INSTITUTE OF ACTUARIAL
& QUANTITATIVE STUDIES

Class: SY MSc

Subject : Statistical & Risk Modelling - 2

Chapter: Unit 1 Chapter 1

Chapter Name: Loss distributions in insurance risk management

Agenda

1. Introduction

1. Generating functions
    1. MGF
    2. CGF

2. Loss distributions
    1. Exponential distribution
    2. Gamma distribution
    3. Normal distribution
    4. Lognormal distribution
    5. Two-parameter Pareto distribution
    6. Burr distribution
    7. Three-parameter Pareto distribution
    8. Weibull distribution

4. Basic Distributional Quantities
    1. Method of moments
    2. Maximum Likelihood Estimator
    3. Method of Percentiles

6. Goodness-of-fit

1 Introduction

- **Insurance Companies**
  - Inflow – **Premiums**
    - – Certain
    - – Predetermined Amount
  - Outflow – **Claims**
    - – Random
    - – Frequency & Severity
    - – Fitting a probability distribution
    - – **LOSS DISTRIBUTIONS**
      - Premium calculation
      - Reserve calculation
      - Solvency
      - Reinsurance arrangement

2 Generating Functions - Recap

Moment Generating Functions – MGF

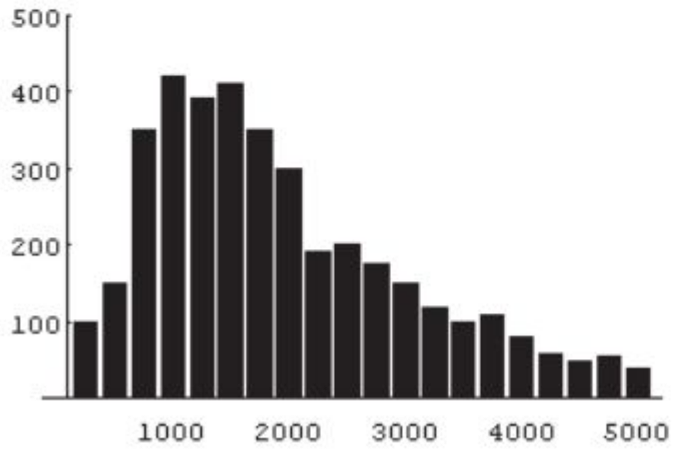
- A moment generating function (MGF) can be used to generate moments (about the origin) of the distribution of a random variable (discrete or continuous).
- Although the moments of most distributions can be determined directly by evaluation using the necessary integrals or summation, utilizing moment generating functions sometimes provides considerable simplifications.

Cumulant Generating Functions – CGF

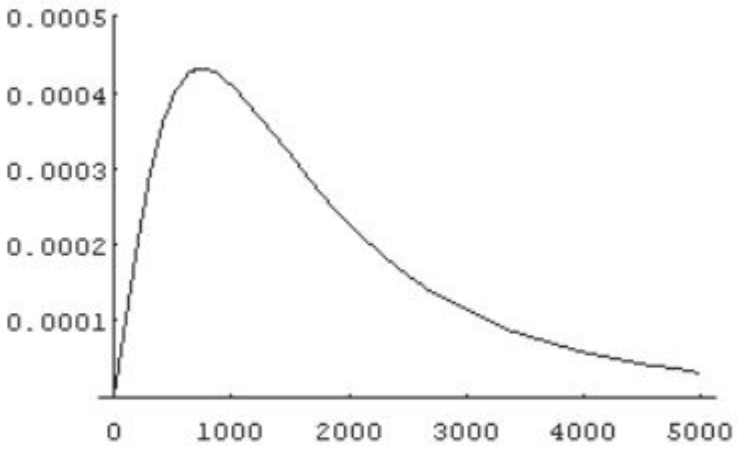
- A cumulant generating function (CGF) takes the moment of a sequence of numbers that describes the distribution in a useful, compact way.
- The first cumulant is the mean, the second the variance, and the third cumulant is the skewness or third central moment.

3 Loss Distributions

1. The frequency of claim amounts when plotted against size might look like this:



2. **The statistical distributions in this chapter are used to approximate this distribution, which is called a loss distribution.** For example, we might decide to use a loss distribution like this as an approximation to the claims arising in the graph besides:



3 Loss Distributions

At a first level, it can be assumed that the claims arise as realizations from a known distribution. For example, it may be possible to assume that the logarithm of the claim amount follows, to a reasonable approximation, a normal distribution with known mean and known standard deviation.

