ETC3420-5342 Tutorial Week 9 - Solutions

Part I: Building Blocks

Question 1: (a) Explain what is meant by an extreme event and give two examples in an insurance context.

Solution: Extreme events are outcomes that have a very low probability of occurrence but involve very large sums of money.

In an insurance context, they may arise as a result of a single cause that has a high financial cost (eg a bodily injury claim or complete destruction of a building) or as an accumulation of events with a related cause (e.g., flood damage to a large number of houses in one town).

(b) Explain why it is important to model extreme events separately from other events.

Solution: The majority of risk events fall within the main body of the fitted distribution and can usually be modelled reasonably accurately by one of the standard statistical distributions.

However, there is usually a lack of past data on extreme events. If a distribution is fitted to the whole dataset, the parameter estimates will reflect where the bulk of the data values lie rather than the extreme events. This might mean the fitted distribution understates the probability of future extreme events.

Therefore, a different approach to modelling extreme events is taken, e.g., by considering the distribution of block maxima or the distribution of threshold exceedances.

Question 2: (a) Describe the generalised extreme value (GEV) distribution.

Solution: The maximum value, X_M, in asample of n IID random variables $X_1, ..., X_n$, tends to a particular distribution as the sample size increases. This is called the generalised extreme value (GEV) distribution.

The GEV distribution has CDF:

$$H(x) = \begin{cases} \exp\left(-\left(1+\frac{\gamma(x-\alpha)}{\beta}\right)^{-\frac{1}{\gamma}}\right) & \gamma \neq 0 \\ \exp\left(-\exp\left(-\frac{x-\alpha}{\beta}\right)\right) & \gamma = 0 \end{cases}$$

The key parameter is the shape parameter γ.

- When $\gamma > 0$, we have the Frchet-type GEV distribution. This is the limiting form for heavy tail underlying distributions with a finite lower bound, such as the Pareto distribution.
- When $\gamma < 0$, we have the Weibull-type GEV distribution. This is the limiting form for underlying distributions with a finite upper bound, such as the uniform distribution.
- When $\gamma = 0$, we have the Gumbel-type GEV distribution. This is the limiting form for most other underlying distributions that have finite moments, such as the normal and lognormal distributions.

The parameters α, β are the location and scale parameters, respectively. These will differ depending on the underlying distribution.

(b) Outline an alternative approach that can be used in place of the GEV distribution to model extreme events.

Solution: As an alternative to focusing upon a single maximum value, we can consider the distribution of all the claim values that exceed some threshold u. The distribution of $X - u|X > u$ is called the threshold exceedances distribution.

A similar theory to GEV predicts that the limiting distribution as $u \to \infty$, is a generalised Pareto distribution (GPD). The GPD has CDF:

$$G(x) = \begin{cases} 1 - \left(1 + \gamma \frac{x}{\beta}\right)^{-\frac{1}{\gamma}} & \gamma \neq 0 \\ 1 - \exp\left(-\frac{x}{\beta}\right) & \gamma = 0 \end{cases}$$

The parameters β and γ are the scale and shape parameters, respectively.

In order to fit the tail of a distribution we need to select a suitably high threshold and then fit the GPD to the values in excess of that threshold.

(c) State the key advantage of the second approach over the first approach.

Solution: The GPD method has the advantage that it uses a larger part of the data and models all the large claims above the threshold, not just the single highest value.

Question 3: The claim amounts in a general insurance portfolio are independent and follow an exponential distribution with mean 2,500.

(a) Calculate the probability that an individual claim will exceed 10,000.

Solution: The claims distribution is $\text{Exp}(\frac{1}{2500})$. The probability

$$\mathbb{P}(X > 10000) = 1 - P(X \leq 10000) = 1 - F(10000) = e^{-10000/2500} = 0.0183.$$

(b) Calculate the probability that, in a sample of 100 claims, the largest claim will exceed 10,000 using:

- an exact method

  Solution: The required probability is

$$
\begin{aligned}
\mathbb{P}(\max\{X_1, ..., X_{100}\} > 10000) =& 1 - \mathbb{P}(\max\{X_1, ..., X_{100}\} \leq 10000) \\
=& 1 - \mathbb{P}(X_1 \leq 10000, ..., X_{100} \leq 10000) \\
=& 1 - \mathbb{P}(X_1 \leq 10000)^{100} \\
=& 1 - \left(1 - \exp\left(-\frac{10000}{2500}\right)\right)^{100} \\
=& 0.8425
\end{aligned}
$$

- an approximation based on a Gumbel-type GEV distribution.

You are given that, for an exponential distribution with parameter λ, the approximate distribution of $\max X_1, , ..., X_n$, for large n is a Gumbel-type GEV distribution with CDF:

$$H(x) = \exp(-\exp(-\frac{x - a_n}{b_n}))$$

where $a_n = \frac{1}{\lambda}\log(n)$ and $b_n = \frac{1}{\lambda}$.

Solution: The approximate distribution of $\max\{X_1, ..., X_{100}\}$ is a Gumbel type GEV distribution with cdf

$$H(x) = \exp(-\exp(-\frac{x - a_{100}}{b_{100}}))$$

where $a_{100} = 2500\log(100) = 11512.93$ and $b_{100} = 2500$. So

$$
\begin{aligned}
\mathbb{P}(\max\{X_1, ..., X_{100}\} > 10000) =& 1 - \mathbb{P}(\max\{X_1, ..., X_{100}\} \leq 10000) \\
=& 1 - \exp(-\exp(-\frac{10000 - a_{100}}{b_{100}})) \\
=& 0.8398.
\end{aligned}
$$

(c) State the two key assumptions made in the exact method.

Solution: The two key assumptions are that all claims come from an exponential distribution with mean 2,500 and that they are statistically independent.

