Queen Mary University of London

**Main Examination period 2022 – May/June – Semester B**

# MTH5126: Statistics for Insurance

**You should attempt ALL questions. Marks available are shown next to the questions.**

**In completing this assessment:**

- **You may use books and notes.**
- **You may use calculators and computers, but you must show your working for any calculations you do.**
- **You may use the Internet as a resource, but not to ask for the solution to an exam question or to copy any solution you find.**
- **You must not seek or obtain help from anyone else.**

All work should be **handwritten** and should **include your student number**.

The exam is available for a period of **24 hours**. Upon accessing the exam, you will have **3 hours** in which to complete and submit this assessment.

When you have finished:

- scan your work, convert it to a **single PDF file**, and submit this file using the tool below the link to the exam;
- e-mail a copy to **maths@qmul.ac.uk** with your student number and the module code in the subject line;
- with your e-mail, include a photograph of the first page of your work together with either yourself or your student ID card.

Please try to upload your work well before the end of the submission window, in case you experience computer problems. **Only one attempt is allowed – once you have submitted your work, it is final**.

**IFoA exemptions.** For actuarial students, this module counts towards IFoA actuarial exemptions. To be eligible for IFoA exemption, **your must submit your exam within the first 3 hours of the assessment period**.

**Examiners: F. Parsa, J. Griffin**

 

**Question 1 [31 marks].** A home insurance company's total monthly claim amounts have a mean of 250 and a standard deviation of 300. The company has estimated that it will face insolvency if the total monthly claim amounts exceed 1,000 in any given month.

(a) Determine the probability that the company faces insolvency in any given month if the company assumes that total monthly claim amounts follow the Normal distribution. [**4**]

(b) Determine the revised value of the probability in part (a) if the company assumes that total monthly claim amounts follow the two-parameter Pareto distribution. [**6**]

(c) An Analyst has determined that the two-parameter Pareto distribution is the best fit for the distribution of the total monthly claim amounts for this company. Outline, using the results from parts (a) and (b), the potential consequences of the company assuming that the total monthly claim amounts follow the Normal distribution rather than the two-parameter Pareto distribution.
Explain why the Normal distribution is unlikely to be a good fit for the distribution of the total monthly claim amounts for this company. [**4**]

Suppose that the aggregate claims from a risk have a compound Poisson distribution with parameter $\mu$, and individual claim amounts have a two-parameter Pareto distribution with a mean of 250 and a standard deviation of 300. The insurer of this risk is considering effecting proportional reinsurance with a retention level of 0.75, and calculates the premium using a premium loading factor of 0.1 (this means they charge 10% in excess of the risk premium). The reinsurance premium would be calculated using a premium loading factor of 0.2.
(The insurer's profit is defined to be the premium charged by the insurer less the reinsurance premium and less the claims paid by the insurer, net of reinsurance.)

(d) Calculate the insurer's expected profit before effecting the reinsurance. [**4**]

(e) Calculate the insurer's expected profit after effecting the reinsurance and hence find the percentage reduction in the insurer's expected profit. [**13**]

**Solution**

(a)

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

$$\begin{aligned} P(X > 1000) &= 1 - P(X \leq 1000) \\ &= 1 - P(\frac{X-\mu}{\sigma} \leq \frac{1000-250}{300}) \\ &= 1 - P(Z \leq 2.5) \\ &= 1 - 0.99379 \\ &= 0.00621 \end{aligned}$$

$$(\text{or} \quad 0.621\%) \qquad [\mathbf{4}]$$

(b)

$$X \sim Pareto(\lambda, \alpha)$$

$$E(X) = \frac{\lambda}{\alpha - 1} = 250$$

$$Var(X) = \frac{\alpha\lambda^2}{(\alpha-1)^2(\alpha-2)} = 300^2$$

$$\frac{\alpha}{\alpha-2}250^2 = 300^2$$

$$\alpha = 6.55$$

$$\lambda = 250 \times (\alpha - 1) = 250(6.55 - 1) = 1,386.36 \qquad [\mathbf{4}]$$

$$P(\text{Claims} > 1000) = (\frac{\lambda}{\lambda + 1000})^{\alpha} = 0.0286$$

$$(\text{or} \quad 2.86\%) \qquad [\mathbf{2}]$$

(c) The probability of insolvency could be underestimated which could lead to the insurance company holding insufficient capital (or taking out insufficient reinsurance)
The Normal distribution is unlikely to be a good fit for the total monthly claim amounts because negative claims can't be incurred by the company and the Normal distribution assigns a non-zero probability to negative claims occurring. In particular, with the given mean and standard deviation, there is a significant probability (around 20%) of claims being negative.

