Non-Life Insurance: Mathematics and Statistics

Solution sheet 6

Solution 6.1 Goodness-of-Fit Test

Let \( Y \) be a random variable following a Pareto distribution with threshold \( \theta = 200 \) and tail index \( \alpha = 1.25 \). Then the distribution function \( G \) of \( Y \) is given by

\[
G(x) = 1 - \left( \frac{x}{\theta} \right)^{-\alpha} = 1 - \left( \frac{x}{200} \right)^{-1.25}
\]

for all \( x \geq \theta \). For example for the interval \( I_2 \) we then have

\[
\mathbb{P}[Y \in I_2] = \mathbb{P}[239 \leq Y < 301] = G(301) - G(239) = 1 - \left( \frac{301}{200} \right)^{-1.25} - \left[ 1 - \left( \frac{239}{200} \right)^{-1.25} \right] \approx 0.2.
\]

By analogous calculations for the other four intervals, we get

\[
\mathbb{P}[Y \in I_1] \approx 0.2, \quad \mathbb{P}[Y \in I_2] \approx 0.2, \quad \mathbb{P}[Y \in I_3] \approx 0.2, \quad \mathbb{P}[Y \in I_4] \approx 0.2, \quad \mathbb{P}[Y \in I_5] \approx 0.2.
\]

Let \( E_i \) and \( O_i \) denote respectively the expected number of observations in \( I_i \) and the observed number of observations in \( I_i \), for all \( i \in \{1, \ldots, 5\} \). As we have 20 observations in our data, we can calculate for example \( E_2 \) as

\[
E_2 = 20 \cdot \mathbb{P}[Y \in I_2] \approx 4.
\]

The values of the expected number of observations and the observed number of observations in the five intervals as well as their squared differences are summarized in the following table:

<table>
<thead>
<tr>
<th>( i )</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>( O_i )</td>
<td>4</td>
<td>0</td>
<td>8</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>( E_i )</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>((O_i - E_i)^2)</td>
<td>0</td>
<td>16</td>
<td>16</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

Now the test statistic of the \( \chi^2 \)-goodness-of-fit test using 5 intervals and 20 observations is given by

\[
X_{20,5}^2 = \sum_{i=1}^{5} \frac{(O_i - E_i)^2}{E_i} = \frac{0}{4} + \frac{16}{4} + \frac{16}{4} + \frac{4}{4} + \frac{4}{4} = 10.
\]

Let \( \alpha = 5\% \). Then the \((1 - \alpha)\)-quantile of the \( \chi^2 \)-distribution with \( 5 - 1 = 4 \) degrees of freedom is given by approximately 9.49. Since this is smaller than \( X_{20,5}^2 \), we can reject the null hypothesis of having a Pareto distribution with threshold \( \theta = 200 \) and tail index \( \alpha = 1.25 \) as claim size distribution at the significance level of 5\%.

Solution 6.2 Log-Normal Distribution and Deductible

(a) Let \( X \sim \mathcal{N}(\mu, \sigma^2) \). Then the moment generating function \( M_X \) of \( X \) is given by

\[
M_X(r) = \mathbb{E}[\exp\{rX\}] = \exp \left\{ r\mu + \frac{r^2\sigma^2}{2} \right\}
\]

Updated: October 24, 2017
for all \( r \in \mathbb{R} \). Since \( Y_1 \) has a log-normal distribution with mean parameter \( \mu \) and variance parameter \( \sigma^2 \), we have
\[
Y_1 \overset{d}{=} \exp(X).
\]

Hence, the expectation, the variance and the coefficient of variation of \( \sigma \) for all
\[
E[\sigma] = \exp\left(\mu + \frac{\sigma^2}{2}\right),
\]
\[
\text{Var}(\sigma) = \exp\left(\mu + \frac{\sigma^2}{2}\right) - \exp\left(\mu + \frac{\sigma^2}{2}\right)^2 = \exp\left(2\mu + \frac{\sigma^2}{2}\right) - \exp\left(2\mu + \sigma^2\right) \exp\{\sigma^2\} - 1 \quad \text{and}
\]
\[
\text{Vco}(\sigma) = \frac{\sqrt{\text{Var}(\sigma)}}{E[\sigma]} = \frac{\exp\left(\mu + \sigma^2/2\right) \sqrt{\exp\{\sigma^2\} - 1}}{\exp\left(\mu + \sigma^2/2\right)} = \sqrt{\exp\{\sigma^2\} - 1}.
\]