In practice the exact claims distribution will hardly ever be known. At this second level a standard method of proceeding is to assume that the claims distribution is a member of a certain family. The parameters of the family must now be estimated using the claim amount records by an appropriate method such as maximum likelihood.

Many studies have been made of the kind of distribution that can be used to describe the variation in claim amounts.
The typical pattern is as shown in the histogram above, with a few small claims, rising to a peak, then tailing off gradually with a few very large claims.

The general conclusion is that claims distributions tend to be positively skewed and long tailed.

3.1 Loss Distributions

The exponential distribution

A random variable $\boldsymbol{X}$ has the exponential distribution with parameter $\lambda > 0$ if it has CDF:

$$F(x) = 1 - e^{-\lambda x}, x > 0$$

In that case we write $X \sim \mathrm{Exp}(\lambda)$.
The PDF is:

$$f(x) = \lambda e^{-\lambda x}, x > 0$$

The mean and variance are $\frac{1}{\lambda}$ and $\frac{1}{\lambda^2}$ respectively.

The MGF is:

$$M(t) = \left(1 - \frac{t}{\lambda}\right)^{-1}, t < \lambda$$

3.2 Loss Distributions

The gamma distribution

The random variable X has a gamma distribution with parameters $\alpha > 0$ and $\lambda > 0$ if it has PDF:

$$f(x) = \frac{\lambda^{\alpha}}{\Gamma(\alpha)} x^{\alpha-1}\, e^{(-\lambda x)},\ x > 0$$

In that case we write $X \sim Ga(\alpha, \lambda)$.

The mean and variance of $\boldsymbol{X}$ are:

$$E(X) = \frac{\alpha}{\lambda}$$

$$\text{var}(X) = \frac{\alpha}{\lambda^2}$$

$$M_X(t) = \left(1 - \frac{t}{\lambda}\right)^{-\alpha},\ \text{for } t < \lambda$$

3.2 Loss Distributions

Relationship between gamma and chi-squared distributions

If $X \sim \text{Gamma}(\alpha, \lambda)$ and 2α is an integer, then:

$$2\lambda x \sim \chi^2_{2\alpha}$$

3.3 Loss Distributions

The normal distribution

$X \sim N(\mu, \sigma^2)$

Moments:

$$E(X) = \mu$$

$$Var\,(X) = \sigma^2$$

The MGF is:

$$M_X(t) = e^{1/2t^2}$$

The normal distribution arises in a variety of contexts. **It is of limited use for modelling loss distributions because of its symmetry (as loss distributions tend to be positively skewed).**

3.4 Loss Distributions

The lognormal distribution

The definition of the lognormal distribution is very simple: $\boldsymbol{X}$ has a lognormal distribution if $\log X$ has a normal distribution.

When $X \sim \log N(\mu, \sigma^2)$, then $Y = \log X \sim N(\mu, \sigma^2)$

$$E(X) = e^{\mu+\frac{1}{2}\sigma^2}$$

$$\mathrm{var}(X) = e^{2\mu+2\sigma^2} - \left(e^{\mu+1/2\sigma^2}\right)^2 = e^{2\mu+2\sigma^2} - e^{2\mu+\sigma^2} = e^{2\mu+\sigma^2}\left(e^{\sigma^2} - 1\right)$$

3.5 Loss Distributions

The two-parameter Pareto distribution

A random variable $\boldsymbol{X}$ has the Pareto distribution with parameters $\alpha > 0$ and $\lambda > 0$ if it has CDF:

$$F(x) = 1 - \left(\frac{\lambda}{\lambda + x}\right)^{\alpha}, x > 0$$

In that case we write $X \sim Pa(\alpha, \lambda)$.