Part II: Example Demonstrations

Question 4: If individual losses, W, follow a Pareto(α, λ) distribution, determine the distribution of the threshold exceedances, $X = W - u|W > u$.

Solution: The CDF of the threshold exceedances is given by

$$F_u(x) = \frac{F(x+u) - F(u)}{1 - F(u)},$$

if the individual losses follow a Pareto(α, λ) distribution, then

$$F_u(x) = \frac{(1 - (\frac{\lambda}{\lambda+x+u})^\alpha)) - (1 - (\frac{\lambda}{\lambda+u})^\alpha))}{1 - (1 - (\frac{\lambda}{\lambda+u})^\alpha))}$$

$$= \frac{(\frac{\lambda}{\lambda+u})^\alpha - (\frac{\lambda}{\lambda+x+u})^\alpha}{(\frac{\lambda}{\lambda+u})^\alpha}$$

$$= 1 - \Big(\frac{\lambda+u}{\lambda+x+u}\Big)^\alpha$$

This is the CDF of the Pareto$(\alpha, \lambda + u)$ distribution.

Question 5: Compare the limiting value of the density functions for a Gamma(α, λ) and an Exp(λ) distribution when $\alpha > 1$ and hence determine which has the heavier tail.

Solution: Comparing the two density functions and taking the limit as $x \to \infty$, we have

$$\lim_{x\to\infty} \frac{f_{Gamma}(x)}{f_{Exp}(x)} = \lim_{x\to\infty} \frac{\frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \exp(-\lambda x)}{\lambda \exp(-\lambda x)}$$

$$= \lim_{x\to\infty} C x^{\alpha-1}$$

for some constant C. For $\alpha > 1$

$$\lim_{x\to\infty} \frac{f_{Gamma}(x)}{f_{Exp}(x)} = \infty$$

Hence the Gamma distribution has the heavier tail.

Part III: Student Practice

Question 6: Determine the hazard rate for the Weibull distribution with parameters $c > 0$ and $\gamma > 0$. Comment on the behaviour of the hazard rate.

Solution: The hazard rate of the Weibull(c, γ) distribution is

$$h(x) = \frac{f(x)}{1 - F(x)}$$

$$= \frac{c\gamma x^{\gamma-1} e^{-cx^{\gamma}}}{e^{-cx^{\gamma}}}$$

$$= c\gamma x^{\gamma-1}$$

If $\gamma > 1$, then this hazard rate is an increasing function of x, which corresponds to a light tail.

If $0 < \gamma < 1$, then the hazard rate is a decreasing function of x, which corresponds to a heavy tail.

(a) Show that:

$$\int_x^{\infty} \exp(-3y^{\frac{1}{2}})dy = \frac{2}{9}\mathbb{P}\left(\chi_4^2 > 6x^{\frac{1}{2}}\right)$$

Hint: use the substitution $u = 3y^{\frac{1}{2}}$ and transform the integrand into the PDF of the Gamma(2,1) distribution.

Solution: Using the change of variable $u = 3y^{\frac{1}{2}}$, such that $y = \frac{u^2}{9}$ and $dy = \frac{2}{9}udu$.

The integral becomes

$$\int_x^\infty \exp(-3y^{\frac{1}{2}})dy = \int_{3x^{\frac{1}{2}}}^\infty \exp(-u)\frac{2}{9}udu$$

$$= \frac{2}{9}\Gamma(2)\int_{3x^{\frac{1}{2}}}^\infty \frac{1^2}{\Gamma(2)}u^{2-1}\exp(-u)du$$

$$= \frac{2}{9}\mathbb{P}(U > 3x^{\frac{1}{2}})$$

$$= \frac{2}{9}\mathbb{P}(2U > 6x^{\frac{1}{2}})$$

where $\Gamma(2) = 1! = 1$ and $U \sim$ Gamma(2,1).
Let $V = 2U$, then

$$f_V(v) = \frac{1}{2}f_U(\frac{v}{2})$$

$$= \frac{1}{2}\frac{1^2}{\Gamma(2)}(\frac{v}{2})^{2-1}\exp(-\frac{v}{2})$$

$$= \frac{(\frac{1}{2})^2}{\Gamma(2)}v^{2-1}\exp(-\frac{v}{2})$$

This is the pdf of Gamma(2,1/2). On the other hand, the $\chi_4^2 \sim$ Gamma(2,1/2). Therefore the above integral is given by $\frac{2}{9}\mathbb{P}(U > 3x^{\frac{1}{2}}) = \frac{2}{9}\mathbb{P}(\chi_4^2 > 6x^{\frac{1}{2}})$.

(b) Hence deduce an expression involving a chi-squared probability for the mean residual life for the Weibull$(3, \frac{1}{2})$ distribution.

Solution: The mean residual life beyond x is given by

$$e(x) = \frac{\int_x^\infty (1 - F(y))dy}{1 - F(x)}$$

$$= \frac{\int_x^\infty \exp(-3y^{\frac{1}{2}})dy}{\exp(-3x^{\frac{1}{2}})}$$

$$= \frac{2\mathbb{P}(\chi_4^2 > 6x^{\frac{1}{2}})}{9\exp(-3x^{\frac{1}{2}})}$$

(c) By calculating the values of mean residual life function when $x = 1$, $x = 4$ and $x = 9$, determine whether the mean residual life of the Weibull$(3, \frac{1}{2})$ distribution is an increasing or decreasing function of x.

Solution: By evaluating $e(x)$ for $x = 1, 4, 9$, we have $e(1) = 0.8887$, $e(4) = 1.5599$ and $e(9) = 2.1608$. Therefore the mean residual life time is an increasing function of x, suggesting a heavy tail.