Exam I: Information and Practice Problems with Solutions

Statistics 509 – Winter 2022

Feb 21, 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. Additional practice problems are the exercise problems from the lectures and exercises in the book.

Practice Exam Problems with Solutions

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 $Y = g(X) \sim \mathrm{DExp}(0,2)$. 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 $\mathrm{DExp}(0,2)$.

(a) The function $g(x) = x^3$ is increasing, so by result in class,

$$y_{.95} = g(x_{.95}) = (1.645)^3 = 4.4514$$

(b) In this case, we want $y_{.95}$ so that

$$\begin{aligned}
.95 &= P(Y \le y_{.95}) \\
&= 1 - P(Y \ge y_{.95}) \\
&= 1 - P(X^2 \ge y_{.95}) \\
&= 1 - [P(X \ge \sqrt{y_{.95}}) + P\left(X \le -\sqrt{y_{.95}}\right)] \\
&= 1 - 2P\left(X \ge \sqrt{y_{.95}}\right) \qquad \text{by symmetry of } \mathcal{N}(0,1) \\
&= 1 - 2\left\{1 - P\left(X \le \sqrt{y_{.95}}\right)\right\}
\end{aligned}$$

Doing the algebra in the above, we have

$$P(X \le \sqrt{y_{.95}}) = .975$$

so that this implies that $\sqrt{y_{.95}} = x_{.975}$, i.e.,

$$y_{.95} = (x_{.975})^2 = (1.96)^2 = 3.8416$$

(c) In this case, note that since X has a pdf symmetric about 0, the rv $\frac{1}{X}$ will also have a pdf symmetric about 0, so we know that $y_{.95} > 0$. Now note that

$$\left\{\frac{1}{X} \leq y_{.95}\right\} = \{X \leq 0\} \cup \left\{X \geq \frac{1}{y_{.95}}\right\}$$

so that

$$\begin{aligned}
.95 = P(Y \leq y_{.95}) &= P\left(\frac{1}{X} \leq y_{.95}\right) \\
&= P(X \leq 0) + P\left(X \geq \frac{1}{y_{.95}}\right) \\
&= .5 + 1 - P\left(X \leq \frac{1}{y_{.95}}\right)
\end{aligned}$$

which after algebra implies that

$$P\left(X \leq \frac{1}{y_{.95}}\right) = .55$$

This in turn implies that

$$\frac{1}{y_{.95}} = x_{.55} = 0.126 \quad \text{which implies} \quad y_{.95} = \frac{1}{0.126} = 7.937$$

(d) Let F_Y denote the cdf DExp$(0, 2)$ and let Φ denote the cdf of $\mathcal{N}(0, 1)$. Then for y, we want

