Exam I: Information and Practice Problems

Statistics 509 – Winter 2022

Feb 16, 2022

Information: The Midterm Exam will be Thursday, February 24, from 7:00-9:00pm in room **SEB1202** (in the School of Education Building). The exam will cover the material in Lectures 1-6, and parts of Lecture 7 covered through class on Monday/Feb 21, and corresponding relevant material in the book. The exam is open textbook (only the textbook) and open notes. Below is a set of practice Exam Problems – solutions for the problems will be posted on Monday, February 24. Additional practice problems are the exercise problems from the lectures and exercises in the book.

Practice Exam Problems

1. Suppose $X \sim \mathcal{N}(0, 1)$ and it is known that the .55, .95 and .975 quantiles of the standard normal distribution are 0.126, 1.645, and 1.96, respectively.

(a) Find the .95 quantile of $Y = X^3$.

(b) Find the .95 quantile of $Y = X^2$.

(c) Find the .95 quantile of $Y = \frac{1}{X}$.

(d) Find transformation/function g such that $Y = g(X) \sim \text{DExp}(0, 2)$. Give a rigorous proof validating your choice of the function g. *Hint:* Can express it in terms of the cdfs and inverse cdfs of $\mathcal{N}(0, 1)$ and $\text{DExp}(0, 2)$.

2. Problem 7, parts (a)-(c) on page 17 of Ruppert/Matteson with the definitions of

$$r_1(2) = r_1 + r_0$$

$$r_2(2) = r_2 + r_1$$

and knowing the following output from R

```
> pnorm(1.44)
[1] 0.9250663
> pnorm(1.28)
[1] 0.8997274
> pnorm(1.65)
[1] 0.9505285
```

3. Suppose that X_1, X_2, X_3, X_4 are iid $\text{DExp}(0, 1)$.

(a) Suppose and $Y = X_1 + X_2$ and $Z = X_2 + X_3 + X_4$. Derive the covariance and correlation of Y, Z.

(b) Suppose $U = X_1\sqrt{|X_3|}$ and $V = X_2\sqrt{|X_3|}$. Show that U, V are uncorrelated.

(c) Show that U, V in **(b)** are not independent.

4. Suppose (X, Y) has a bivariate normal distribution with expected values $\mu_X = \mu_Y = 0$ and matrix

$$\boldsymbol{\Lambda} = \begin{bmatrix} \sigma_X^2 & \rho\sigma_X\sigma_Y \\ \rho\sigma_X\sigma_Y & \sigma_Y^2 \end{bmatrix}$$

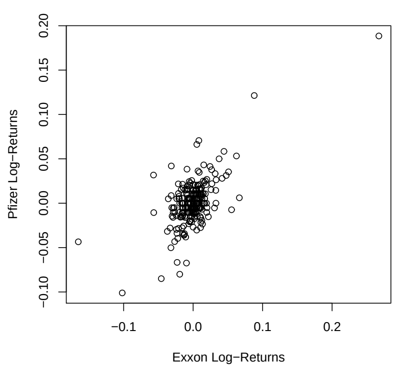
Suppose $U = e^X, \;\; V = e^{Y^3}, \;\; W = e^{X-Y}$

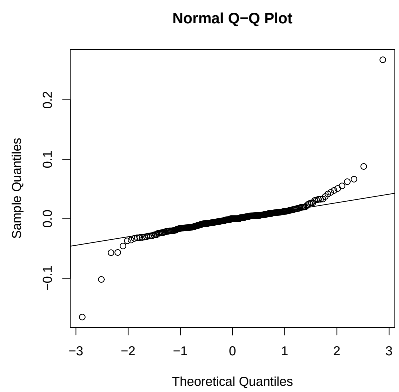
(a) Derive the form of the appropriate copula model for U, V, i.e., specify the parametric copula model (and the parameter values). Justify your answer.

(b) Derive the form of the appropriate copula model for U, W, i.e., specify the parametric copula model (and the parameter values). Justify your answer.

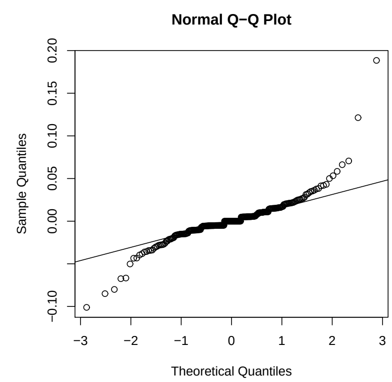
5. The following is a

- scatter diagram of Exxon daily log-returns vs Pfizer daily log-returns (2005-2006)
- Normal QQ plot of Exxon daily log-returns
- Normal QQ plot of Pfizer daily log-returns





Exxon - Normal QQ



Pfizer - Normal QQ

(a) Suppose for this data have a Pearson correlation of 0.655 and Spearman correlation of 0.419. Which do you believe to be a better quantification of dependency? Justify your answer.

