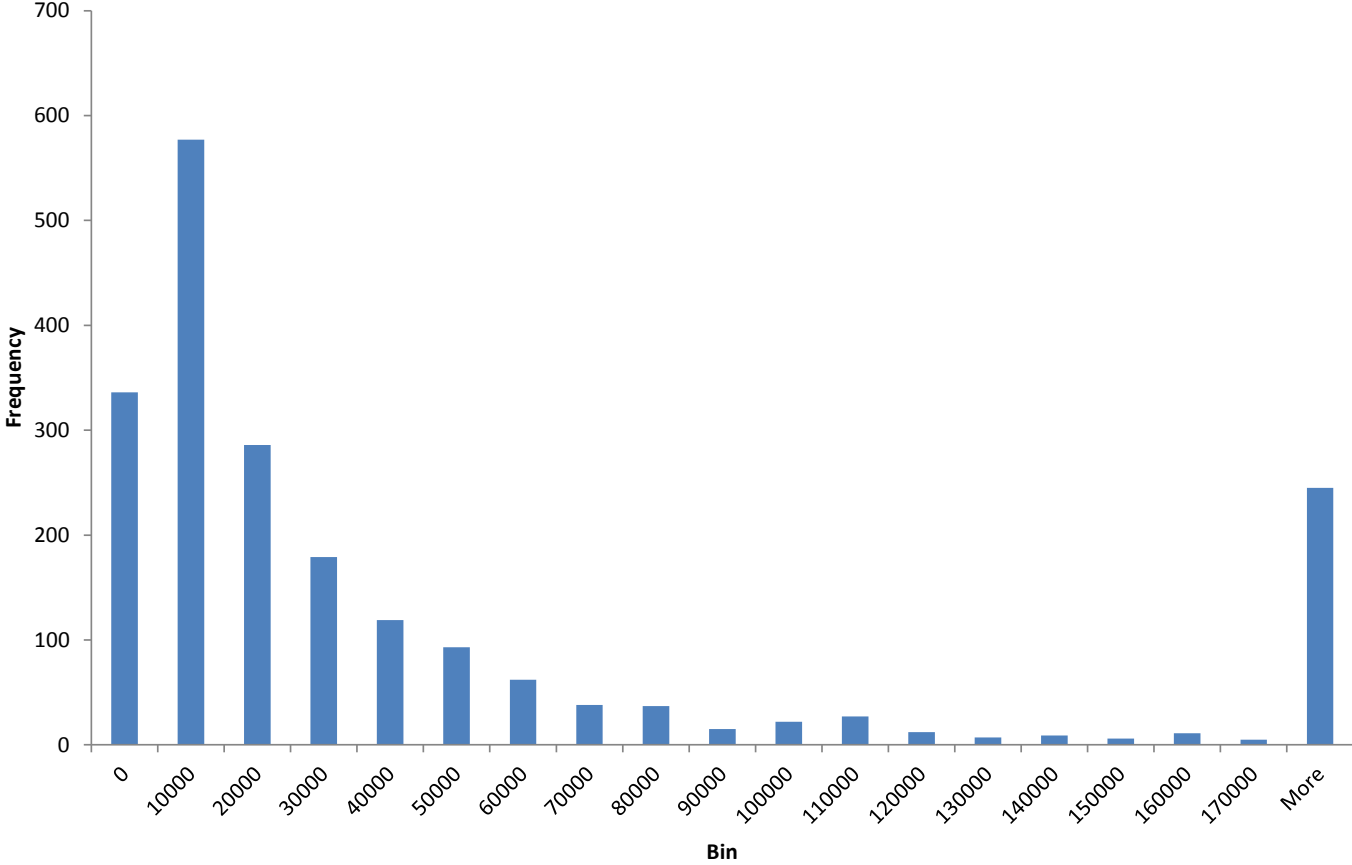
6 $X^* \leftarrow X^* + H_j(z)$

7 Draw $U^* \sim Uniform$ and $S^* \leftarrow S^* - \log(U^*)$