$$P(g(X) \leq y) = F_Y(y)$$

or

$$P(X \leq g^{-1}(y)) = F_Y(y)$$

if g is increasing. But the above implies that want

$$\Phi(g^{-1}(y)) = F_Y(y)$$

or

$$g^{-1} = \Phi^{-1}F_Y$$

or

$$g = F_Y^{-1}\Phi$$

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
```

(a) The distribution of

$$r_t(4) = r_t + r_{t-1} + r_{t-2} + r_{t-3}$$

is normal with mean and variance of

$$E(r_t(4)) = 4 \cdot .06 = .24, \quad \mathrm{Var}(r_t(4)) = 4 \cdot (.47) = 1.88$$

(b) Note that

$$P(r_t(4) < 2) = P\left(\frac{r_t(4) - .24}{\sqrt{1.88}} \leq \frac{2 - .24}{\sqrt{1.88}}\right) = \Phi(1.28) = .8997$$

where we have invoked that the $\frac{r_t(4)-.24}{\sqrt{1.88}}$ is standard normal, i.e., mean of 0 and variance of 1, and Φ is the pdf of this distribution.

(c) The covariance is

$$\begin{aligned}
\mathrm{Cov}(r_1(2), r_2(2)) &= \mathrm{Cov}(r_1 + r_0\ ,\ \ r_2 + r_1) \\
&= \mathrm{Cov}(r_1, r_2) + \mathrm{Cov}(r_1, r_1) + \mathrm{Cov}(r_0, r_2) + \mathrm{Cov}(r_0, r_1) \\
&= \mathrm{Cov}(r_1, r_1) \qquad \text{(since independence implies the other covariances are 0)} \\
&= .47
\end{aligned}$$

3. Suppose that X_1, X_2, X_3, X_4 are iid 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.

Based on results on double exponential,

$$\mathrm{Var}(X_i) = E(X_i^2) = \frac{2}{1^2} = 2.$$

(a) Now variances are

$$\mathrm{Var}(Y) = \mathrm{Var}(X_1) + \mathrm{Var}(X_2) = 2 \cdot 2 = 4$$

$$\mathrm{Var}(Z) = \mathrm{Var}(X_2) + \mathrm{Var}(X_3) + \mathrm{Var}(X_4) = 3 \cdot 2 = 6$$

and since $\text{Cov}(X_i, X_{i'}) = 0$ for $i \neq i'$,

$$\text{Cov}(Y, Z) = \text{Cov}(X_1 + X_2, X_2 + X_3 + X_4) = \text{Cov}(X_2, X_2) = 2.$$

Thus the correlation is

$$\text{Corr}(Y, Z) = \frac{\text{Cov}(Y, Z)}{\sqrt{\text{Var}(Y) \cdot \text{Var}(Z)}} = \frac{2}{\sqrt{4 \cdot 6}} = \frac{1}{\sqrt{6}}$$

(b) By independence of the rvs,

$$E(U) = E\left(X_1\sqrt{|X_3|}\right) = E(X_1)E\left(\sqrt{|X_3|}\right) = 0$$

$$E(V) = E\left(X_2\sqrt{|X_3|}\right) = E(X_2)E\left(\sqrt{|X_3|}\right) = 0$$

so that

$$\text{Cov}(U, V) = E(UV) = E\left(X_1\sqrt{|X_3|}X_2\sqrt{|X_3|}\right) = E(X_1)E(X_2)E(|X_3|) = 0$$

(c) Note that it is easy to show that

$$E(|X_3|) = 1$$

and based on this

$$E(U^2) = E(X_1^2|X_3|) = E(X_1^2)E(|X_3|) = 2$$
$$E(V^2) = E(X_2^2|X_3|) = E(X_2^2)E(|X_3|) = 2$$

But

$$E(U^2V^2) = E(X_1^2X_2^2X_3^2) = E(X_1^2)E(X_2^2)E(X_3^2) = 8$$

and this is not equal to the product of $E(U^2) \cdot E(V^2)$, implying that U and V are not independent.

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

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

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.

