Module 4

Short-Term Actuarial Mathematics

Errata

Chapter 1, page 12

The summary of the tail weights should read:

Tail probabilities

In order from lightest tail to heaviest tail, we have:

$W(\lambda,\gamma)$ with $\gamma > 1$

$Exp(\lambda)$ (which is the same as $W(\lambda,1)$)

$W(\lambda,\gamma)$ with $\gamma < 1$

$Pa(\alpha,\lambda)$

Chapter 1, page 21

In example 1.16 it states that the sample size is 65 when clearly it is 60. This changes the numbers throughout the solution.

Replacement pages are provided at the end of this document.

Chapter 3, page 61

There is a typo in the last line of example 3.6. For the formula to hold, we need $t < \dfrac{1}{800}$ and $t < \dfrac{1}{700}$, which means that it holds for $t < \dfrac{1}{800}$.

Chapter 3, page 63

In example 3.8 there is a typo in the solution for the expected value of Y. The second term should read:

$$1{,}000P(X > 1{,}000)$$

Chapter 5, page 112

The assumptions for the average cost per claim method should read:

- there are no further claims to come from the first accident year (*ie* the first accident year is fully run-off)
- the development pattern for the average cost per claim is the same for each accident year
- the development pattern for the number of claims is the same for each accident year
- there is no explicit allowance for inflation, and any inflation component is absorbed into the grossing-up factors.

Finally, substituting this value into (1) gives:

$$c = \frac{-\ln 0.75}{(350)^{\gamma}} = 0.003155$$ ♦♦

When using the method of percentiles to fit a single parameter distribution (such as the exponential distribution), we just equate the median of the distribution with the sample median.

1.4 Testing the goodness of fit of a loss distribution

Once we have estimated the parameter(s) of a loss distribution, we should check how well the distribution fits our sample data. In other words, we should check whether the assumed loss distribution has the correct sort of shape. We can do this by carrying out a chi-squared goodness-of-fit test.

Example 1.16

Claims from a particular group of policies are expected to follow an exponential distribution. A random sample of 60 claims has produced the following results:

Claim amount (£)	Observed frequency
$x \le 250$	24
$250 < x \le 500$	15
$500 < x \le 750$	9
$750 < x \le 1{,}000$	2
$1{,}000 < x \le 1{,}250$	2
$1{,}250 < x \le 1{,}500$	4
$x > 1{,}500$	4

The sample mean is £509.78.

Carry out a chi-squared test to assess the goodness of fit of the proposed exponential model.

Solution

We are testing:

H_0 : the claim amounts are exponentially distributed

against:

H_1 : the claim amounts are not exponentially distributed

The test statistic for this test is:

$$\sum \frac{(O-E)^2}{E}$$

In this formula, O denotes the observed frequency, E denotes the expected frequency, and the sum is taken over all claim size bands. To calculate the expected frequencies we must first estimate the value of the exponential parameter.

From Example 1.13 and the comment that follows it, we know that the exponential parameter is (usually) estimated by $\frac{1}{\bar{x}}$. So here we have:

$$\hat{\lambda} = \frac{1}{509.78}$$

Now, under the assumption that $X \sim Exp\left(\frac{1}{509.78}\right)$, we have:

$$P(X \leq 250) = 1 - e^{-250/509.78} = 0.38762$$

$$P(250 < X \leq 500) = e^{-250/509.78} - e^{-500/509.78} = 0.23737$$

$$P(500 < X \leq 750) = e^{-500/509.78} - e^{-750/509.78} = 0.14536$$

$$P(750 < X \leq 1{,}000) = e^{-750/509.78} - e^{-1{,}000/509.78} = 0.08902$$

$$P(1{,}000 < X \leq 1{,}250) = e^{-1{,}000/509.78} - e^{-1{,}250/509.78} = 0.05451$$

$$P(1{,}250 < X \leq 1{,}500) = e^{-1{,}250/509.78} - e^{-1{,}500/509.78} = 0.03338$$

$$P(X > 1{,}500) = e^{-1{,}500/509.78} = 0.05274$$

To calculate the expected frequencies, we multiply each of these probabilities by 60 (the sample size). For example, the expected number of claims that are less than or equal to £250 is:

$$60\,P(X \leq 250) = 23.257$$

The full set of expected frequencies is shown in the table below:

Claim amount (£)	Observed frequency	Expected frequency
$x \leq 250$	24	23.257
$250 < x \leq 500$	15	14.242
$500 < x \leq 750$	9	8.722
$750 < x \leq 1{,}000$	2	5.341
$1{,}000 < x \leq 1{,}250$	2	3.271
$1{,}250 < x \leq 1{,}500$	4	2.003
$x > 1{,}500$	4	3.164

Since the expected frequencies in the last 3 claim size bands are less than 5, we will combine these as shown below:

Claim amount (£)	Observed frequency	Expected frequency
$x \le 250$	24	23.257
$250 < x \le 500$	15	14.242
$500 < x \le 750$	9	8.722
$750 < x \le 1{,}000$	2	5.341
$x > 1{,}000$	10	8.438

The value of the test statistic (based on the 5 claim size bands given in the table above) is:

$$\sum \frac{(O-E)^2}{E} = \frac{(24-23.257)^2}{23.257} + \frac{(15-14.242)^2}{14.242} + \frac{(9-8.722)^2}{8.722} + \frac{(2-5.341)^2}{5.341} + \frac{(10-8.438)^2}{8.438}$$

$$= 2.452$$

We compare this against a χ^2 distribution. We have 5 claim size bands, but we have estimated the exponential parameter using the data, and we have the constraint that the expected frequencies must sum to 60. So we have $5-1-1=3$ degrees of freedom.

This is a one-tailed test. The upper 5% point of χ_3^2 is 7.815. As the value of the test statistic is less than 7.815, there is insufficient evidence to reject the null hypothesis at the 5% significance level. So we conclude that it is reasonable to model these claim amounts using an exponential distribution. ♦♦

1.5 Mixture distributions

Suppose that losses suffered by each policyholder are exponentially distributed, and the exponential parameter is different for each policyholder. This means that the loss random variable, X, depends on the value of the exponential parameter.

Now let's model the exponential parameter as a random variable. Recall that we usually use capital letters to denote random variables. In keeping with this convention, we will use Λ to denote the exponential parameter. In particular, let's assume that $\Lambda \sim Gamma(\alpha, \delta)$, so that the PDF of Λ is:

$$f_\Lambda(\lambda) = \frac{\delta^\alpha}{\Gamma(\alpha)} \lambda^{\alpha-1} e^{-\delta\lambda}, \quad \text{for } \lambda > 0$$

We may be interested in the unconditional distribution of X. This is also known as the *mixture distribution* or *marginal distribution* of X. We can determine the unconditional PDF of X by integrating the joint PDF of X and Λ over all possible values of Λ, *ie*:

$$f_X(x) = \int_\lambda f_{X,\Lambda}(x,\lambda)\, d\lambda$$