(b) From part (a), we know that

\[
\text{E}[Y_1] = \text{E}[\exp(X)] = \text{E}[\exp\{1 \cdot X\}] = M_X(1) = \exp\left(\mu + \frac{\sigma^2}{2}\right),
\]
\[
\text{Var}(Y_1) = \text{E}[Y_1^2] - \text{E}[Y_1]^2 = \text{E}[\exp\{2X\}] - M_X(1)^2 = M_X(2) - M_X(1)^2
\]
\[
= \exp\left(2\mu + \frac{4\sigma^2}{2}\right) - \exp\left(2\mu + \frac{2\sigma^2}{2}\right) = \exp\left(2\mu + \sigma^2\right) \exp\{\sigma^2\} - 1
\]
\[
\text{Vco}(Y_1) = \sqrt{\frac{\text{Var}(Y_1)}{\text{E}[Y_1]}} = \frac{\exp\left(\mu + \frac{\sigma^2}{2}\right) \sqrt{\exp\{\sigma^2\} - 1}}{\exp\left(\mu + \frac{\sigma^2}{2}\right)} = \sqrt{\exp\{\sigma^2\} - 1}.
\]

Note that the introduction of this deductible reduces the administrative burden a lot, because 41% of (small) claims disappear.

\[
\text{(i) The claims frequency } \lambda \text{ is given by } \lambda = \text{E}[N]/v. \text{ With the introduction of the deductible } d = 500, \text{ the number of claims changes to}
\]
\[
N_{\text{new}} = \sum_{i=1}^{N} 1_{\{Y_i > d\}}.
\]

Using the independence of \( N \) and \( Y_1, Y_2, \ldots \), we get
\[
\text{E}[N_{\text{new}}] = \text{E}\left[\sum_{i=1}^{N} 1_{\{Y_i > d\}}\right] = \text{E}[N]\text{E}[1_{\{Y_1 > d\}}] = \text{E}[N]\text{P}[Y_1 > d].
\]

Let \( \Phi \) denote the distribution function of a standard Gaussian distribution. Since \( \log Y_1 \) has a Gaussian distribution with mean \( \mu \) and variance \( \sigma^2 \), we have
\[
\text{P}[Y_1 > d] = 1 - \Phi\left(\frac{\log Y_1 - \mu}{\sigma}\right) = 1 - \Phi\left(\frac{\log d - \mu}{\sigma}\right).
\]

Hence, the new claims frequency \( \lambda_{\text{new}} \) is given by
\[
\lambda_{\text{new}} = \frac{\text{E}[N_{\text{new}}]}{v} = \frac{\text{E}[N]\text{P}[Y_1 > d]}{v} = \lambda\text{P}[Y_1 > d] = \lambda\left[1 - \Phi\left(\frac{\log d - \mu}{\sigma}\right)\right].
\]

Inserting the values of \( d, \mu \) and \( \sigma \), we get
\[
\lambda_{\text{new}} \approx \lambda\left[1 - \Phi\left(\frac{\log 500 - 6.59}{1.68}\right)\right] \approx 0.59 \cdot \lambda.
\]
(ii) With the introduction of the deductible $d = 500$, the claim sizes change to
\[ Y_i^{\text{new}} = Y_i - d \mid Y_i > d. \]

Thus, the new expected claim size is given by
\[ E[Y_i^{\text{new}}] = E[Y_i - d \mid Y_i > d] = e(d), \]
where $e(d)$ is the mean excess function of $Y_i$ above $d$. According to the lecture notes, $e(d)$ is given by
\[ e(d) = E[Y_1] \left[ 1 - \Phi \left( \frac{\log d - \mu - \sigma^2}{\sigma} \right) \right] \]
\[ \cdot \left[ 1 - \Phi \left( \frac{\log d - \mu}{\sigma} \right) \right] - d. \]

Inserting the values of $d, \mu, \sigma$ and $E[Y_1]$, we get
\[ E[Y_i^{\text{new}}] \approx 3'000 \left[ 1 - \Phi \left( \frac{\log 500 - 6.59 - 1.68^2}{1.68} \right) \right] - 500 \approx 4'456 \approx 1.49 \cdot E[Y_1]. \]

(iii) According to Proposition 2.2 of the lecture notes, the expected total claim amount $E[S]$ is given by
\[ E[S] = E[N]E[Y_1]. \]

With the introduction of the deductible $d = 500$, the total claim amount $S$ changes to $S^{\text{new}}$, which can be written as
\[ S^{\text{new}} = \sum_{i=1}^{N^{\text{new}}} Y_i^{\text{new}}. \]

Hence, the expected total claim amount changes to
\[ E[S^{\text{new}}] = E[N^{\text{new}}]E[Y_1^{\text{new}}] \]
\[ = E[N]P[Y_1 > d]e(d) \]
\[ = \lambda v \left[ 1 - \Phi \left( \frac{\log d - \mu}{\sigma} \right) \right] \cdot \left( E[Y_1] \left[ 1 - \Phi \left( \frac{\log d - \mu - \sigma^2}{\sigma} \right) \right] \right). \]