(a) Let $g(x) = e^x$ and let $h(y) = e^{y^3}$ and note that these are monotonically increasing functions. By results in class, this implies that the copula model for (U, V) is identical to the copula model for (X, Y) and as defined in class (and in the book), this corresponds to the Gaussian copula model with parameter ρ.

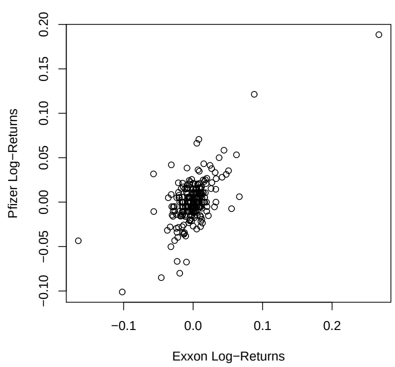
(b) In this case, note that $(X, X - Y)$ are multivariate normal since there are linear combinations of X and Y. The correlation is given by

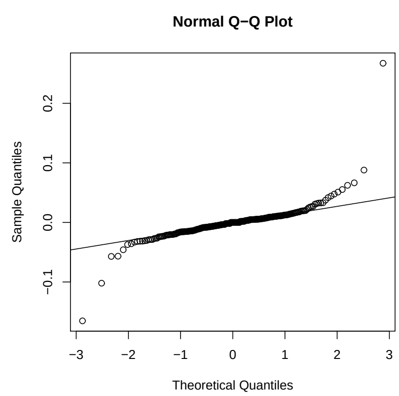
$$
\begin{aligned}
\mathrm{Corr}(X, X-Y) &= \frac{\mathrm{Cov}(X, X-Y)}{\sqrt{\mathrm{Var}(X)\cdot \mathrm{Var}(X-Y)}} \\
&= \frac{\mathrm{Cov}(X,X) - \mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\cdot [\mathrm{Var}(X) + \mathrm{Var}(Y) - 2\mathrm{Cov}(X,Y)]}} \\
&= \frac{\sigma_X^2 - \rho\sigma_X\sigma_Y}{\sqrt{\sigma_X^2 \cdot (\sigma_X^2 + \sigma_Y^2 - 2\rho\sigma_X\sigma_Y)}} \\
&= \frac{\sigma_X - \rho\sigma_Y}{\sqrt{\sigma_X^2 + \sigma_Y^2 - 2\rho\sigma_X\sigma_Y}}
\end{aligned}
$$

Since $\ell(z) = e^z$ is an increasing function, U, W have the same copula model as $(X, (X-Y))$ and this is the Gaussian copula model with correlation parameter of $\rho' = \frac{\sigma_X - \rho\sigma_Y}{\sqrt{\sigma_X^2+\sigma_Y^2-2\rho\sigma_X\sigma_Y}}$.

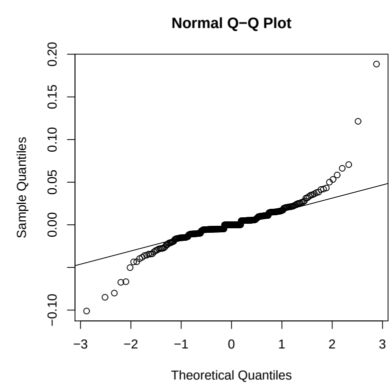
5. The following is a

- 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?

(a) The normal QQ plot for the Pfizer log-returns seems to suggests that the distribution is significantly heavier tailed than the normal distribution. This renders the Pearson correlation coefficient suspect in terms of estimating dependency. Thus a better quantification of dependency would be the Spearman correlation coefficient.

(b) The Spearman correlation coefficient would remain unchanged since it is unchanged under monotonically increasing functions of the data. The Pearson correlation coefficient would be more variable due to the heavy-tailed nature of the data and hence it would change, but it is not clear which direction.

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.

(a) The copula model for (U, V) is again the same Gumbel copula model since both of these transformations of X and Y are monotonically increasing.

(b) The first probability is given by

$$\begin{aligned}
P(X \le 0, Y \le 0) &= C_4(F_X(0), F_Y(0)) \\
&= C_4(.5, .5) \\
&= e^{-\left[(-\log(.5))^4 + (-\log(.5))^4\right]^{\frac{1}{4}}} \\
&= 0.439
\end{aligned}$$

and the second probability is given by

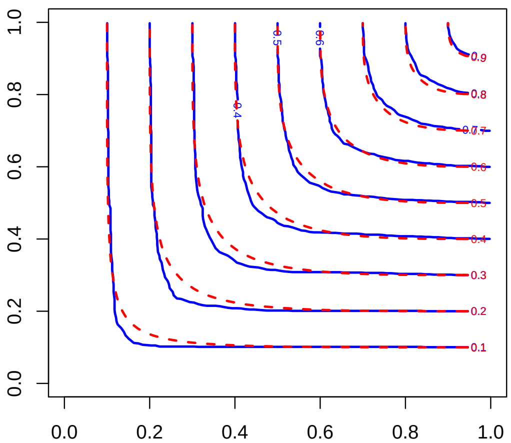
$$\begin{aligned}
P(X \le -.03, Y \le -.03) &= C_4(F_X(-.03), F_Y(-.03)) \\