8 Return X^*

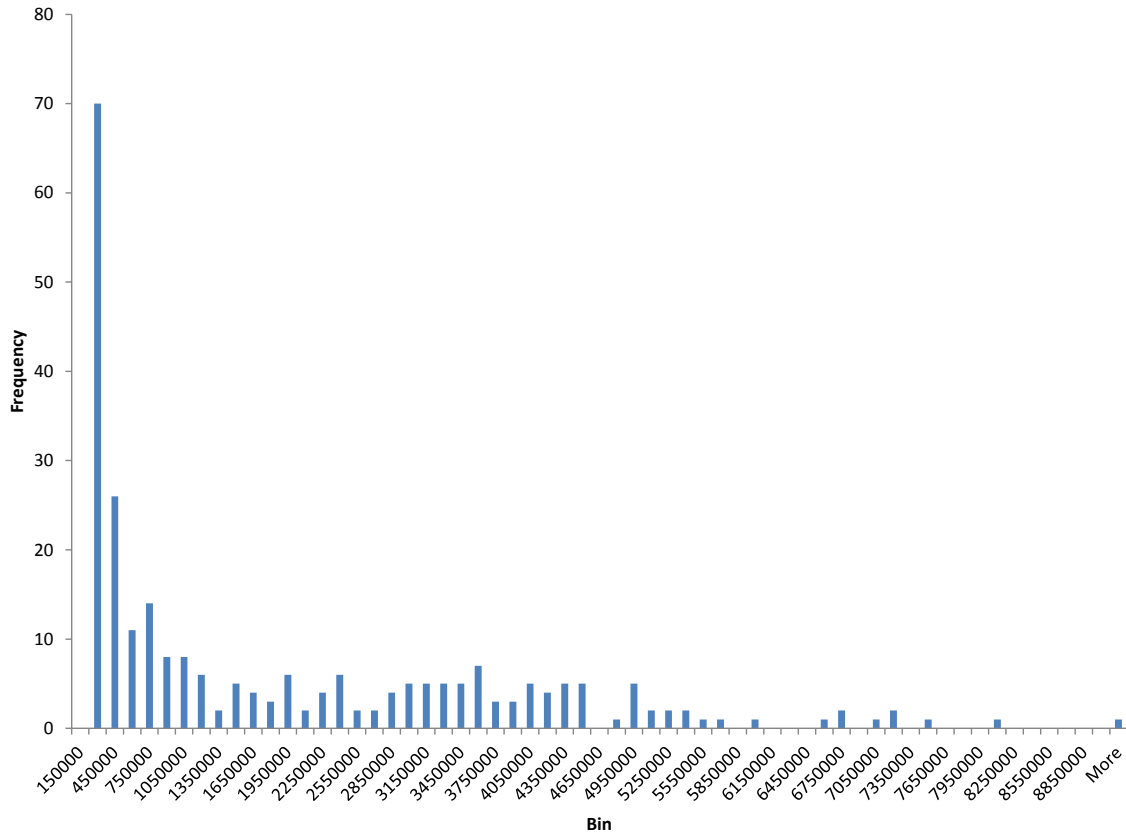
- Non parametric
- Log-normal, Gamma
- The Pareto
- Extreme value
- Searching