The Normal distribution is also unlikely to be a good fit because the distribution of claims incurred by the company is likely to be positively skewed and the Normal distribution is symmetric and has zero skewness.
Additionally, the Normal distribution is thin-tailed and therefore, not suitable for modelling situations where extreme events occur reasonably frequently which would be expected to be the case for home insurance [**4** from 9 available marks]

(d) Aggregate claims before reinsurance, $S$, follows compound Poisson distribution, with parameter $\mu$. If $N$ is the number of claims and the individual claim amounts $X \sim Pareto(\alpha, \lambda)$, we can write

$$E(S) = E(N)E(X) = 250\mu.$$

The insurers expected profit without reinsurance is equal to premiums minus expected claims. But if a loading factor of 0.1 is in use, then the total premium is $1.1 \times 250\mu = 275\mu$. So,

$$\text{the expected profit} = 275\mu - 250\mu = 25\mu. \qquad [\mathbf{4}]$$

(e) For the reinsurer, total aggregate claims are given by

$$S_R = Z_1 + Z_2 + \quad + Z_N$$

where $Z_i = (1 - 0.75)X_i$ and it follows a compound Poisson distribution. So

$$E(S_R) = E(N)E(Z) = \mu E(Z) \qquad [\mathbf{3}]$$

but

$$E(Z) = \int_0^\infty (1 - 0.75)x \frac{\alpha\lambda^\alpha}{(\lambda + x)^4} dx$$

$$= 0.25 \int_0^\infty x \frac{\alpha\lambda^\alpha}{(\lambda + x)^4} dx$$

$$= 0.25E(X)$$

$$= 0.25 \times 250$$

$$= 62.5 \qquad [\mathbf{4}]$$

so

$$E(S_R) = 62.5\mu \qquad [\mathbf{1}]$$

The reinsurance premium is given by $1.2E(S_R)$, and

$$1.2E(S_R) = 1.2 \times 62.5\mu = 75\mu$$

The expected profit with reinsurance is

$E$(Gross premium-reinsurance premium-net claim amounts) $= 275\mu - 75\mu - E(S - S_R)$

where net claim amounts, $S - S_R$ , is what the insurer pays net of reinsurance.
We have

$$E(S - S_R) = E(S) - E(S_R) = 250\mu - 62.5\mu = 187.5\mu$$

So, the expected profit is $12.5\mu$. [**4**]

And the percentage reduction in the expected profit (which was $25\mu$ without reinsurance) is 50% [**1**].

**Question 2 [27 marks].** Claims on a portfolio of insurance policies arrive as a Poisson process with parameter 200. Individual claim amounts follow a normal distribution with the mean 20 and variance 16. The insurer calculates premiums using a premium loading of 15% and has initial surplus of 200.

(a) Define the ruin probabilities $\Psi(200)$,$\Psi(200, 1)$ and $\Psi_1(200, 1)$. [**7**]

(b) Define the adjustment coefficient R. [**3**]

(c) Show that for this portfolio the value of R is 0.013 correct to 3 decimal places. [**7**]

(d) Calculate an upper bound for $\Psi(200)$ and an estimate of $\Psi_1(200, 1)$. [**8**]

(e) Comment on the results in part (d). [**2**]

**Solution**

(a) Let $S(t)$ denote the claims to time $t$. Let the annual rate of premium income be $c$ and let the insurer's initial surplus $U$ be $U = 200$. Then the surplus at time $t$ is given by:

$$U(t) = U + ct - S(t). \qquad [\mathbf{4}]$$

And the relevant probabilities are defined by:

$$\Psi(200) = P(U(t) < 0 \quad \text{for some} \quad t > 0)$$

$$\Psi(200, 1) = P(U(t) < 0 \quad \text{for some} \quad t, \quad 0 < t \leq 1)$$

$$\Psi_1(200, 1) = P(U(1) < 0) \qquad [\mathbf{3}]$$

(b) The adjustment coefficient is the unique positive root of the equation:

$$\lambda M_X(R) = \lambda + cR$$

Where $\lambda$ is the rate of the Poisson process(i.e. 200) and $X$ is the normal distribution with mean 20 and standard deviation 4.[**3**]