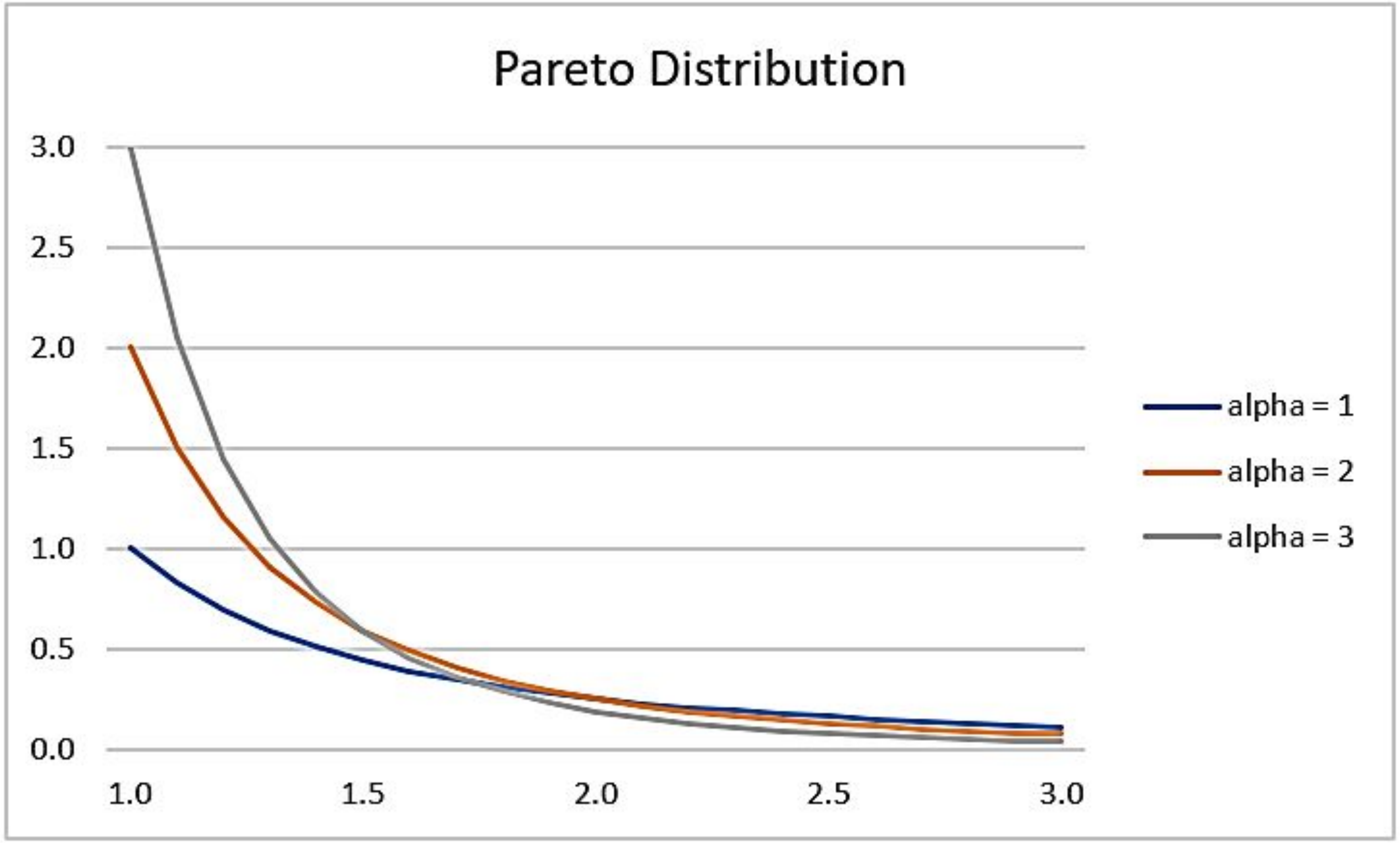
It is easily checked by differentiating $F(x)$ with respect to x that the Pareto distribution has PDF:

$$f(x) = \frac{\alpha\lambda^{\alpha}}{(\lambda + x)^{\alpha+1}}, x > 0$$

$$E(X) = \alpha\lambda^{\alpha}\left[\frac{t^{-\alpha+1}}{-\alpha+1}\right]_{\lambda}^{\infty} - \alpha\lambda^{\alpha+1}\left[\frac{t^{-\alpha}}{-\alpha}\right]_{\lambda}^{\infty} = \frac{\alpha\lambda}{\alpha-1} - \lambda = \frac{\lambda}{\alpha-1}$$

3.5 Loss Distributions

The two-parameter Pareto distribution



3.6 Loss Distributions

The Burr distribution

The CDF of the Pareto distribution $\mathrm{Pa}(\alpha, \lambda)$ is:

$$F(x) = 1 - \frac{\lambda^\alpha}{(\lambda + x)^\alpha}, x > 0$$

A further parameter $\boldsymbol{\gamma} > 0$ can be introduced by setting:

$$F(x) = 1 - \frac{\lambda^\alpha}{(\lambda + x^\gamma)^\alpha}, x > 0$$

This is the CDF of the transformed Pareto or Burr distribution. The additional parameter gives extra flexibility when a fit to data is required.