(b) Suppose that you computed the Pearson correlation and Spearman on the raw return (no logarithm taken) data? What changes do you expect in these correlation parameters and why?

6. Suppose X, Y are two random variables corresponding to the relative returns of two securities, and suppose that $X \sim \text{DExp}(0, 40)$, $Y \sim \text{DExp}(0, 40)$ and suppose (X, Y) have a Gumbel copula model with $\theta = 4$, i.e., with $\theta = 4$ the bivariate cdf is

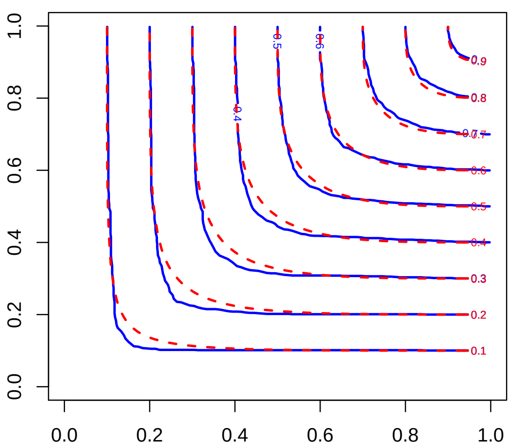
$$C(u_1, u_2) = e^{-[(-\log u_1)^\theta + (-\log u_2)^\theta]^{\frac{1}{\theta}}}$$

(a) Suppose X', Y' are given as $X' = e^X$ and $Y' = Y^3$. Derive the copula model for (X', Y') and justify your answer.

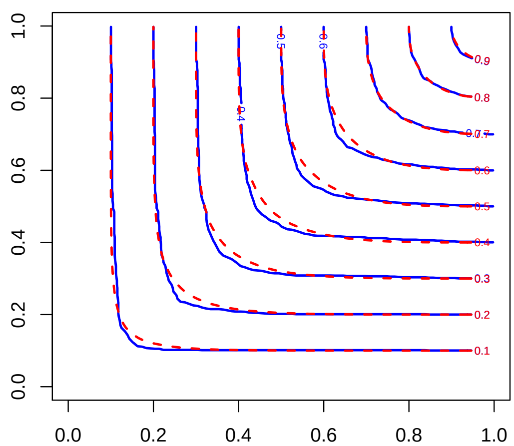
(b) Derive the probability that both of the securities will have negative returns (at the same time). Also derive the probability that both of the securities will have returns below -.03 (at the same time)

(c) Based on the copula model, what can you say about whether X and Y are independent or not. Justify your answer.

7. Suppose that the log-return data from 2 assets are both modelled with a normal distribution and the parameters estimated. The log-return data is then converted, via the appropriate estimated normal distribution cumulative distribution functions to Copula data, and two different copula models estimated – one with Gumbel copula and one with a multivariate Gaussian copula. The plots of the empirical and theoretical bivariate cumulative distributions are shown below (solid is empirical and dashed is the theoretical).



Contour Plot - Gumbel Model Fit



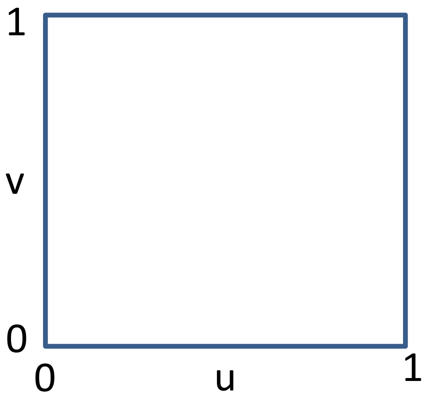
Contour Plot - Gauss Model Fit

(a) Which model do you prefer and why?

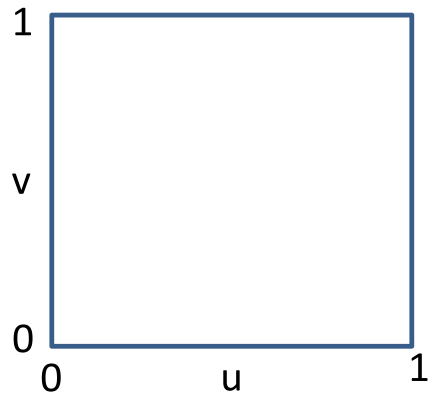
(b) Why might you not be satisfied with the model you selected in **(a)**, and what would you do? Be specific about what steps you might you do to improve the overall joint modelling of the return, i.e., be specific about better models/modelling for the marginal distributions and the copula models.

(c) What "quantitative" statistical methods, as presented in class, can help you assess which copula models are best?

8. Sketch out contours (for values of .2, .4, .6, and .8) for the bivariate cumulative distribution function for the independent copula model and the co-monotonicity copula model below.