(c) In this case we have:

$$c = 200 \times 20 \times 1.15 = 4,600$$

And

$$M_X(R) = \exp(20R + 8R^2)$$

So $R$ is the root of:

$$200 \exp(20R + 8R^2) - 200 - 4,600R = 0 \qquad [\mathbf{3}]$$

Denote the left hand side of the equation by $f(R)$ When $R = 0.0135$ we have:

$$f(0.0135) = 200\exp(0.271458) - 200 - 62.1 = 0.2751544 > 0$$

And when $R = 0.0125$ we have:

$$f(0.0125) = 200\exp(0.25125) - 200 - 57.5 = -0.3737096 < 0 \qquad [\mathbf{2}]$$

Since the function changes sign between 0.0135 and 0.0125 the unique positive root must lie between these values and hence the root is 0.013 correct to 3 decimal places. [**2**]

(d) By Lundberg's inequality:

$$\Psi(200) < \exp(-200 \times 0.013) = 0.074273578 \qquad [\mathbf{2}]$$

Claims in the first year are approximately Normal with:

mean $200 \times 20 = 4000$
variance $200 \times (16 + 20^2) = 83200$ [**2**]

So approximately:

$$\Psi_1(200, 1) = P(200 + 4600 - N(4000, 83200) < 0)$$

$$= P(N(4000, 83200) > 4800)$$

$$= P(N(0,1) > \frac{4800 - 4000}{\sqrt{83200}})$$

$$= P(N(0,1) > 2.774)$$

$$= 1 - (0.99728 \times 0.4 + 0.99720 \times 0.6)$$

$$= 0.0028 \qquad [\mathbf{4}]$$

(e) The probability of ruin is much smaller in the first year than in the long-term bound provided by Lundberg's inequality.
This suggests that either the bound in Lundberg's inequality may not be that tight
or that there is significant probability of ruin at times greater than 1 year. [**2**]

**Question 3 [22 marks].**

(a) Derive that the coefficient of lower tail dependence for the Clayton copula when the parameter $\alpha > 0$, is $2^{-\frac{1}{\alpha}}$. [**8**]

(b) Comment on how the value of the parameter $\alpha$ affects the degree of upper tail dependence in the case of the Clayton copula. [**3**]

(c) Derive an expression for the Clayton copula for the case where the parameter $\alpha > 0$ and there are 3 variables. The Clayton copula has a generator function:

$$\Psi(t) = \frac{1}{\alpha}(t^{-\alpha} - 1).$$

[**11**]

**Solution**

(a) Coefficient of lower tail dependence in terms of the copula function is

$$\lambda_L = \lim_{u \to 0^+} \frac{C[u,u]}{u}$$

for the Clayton copula:

$$C[u,v] = (u^{-\alpha} + v^{-\alpha} - 1)^{-\frac{1}{\alpha}}$$

for $\alpha > 0$. Setting $u = v$, we have

$$C[u,u] = (2u^{-\alpha} - 1)^{-\frac{1}{\alpha}} \qquad [\mathbf{4}]$$

so

$$\begin{aligned}\lambda_L &= \lim_{u \to 0^+} \frac{(2u^{-\alpha} - 1)^{-\frac{1}{\alpha}}}{u} \\ &= \lim_{u \to 0^+} (2 - \frac{1}{u^{-\alpha}})^{-\frac{1}{\alpha}} \\ &= 2^{-\frac{1}{\alpha}} \qquad [\mathbf{4}]\end{aligned}$$

(b) One can see that as $\alpha$ increases, $2^{-\frac{1}{\alpha}}$ increases. So increasing the value of the parameter $\alpha$ increases the degree of lower tail dependence of the Clayton copula.[**3**]

(c) The Clayton copula is an example of Archimedean copulas. For the case where there are 3 variables, Archimedean copulas take the form:

$$C[u,v,w] = \Psi^{[-1]}(\Psi(u) + \Psi(v) + \Psi(w)) \qquad [\mathbf{2}]$$

So, we need to find $\Psi^{[-1]}$, the pseudo-inverse generator function. Check

$$\Psi(0) = \lim_{t \to 0} \frac{1}{\alpha}(t^{-\alpha} - 1) = \infty$$

So, the pseudo-inverse $\Psi^{[-1]}$ is equal to the ordinary inverse $\Psi^{(-1)}$ [**3**].