3.7 Loss Distributions

The three-parameter Pareto distribution

The PDF of the Pareto distribution $\mathrm{Pa}(\alpha, \lambda)$ is:

$$f(x) = \frac{\alpha\lambda^{\alpha}}{(\lambda + x)^{\alpha+1}}, x > 0$$

Another generalization of the Pareto distribution is to add a further parameter $\boldsymbol{k}$ so that the PDF becomes:

$$f(x) = \frac{\Gamma(\alpha + k)\lambda^{\alpha}}{\Gamma(\alpha)\Gamma(k)} \frac{x^{k-1}}{(\lambda + x)^{\alpha+k}},\ x > 0$$

The three-parameter Pareto distribution is equivalent to the two-parameter Pareto distribution when $\boldsymbol{k = 1}$.

3.8 Loss Distributions

The Weibull distribution

This distribution is called the Weibull distribution, a very flexible distribution, which can be used as a model for losses in insurance, usually with $\gamma < 1$. A random variable X has a Weibull distribution with parameters $c > 0$ and $\gamma > 0$ if it has CDF:

$$F(x) = 1 - \exp(-cx^{\gamma}), x > 0$$

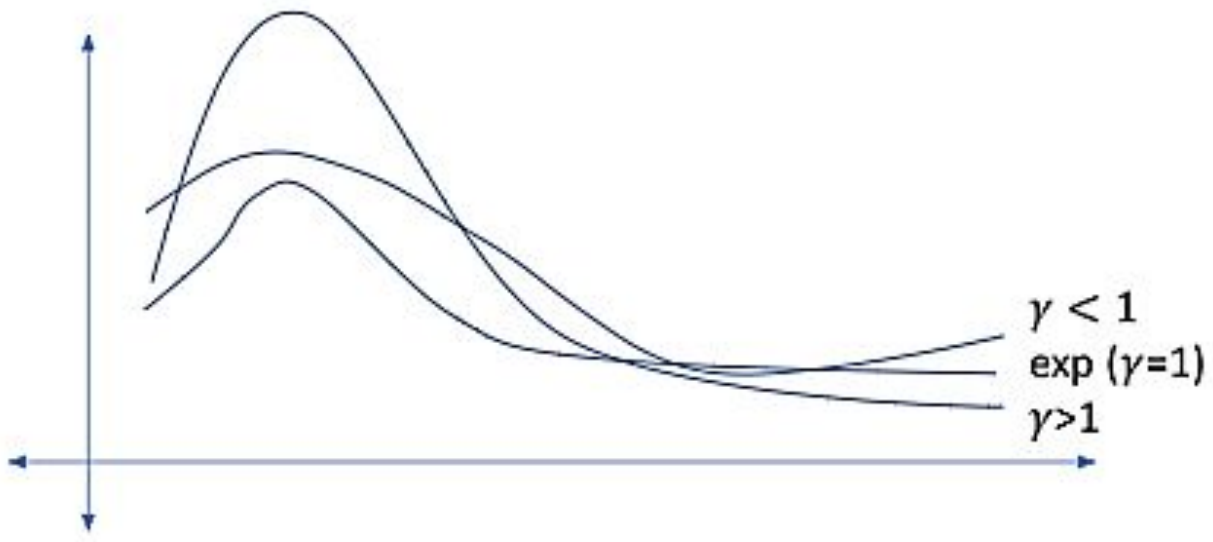
In that case we write $X \sim W(c, \gamma)$. (Note the change from λ to c; this is the notation used in the Tables for Actuarial Examinations).

The PDF of the $\boldsymbol{W(c, \gamma)}$ distribution is:

$$f(x) = c\gamma x^{\gamma-1}\exp(-cx^{\gamma}), x > 0$$

3.8 Loss Distributions

The Weibull distribution – Graph



4 Methods of Estimation

Practically, we will not have a ready distribution for claims.

We need to fit a distribution to the available data.

For this we will need to estimate the parameters of such distributions.