Fire up to 88 percentile



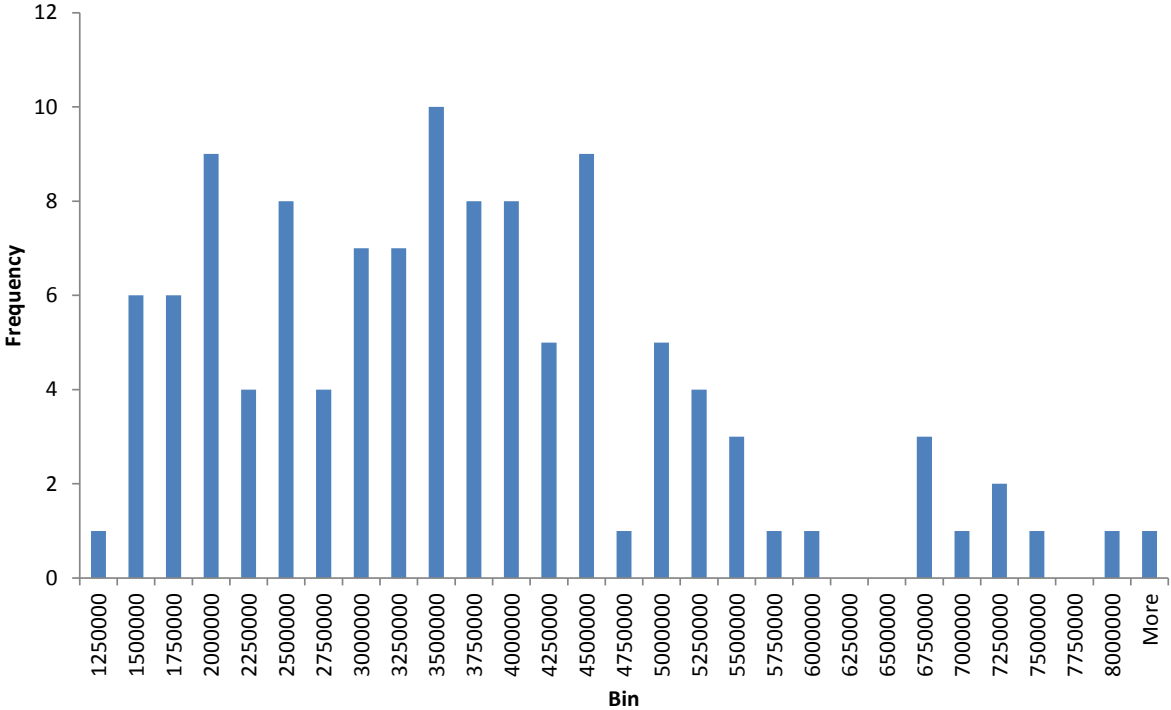
- Non parametric
- Log-normal, Gamma
- The Pareto
- Extreme value
- Searching

Fire above 88th percentile



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

Fire above 95th percentile



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)

Non parametric

Log-normal, Gamma

The Pareto

Extreme value

Searching

Non parametric

Log-normal, Gamma

The Pareto

Extreme value

Searching

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

Non parametric

Log-normal, Gamma

The Pareto

Extreme value

Searching

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

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

Check out bimodal distributions on wikipedia

Search | Facebook | Listen to music | Amazon | YouTube | 16° Oslo, Norway | News | Fun Games | Personas | Celebrity | E | Options

Article | **Talk** | Read | Edit | View history | Search

Bimodal distribution

From Wikipedia, the free encyclopedia

"Bimodal" redirects here. For the musical concept, see Bimodality.

In statistics, a **bimodal distribution** is a continuous probability distribution with two different modes. These appear as distinct peaks (local maxima) in the probability density function, as shown in Figure 1.

More generally, a **multimodal distribution** is a continuous probability distribution with two or more modes, as illustrated in Figure 3.

Contents [hide]

- 1 Terminology
- 2 Gattung's classification
- 3 Examples
- 4 Origins
- 5 Properties
 - 5.1 Moments of mixtures
- 6 Summary statistics
 - 6.1 Ashman's D
 - 6.2 Bimodality index
 - 6.3 Bimodality coefficient
- 7 Statistical tests
 - 7.1 Unimodal vs. bimodal distribution
 - 7.2 Antimode
 - 7.3 General tests
- 8 See also
- 9 References



Figure 1. A simple bimodal distribution, in this case a mixture of two normal distributions with the same variance but different means. The figure shows the probability density function (p.d.f.), which is an average of the bell-shaped p.d.f.s of the two normal distributions.



Figure 2. Histogram of body lengths of 300 weaver ant workers.^[1]

Terminology [edit]

When the two modes are unequal the larger mode is known as the major mode and the other as the minor mode. The least frequent value between the modes is

NO 19:58 09.10.2013

Or...check out mixture distributions on wikipedia

[References](#) [\[edit\]](#)

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

$$\mu = \int xf(x)dx$$

$$\delta_i = \mu_i - \mu$$

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

