INSTITUTE OF ACTUARIAL & QUANTITATIVE STUDIES

Subject: Statistical Techniques and Risk Management 4

Chapter: Unit 1&2

Category: Practice Questions

Extreme Value Theory – Unit 1

1. CS2 September 2020 Q2

i) Explain why Extreme Value Theory (EVT) models can be useful.

A sports scientist is interested in analysing the probability that the javelin world record may be broken next year and is intending to use EVT to do this. The sports scientist has obtained data for the distances of all javelin throws from all javelin competitions in 2019. The total number of throws recorded was 3,000. The sports scientist has carried out an EVT analysis using the Generalised Pareto Distribution by selecting only those throws that exceeded 50 metres.

This resulted in the longest 150 throws being selected for the analysis.

The following parameters were obtained from the EVT analysis:

$\beta = 15$, and $\gamma = 3$.

ii) Determine the percentage of javelin throws that would be expected to exceed 70 metres next year.

iii) Comment on the limitations of this analysis.

2. CS2 September 2021 Q1

An Analyst is assessing the risks of an equity portfolio and wishes to estimate the probability that the portfolio will incur at least one daily loss exceeding 5% next month.

Explain how a generalised extreme value distribution and the block maxima method could be used to estimate this probability.

3. CS2A S2022 Q9

An airline company is concerned about potential financial losses due to currency movements in various countries. An analyst has been hired to model daily currency movements in these economies. The objective is to understand the likelihood of extreme currency movements that may have major financial consequences for the airline.

The analyst has suggested fitting several types of standard statistical distributions to the entire datasets, using maximum likelihood estimation, including Normal, Gamma, Pareto and lognormal. Estimates can then be made regarding probabilities of extreme currency movements based on the chosen model.

(i) Discuss the problems with this approach. [4]

The analyst proceeds to use the Generalised Pareto Distribution (GPD) to assess only extreme daily losses. For this analysis, currency losses were treated as positive, and gains were treated as negative.

A threshold was set at daily losses of +2.2%, resulting in a data subset of 150 'extreme' values from the 3,000 original data points to be used in the extreme value analysis. The following results were obtained from the subsequent analysis:

- Beta = 0.009785.
- Gamma = 5.417 (shape).

(ii) Calculate G(0.02) using the above parameter values, where G is the distribution function of the GPD, and write down its meaning. [3]

(iii) Estimate the probability that the currency losses tomorrow exceed 5%, based on this GPD model. [3]

A Normal distribution was also fitted to the data. The best fitting Normal distribution was determined to be N(0, 0.014^2). However, currency returns are generally considered to exhibit leptokurtic behaviour.

(iv) Calculate the probability that daily losses exceed 5% using this Normal distribution, and comment on the results. [3]

(v) Discuss further investigations the analyst should proceed with, including any problems with using extreme value theory and/or the GPD in this case. [2] [Total 15]

4. CS2A A2023 Q6

A hydroelectric company is managing a water reservoir created from a dam in a river valley. The dam was originally chosen so that the water level would exceed a threshold of 50 metres in about 2 days in every 300 days. In these extreme events, the excess water is left to escape the reservoir so that the water level is kept below the safety 50-metre limit. It is believed that the daily water level in the reservoir follows an exponential distribution with mean μ.

(i) Estimate the value of μ. [3]

(ii) Determine the expected threshold exceedance of the water level over the 50-metre threshold. [2]

In order to better manage the excess water, it is now assumed that the excess water level follows a Generalised Pareto distribution with scale parameter $\beta = 1$.

(iii) Explain the circumstances in which the Generalised Pareto distribution would be preferred to the exponential distribution. [2]

(iv) Estimate the value of the parameter γ if the expected threshold exceedance is the same as that in part (ii). [3] [Total 10]

Copulas – Unit 2

1. CS2 April 2019 Q7

(i) Write down Sklar's theorem.

(ii) Explain, in words, the meaning of the following copula expression: C(u1, u2, u3)

The Gumbel copula has a generating function:

$$\text{Gumbel}\psi_\alpha [F(x)] = (-\ln F(x))^\alpha$$

(iii) Derive an expression for the Gumbel (Hougaard) copula for the case where there are three variables.

A student has fitted a Gumbel copula to investment returns from three developing markets, and has calculated a value for the dependency parameter, α, of 4.0. She has separately determined that the probability of making a loss over the next calendar year (i.e. the probability that the return is less than 0%) in each of the three markets is 5%, 7.5% and 10% respectively.

(iv) Calculate the probability that all three markets have returns of less than 0% over the next calendar year. (v) State what type of copula is equivalent to a Gumbel copula if α = 1.0.

(vi) Calculate the probability that all three markets have returns of less than 0% over the next calendar year, assuming that each of the markets were independent.

2. CS2 September 2020 Q1

Consider a two-dimensional Gaussian copula function, $C_{\text{Gauss}}(u_1, u_2)$, with parameter $\rho = 0$:

(i) Give the solution to the copula function $C_{\text{Gauss}}(1, 1)$. [1]

(ii) Give the solution to the copula function $C_{\text{Gauss}}(1, 0.2)$. [1]

(iii) Give the solution to the copula function $C_{\text{Gauss}}(0.2, 0.2)$. [1]

(iv) Outline how your answers to parts (i), (ii) and (iii) would change if $\rho = 1$. [2]

An insurer uses copulas to model the dependencies between various types of claims. A statistical analysis of the insurer's claims shows that a Gumbel copula is a better representation of historic claims data than a Gaussian copula.

(v) Discuss why the use of a Gaussian copula in a claims model could result in solvency issues for this insurer. [3]

[Total 8]

3. CS2 April 2021 Q1

The Frank copula, C_F, for a bivariate distribution is defined as:

$$C_F(u, v) = -\frac{1}{\alpha} ln\left(1 + \frac{(e^{-\alpha u} - 1)(e^{-\alpha v} - 1)}{(e^{-\alpha} - 1)}\right), \alpha > 0$$

(i) Determine the probability that two jointly distributed random variables, X and Y, are both less than or equal to their median values where X and Y follow the Frank copula, C_F, with $\alpha = 1$. [2]

(ii) Determine the revised value of the probability in part (i) when $\alpha = 0.1$. [1]

(iii) Determine the probability that two jointly distributed random variables, X and Y, are both less than or equal to their median values where X and Y follow the product copula. [1]

(iv) Comment on your answers to parts (i), (ii) and (iii) with reference to the sign and level of dependence exhibited by the Frank copula. [2]

[Total 6]

4. CS2A S2022 Q7

An insurer writes three different classes of insurance business: X, Y and Z. The classes have the following total annual claims distributions:

$X \sim Exp(0.08)$
$Y \sim Normal(10, 2^2)$
$Z \sim Normal(20, 3^2)$

The insurer models the level of dependency between the classes' total claim amounts using a Clayton (α = 2) copula.

i) Calculate the probability, using the Clayton (2) copula, that both $X < 3$ and $Y < 8$. [3]

ii) Calculate the probability, using the Clayton (2) copula, that all of the following occur: $X < 3$, $Y < 8$ and $Z < 20$. [3]

iii) Calculate the probability, using the Clayton (2) copula, that both $X > 10$ and $Y > 12$. [3]

A student has noted that copulas are a useful modelling tool as 'they allow us to model different degrees of dependency'.

iv) Comment on this statement. [3] [Total 12]