Methods of estimation

1. Method of moments
2. Maximum Likelihood Estimator
3. Method of Percentiles

4.1 Methods of Moments

For a distribution with r parameters, the moments are as follows:

$$\boldsymbol{m_j = \frac{1}{n}\sum_{i=1}^{n} x_i^j \; j = 1, 2 \ldots r}$$

where:
$m_j = E\left(X^j \mid \theta\right)$, a function of the unknown parameter, θ, being estimated
$\boldsymbol{n}$ = the sample size
x_i = the ith value in the sample

The estimate for the parameter, $\boldsymbol{\theta}$, can be determined by solving the equation above.

Where there is more than one parameter, they can be determined by solving the simultaneous equations for each $\boldsymbol{m_j}$.

Question

CT6 September 2012 Q4

Claims arising on a particular type of insurance policy are believed to follow a Pareto distribution. Data for the last several years shows the mean claim size is 170 and the standard deviation is 400.

(i) Fit a Pareto distribution to this data using the method of moments. [4]

(ii) Calculate the median claim using the fitted parameters and comment on the result. [3]

[Total 7]

Solution

(i) For the Pareto distribution with parameters α, λ as per the tables we have:

$$E(X) = \frac{\lambda}{\alpha - 1}$$

And

$$Var(X) = \frac{\alpha\lambda^2}{(\alpha-1)^2(\alpha-2)} = E(X)^2 \frac{\alpha}{\alpha-2}$$

And so

$$E\left(X^2\right) = Var(X) + E(X)^2 = E(X)^2\left(\frac{\alpha}{\alpha-2}+1\right) = E(X)^2\left(\frac{2\alpha-2}{\alpha-2}\right)$$

The observed values we are trying to fit are

$$E(X) = 170$$
$$E\left(X^2\right) = 400^2 + 170^2 = 434.626^2$$

So we have

$$\frac{2\alpha-2}{\alpha-2} = \frac{E(X^2)}{E(X)^2} = \frac{434.626^2}{170^2} = 6.53633$$

And so

$$\alpha = \frac{2 - 2\times 6.53633}{(2 - 6.53633)} = 2.441$$

And finally $\lambda = 1.441 \times 170 = 244.95$

Solution

(ii) We must solve

$$0.5 = 1 - \left(\frac{244.95}{244.95 + x}\right)^{2.441}$$

Re-arranging and taking roots gives

$$0.5^{\frac{1}{2.441}} = 0.7527965 = \frac{244.95}{244.95 + x}$$

And so

$$x = \frac{244.95 - 244.95 \times 0.7527965}{0.7527965} = 80.44$$

So the median is significantly lower than the mean. This demonstrates how skew the Pareto distribution is.

4.2 Maximum Likelihood Estimation

- Maximum likelihood estimation (MLE) is a technique used for estimating the parameters of a given distribution, using some observed data.

- Using a limited sample of the population, we find particular values of the mean and variance such that the observation is the most likely result to have occurred.

- For this we define a likelihood function.
- The likelihood function of a random variable, X, will give us the probability (or PDF) using a hypothetical parameter, θ.

- The maximum likelihood estimate (MLE) is that parameter which gives the highest probability (or PDF), i.e. that maximizes the likelihood function.

4.2 Maximum Likelihood Estimation

Step 1

the likelihood function $L(\theta)$ can be expressed as:

$$L(\theta) = \prod_{i=1}^{n} P(X = x_i \mid \theta) \text{ for a discrete random variable, } x$$

or:

$$L(\theta) = \prod_{i=1}^{n} f(x_i \mid \theta) \text{ for a continuous random variable, } x$$

To determine the MLE the likelihood function needs to be maximized.

4.2 Maximum Likelihood Estimation

Step 2
Often it is practical to consider the log-likelihood function:

$$I(\theta) = \log L(\theta) = \sum_{i=1}^{n} \log P(X = x_i \mid \theta) \text{ for a discrete random variable, } X$$

or:

$$I(\theta) = \log L(\theta) = \sum_{i=1}^{n} \log f(x_i \mid \theta) \text{ for a continuous random variable, } x$$

Step 3
If $I(\theta)$ can be differentiated with respect to θ, the MLE, expressed as $\hat{\theta}$, satisfies the expression:

$$\frac{d}{d\theta} I(\hat{\theta}) = 0$$

Where there is more than one parameter, the MLEs for each parameter can be determined by taking partial derivatives of the log-likelihood function and setting each to zero.