Let $y = \Psi^{(-1)}(x)$, then $\Psi(y) = x$. So

$$\frac{1}{\alpha}(y^{-\alpha} - 1) = x$$

$$y = (\alpha x + 1)^{-\frac{1}{\alpha}}$$

$$\Psi^{[-1]}(x) = (\alpha x + 1)^{-\frac{1}{\alpha}} \qquad [\mathbf{3}]$$

Therefore

$$C[u,v,w] = \Psi^{[-1]}(\Psi(u) + \Psi(v) + \Psi(w))$$

$$= (\alpha(\frac{1}{\alpha}(u^{-\alpha} - 1) + \frac{1}{\alpha}(v^{-\alpha} - 1) + \frac{1}{\alpha}(w^{-\alpha} - 1)) + 1)^{-\frac{1}{\alpha}}$$

$$= (u^{-\alpha} + v^{-\alpha} + w^{-\alpha} - 2)^{-\frac{1}{\alpha}} \qquad [\mathbf{3}]$$


**Continue to next page**

**Question 4 [20 marks].** An actuary assumes that the underlying gross claims follow an exponential distribution of some unknown rate $\lambda$. The payments of $n$ claims are to be split between an insurance company and its reinsurer under an Excess of Loss reinsurance arrangement with a retention level M. The actuary needs to find the maximum likelihood estimate of $\lambda$ using only the claims amount paid by the insurer. If the amount of only $r$ claims are above the retention level, find the maximum likelihood estimate of $\lambda$. Assume that all claims are independent. [**20**]

**Solution**

We have:

$$X_i \sim \exp(\lambda), i = 1, 2, ..., n$$

$$f(x_i|\lambda) = \lambda e^{-\lambda x_i}$$

$$F(x_i|\lambda) = \int_0^{x_i} \lambda e^{-\lambda u} du = [-e^{-\lambda u}]_0^{x_i} = 1 - e^{-\lambda x_i}$$

There are $n - r$ claims below the retention level and they will be paid by insurer, and the insurer will pay only amount $M$ for other $r$ claims which are above the retention level. So our data is censored. Therefore, our log-likelihood function has 2 parts. [**3**]

$$\begin{aligned} L_1(\lambda) &= \prod_{i=1}^{n-r} f(x_i|\lambda) \\ &= \prod_{i=1}^{n-r} \lambda e^{-\lambda x_i} \\ &= \lambda^{n-r} e^{-\lambda \sum_{i=1}^{n-r} x_i} \qquad [\mathbf{3}] \end{aligned}$$

for all the claims which are less than the retention level $M$, and

$$\begin{aligned} L_2(\lambda) &= \prod_{i=1}^{r} P(y_i > M) \\ &= \prod_{i=1}^{r} 1 - F(M|\lambda) \\ &= \prod_{i=1}^{r} 1 - (1 - e^{-\lambda M}) \\ &= e^{-\lambda(rM)} \qquad [\mathbf{3}] \end{aligned}$$

for claims which are greater than $M$.

Therefore, the likelihood function of $\lambda$ is:

$$L(\lambda) = L_1(\lambda) \times L_2(\lambda)$$

$$= (\lambda^{n-r} e^{-\lambda \sum_{i=1}^{n-r} x_i}) e^{-\lambda(rM)}$$

$$= \lambda^{n-r} e^{-\lambda(\sum_{i=1}^{n-r} x_i + rM)} \qquad [\mathbf{2}]$$

Now we need to take the natural log to get the log-likelihood function of $\lambda$:

$$l(\lambda) = \log[L(\lambda)]$$

$$= \log[\lambda^{n-r} e^{-\lambda(\sum_{i=1}^{n-r} x_i + rM)}]$$

$$= (n-r)\log\lambda - \lambda(\sum_{i=1}^{n-r} x_i + rM). \qquad [\mathbf{3}]$$

Diffrentiating, we get

$$\frac{d}{d\lambda} l(\lambda) = \frac{n-r}{\lambda} - (\sum_{i=1}^{n-r} x_i + rM) \qquad [\mathbf{2}]$$

and equating to zero:

$$\hat{\lambda} = \frac{n-r}{\sum_{i=1}^{n-r} x_i + rM}. \qquad [\mathbf{2}]$$

Since

$$\frac{d^2}{d\lambda} l(\lambda) = -\frac{n-r}{\lambda^2} < 0,$$

then $\hat{\lambda}$ maximises the function. [**2**]

---

**End of Paper.**