Inserting the values of $d, \mu, \sigma$ and $E[Y_1]$, we get
\[ E[S^{\text{new}}] \approx \lambda v \left[ 1 - \Phi \left( \frac{\log 500 - 6.59}{1.68} \right) \right] \cdot \left( 3'000 \left[ 1 - \Phi \left( \frac{\log 500 - 6.59 - 1.68^2}{1.68} \right) \right] - 500 \right) \]
\[ \approx 0.88 \cdot 4'456 \approx 1.49 \cdot E[S]. \]

In particular, the insurance company can grant a discount of roughly 12% on the pure risk premium. Note that also the administrative expenses on claims handling will reduce substantially because we only have 59% of the original claims, see the result in (i).
Solution 6.3 Inflation and Deductible

Let $Y$ be a random variable following a Pareto distribution with threshold $\theta > 0$ and tail index $\alpha > 1$. Then the expectation $E[Y]$ of $Y$ and the mean excess function $e_Y(u)$ of $Y$ above $u > \theta$ are given by

$$E[Y] = \frac{\alpha}{\alpha - 1} \theta \quad \text{and} \quad e_Y(u) = \frac{1}{\alpha - 1} u.$$

Since the insurance company only has to pay the part that exceeds the threshold $\theta$, this year’s average claim payment $z$ is

$$z = E[Y] - \theta = \frac{\alpha}{\alpha - 1} \theta - \theta = \frac{\theta}{\alpha - 1}.$$

For the total claim size $\tilde{Y}$ of a claim next year we have

$$\tilde{Y} \stackrel{d}{=} (1 + r)Y \sim \text{Pareto}([1 + r]\theta, \alpha).$$

Let $\rho \theta$ for some $\rho > 0$ denote the increase of the threshold that is needed such that the average claims payment remains unchanged. Then next year’s average claim payment is given by

$$\tilde{z} = E[(\tilde{Y} - [1 + \rho]\theta_+)].$$

Let’s first assume that we can choose a $\rho < r$ such that $z = \tilde{z}$. In this case we get

$$\tilde{Y} \geq (1 + r)\theta \quad \text{a.s.} \quad \implies \quad \tilde{Y} \geq (1 + \rho)\theta \quad \text{a.s.}$$

and thus

$$\tilde{z} = E[\tilde{Y} - (1 + \rho)\theta] = E[\tilde{Y}] - (1 + \rho)\theta = \frac{\alpha}{\alpha - 1} (1 + r)\theta - (1 + \rho)\theta.$$

Now we have $z = \tilde{z}$ if and only if

$$\frac{\alpha}{\alpha - 1} \theta - \theta = \frac{\alpha}{\alpha - 1} (1 + r)\theta - (1 + \rho)\theta \iff 0 = \frac{\alpha}{\alpha - 1} r\theta - \rho \theta \iff \rho = \frac{\alpha}{\alpha - 1} r > r,$$

which is a contradiction to the assumption $\rho < r$. Hence, we conclude that $\rho \geq r$, i.e. the percentage increase in the deductible has to be bigger than the inflation. Assuming $\rho \geq r$, we can calculate

$$\tilde{z} = E[(\tilde{Y} - [1 + \rho]\theta) \cdot 1_{\{\tilde{Y} > (1 + \rho)\theta\}}] = E[\tilde{Y} - (1 + \rho)\theta | \tilde{Y} > (1 + \rho)\theta] \cdot P[\tilde{Y} > (1 + \rho)\theta] = e_\tilde{Y}((1 + \rho)\theta) \cdot P[\tilde{Y} > (1 + \rho)\theta] = \frac{1}{\alpha - 1} (1 + \rho)\theta \cdot \left(\frac{(1 + \rho)\theta}{(1 + r)\theta}\right)^{-\alpha} = \frac{\theta}{\alpha - 1} (1 + r)^{\alpha - 1} (1 + \rho)^{-\alpha + 1} = z \cdot (1 + r)^{\alpha - 1} (1 + \rho)^{-\alpha + 1}.$$

Now we have $z = \tilde{z}$ if and only if

$$(1 + r)^{\alpha} (1 + \rho)^{-\alpha + 1} = 1 \iff \rho = (1 + r)^{\frac{\alpha}{\alpha - 1}} - 1.$$

We conclude that if we want the claim payment to remain unchanged, we have to increase the deductible $\theta$ by the amount

$$\theta \left( (1 + r)^{\frac{\alpha}{\alpha - 1}} - 1 \right).$$