Question

CT6 September 2013 Q8

The number of claims per month Y arising on a certain portfolio of insurance policies is to be modelled using a modified geometric distribution with probability density given by

$$p(y|\alpha) = \frac{\alpha^{y-1}}{(1+\alpha)^y} \quad y = 1, 2, 3, \ldots$$

where α is an unknown positive parameter.

The most recent four months have resulted in claim numbers of 8, 6, 10 and 9.

(i) Derive the maximum likelihood estimate of α [5]

Solution

(i) We have 4 years of observations such that $y_1 + y_2 + y_3 + y_4 = 33$. The likelihood function is then:

$$L = \prod_{i=1}^{4} \frac{\alpha^{y_i - 1}}{(1+\alpha)^{y_i}} = \frac{\alpha^{33-4}}{(1+\alpha)^{33}} = \frac{\alpha^{29}}{(1+\alpha)^{33}}$$

The log-likelihood is then:

$$l = 29 \log \alpha - 33 \log(1+\alpha)$$

Taking its derivative w.r.t. α and equation it to zero we have:

$$\frac{29}{\alpha} - \frac{33}{1+\alpha} = 0$$

$$29(1+\alpha) = 33\alpha$$

which implies that $29 = 4\alpha$

therefore $\hat{\alpha} = \frac{29}{4} = 7.25$.

Differentiating the log likelihood again gives $-\frac{29}{\alpha^2} + \frac{33}{(1+\alpha)^2}$ which is negative at $\hat{\alpha} = 7.25$.

4.3 The method of percentiles

The method involves **equating selected sample percentiles to the distribution function**; for example, equate the sample quartiles, the 25th and 75th sample percentiles, to the population quartiles.
This corresponds to the way in which sample moments are equated to population moments in the method of moments. This method will be referred to as the method of percentiles.

In a similar fashion, when using the method of percentiles, the median would be used if there were one parameter to estimate.

Example: The distribution function of the $\boldsymbol{W}(\boldsymbol{c}, \boldsymbol{\gamma})$ distribution is an elementary function, and a simple method of estimation of both $\boldsymbol{c}$ and $\boldsymbol{\gamma}$ is based on this method.

Question

CT6 September 2010 Q3

An underwriter has suggested that losses on a certain class of policies follow a Weibull distribution. She estimates that the 10th percentile loss is 20 and the 90th percentile loss is 95.

(i) Calculate the parameters of the Weibull distribution that fit these percentiles. [3]

(i) Calculate the 99.5th percentile loss. [2] [Total 5]

Solution

(i) Let the parameters be c and γ as per the tables.

Then we have:

$$1-e^{-c\times 20^{\gamma}}=0.1 \text{ so } e^{-c\times 20^{\gamma}}=0.9 \text{ and so } c\times 20^{\gamma}=-\log 0.9 \text{ (A)}$$

And similarly $c\times 95^{\gamma}=-\log 0.1$ (B)

(A) divided by (B) gives $\left(\frac{20}{95}\right)^{\gamma}=\frac{\log 0.9}{\log 0.1}=0.0457575$

So $\gamma=\frac{\log 0.0457575}{\log\left(\frac{20}{95}\right)}=1.9795337$

And substituting into (A) we have $c=-\frac{\log 0.9}{20^{1.9795337}}=0.000280056$

Solution

(ii) The 99.5th percentile loss is given by

$$1-e^{-0.00280056x^{1.9795337}}=0.995$$

So that $-0.000280056x^{1.9795337}=\log 0.005$

$$\log x=\frac{\log\left(\dfrac{\log 0.005}{-0.000280056}\right)}{1.9795337}=4.97486366$$

So $x=e^{4.97486366}=144.73$

5 Goodness of fit tests

One way of testing whether a given loss distribution provides a good model for the observed claim amounts is to apply a chi-squared goodness-of-fit test.

Recall that the formula for the test statistic is $\sum \frac{(O-E)^2}{E}$, where:

- O is the observed number in a particular category
- E is the corresponding expected number predicted by the assumed probabilities
- the sum is over all possible categories.

Under the null hypothesis (that the model is correct), the test statistic has a chi-squared distribution.

A high value for the total indicates that the overall discrepancy is quite large and would lead us to reject the model.