Independent Copula



Co-monotonicity Copula

9. Suppose one has the daily log-returns of a particular assets over 10-yr period.

(a) The following output statistics were computed

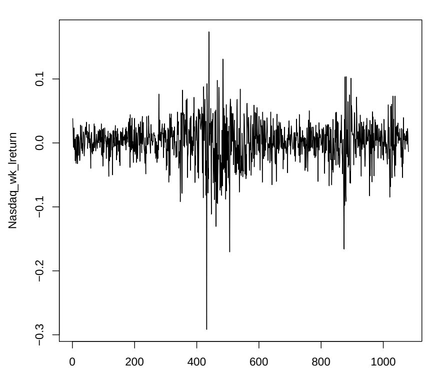
median	mean	stand. dev	skewness	(excess) kurtosis
0.0036	0.0018	0.0274	-0.190	6.870

Give a brief discussion on what these statistics suggests, and make sure to include discussions about the whether log-returns follow a normal distribution.

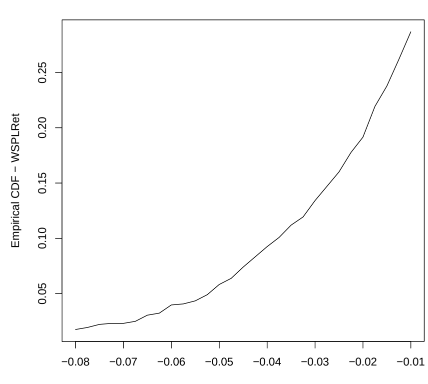
(b) Based on these descriptive statistics, what test might you apply to statistically test whether the log-returns are normal.

(c) Suppose for the test you suggested in **(b)** rejected the hypothesis of normally distributed log-returns at the $\alpha = .01$ level. Does that automatically require you look to use a different distributional family (e.g., t-distribution or double exponential)? Why or why not? Note that 10 years corresponds to approximately 2500 log-returns in the data set.

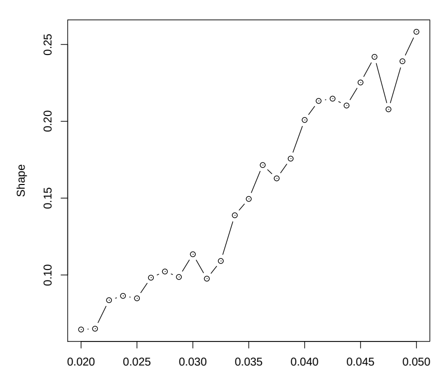
10. Suppose for a GPD analysis for Nasdaq data, we carried out the shape plot on lower tails and generated shape plots and this also shown below.



Nasdaq wkly log-returns vs. time



Empircal CDF Plot



Lower-Tail Shape Plot

(a) If you had to pick a threshold and associated estimate for the shape parameter for a POT model, what would it be and discuss/justify your selection.

(b) What is the underlying statistical model assumptions that is made in doing the MLE of the shape parameters for the log-returns over a specific threshold? Be specific.

(c) Does the shape plot on the right provide a good justification for the POT threshold model you specified in **(a)**? Based on the plot, what important assumptions appears to be violated, and be specific about the nature of the violation.

11. If random variable X is exponential with rate parameter λ, then $Y = X + \mu$ is said to have a shifted exponential distribution with shift parameter μ and rate parameter λ.
(a) What is the mean and variance of Y?

(b) What is the cumulative distribution function (cdf) of Y. Note that the cdf needs to be defined for all values of y.

(c) Suppose $R_1, R_2, \ldots, R_{1000}$ is a time series of negative returns (i.e., (-1) times the original returns) with

(i) a sample .80-quantile of .02, and

(ii) for POT, the model for the tail distribution of the negative returns above the .02 threshold as being exponential with MLE's $\hat{\mu} = .02$ and $\hat{\lambda} = 60$.

Derive the semiparametric estimate of the relative VaR and the expected shortfall for $\alpha = .005$.

12. **(a)** Problem 1 on page 491 in Ruppert/Matteson.

(b) Problem 2 on page 492 in Ruppert/Matteson.

13. Problem 5 on page 492 in Ruppert/Matteson.

14. **(a)** Problem 6 on page 492 in Ruppert/Matteson, assuming multivariate normal distribution for the log-returns. Also derive the VaR corresponding to .05, given that the .05-quantile for the standard normal distribution is -1.645.

(b) Do Problem 6 on page 492 in Ruppert/Matteson, assuming a multivariate t-distribution for the log-returns with shape parameter $\nu = 6.5$. Also derive the VaR corresponding to .05, and given that the .05-quantile of the standard univariate t-distribution with $\nu = 6.5$ is -1.917.

15. Problem 8 on page 513 in Ruppert/Matteson, but change part (d) to be a 4% expected return.

16. Problem 11 on page 513 in Ruppert/Matteson.