&= C_4(.151, 151) \\
&= e^{-\left[(-\log(.151))^4 + (-\log(.151))^4\right]^{\frac{1}{4}}} \\
&= 0.106
\end{aligned}$$

since

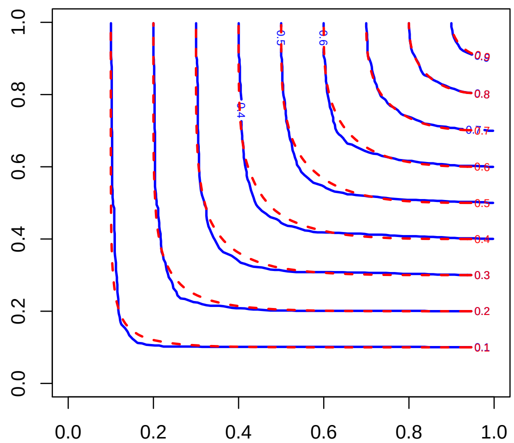
$$F_X(-.03) = F_Y(-.03) = \frac{1}{2}e^{-40*|-.03|} = \frac{1}{2}e^{-1.2} = .151$$

(c) If X and Y were independent, then $F_X(X)$ and $F_Y(Y)$ would be independent uniform random variables, and that corresponds to the Gumbel copula model with $\theta = 1$, i.e., $C_1(u, v) = uv$. With $\theta = 4$, it is clear that $C_4(u, v) \ne uv$, and so they are not independent, so X and Y are not independent.

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.

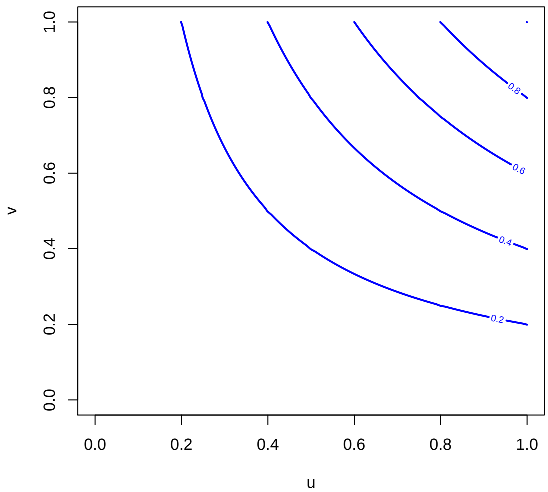
(c) What "quantitative" statistical methods, as presented in class, can help you assess which copula models are best?

(a) The Gauss copula model seems to have an overall better fit, and in particular it fits better in the southwest corner, which is really important for risk assessment. The fit is pretty good in that region, but it could be better.

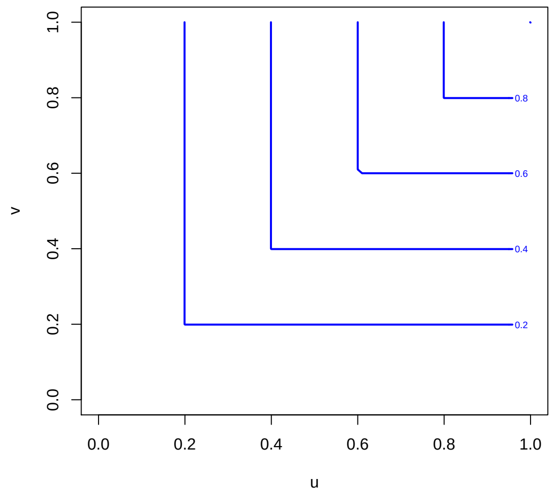
(b) One of the reasons is that the Gaussian copula model has a coefficient of tail dependence that is 0, and that is a strong assumption on tail dependencies. To do a better fit to marginals, I would consider utilizing a t-distribution to ensure that fit the tails even a little better (the normal is a special case of the t with large degrees of freedom). Similarly, I would consider a t-copula, again trying to get an overall better fit, and in particular a better fit in the southeast corner. This again makes sense since the t-copula distribution is again a sort of generalization of the Gaussian copula, when you have a large degrees of freedom parameter ν.

(c) I would consider the AIC criteria as a quantitative measure for helping to select the copula models which provide a best fit.

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.

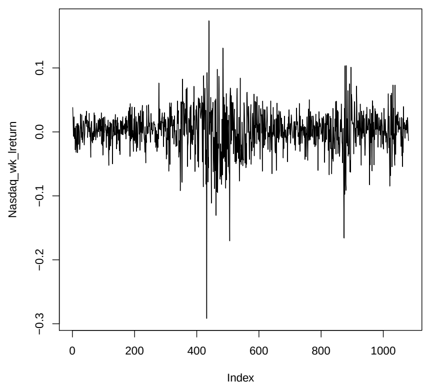
(c) Suppose for the test you suggested in **(b)** rejected the hypothesis of normally distributed log-returns at the $\alpha = .01$ level. Does that automatically require you look to use a different distributional family (e.g., t-distribution or double exponential)? Why or why not? Note that 10 years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

(a) The means are close to 0, which is fairly standard. The skewness is not very large, suggesting that the distribution is fairly symmetric. The kurtosis is quite large, suggesting the tails of the distribution of the log-returns are quite heavy. It appears heavier than even double exponential (since that would have value of approximately 3).

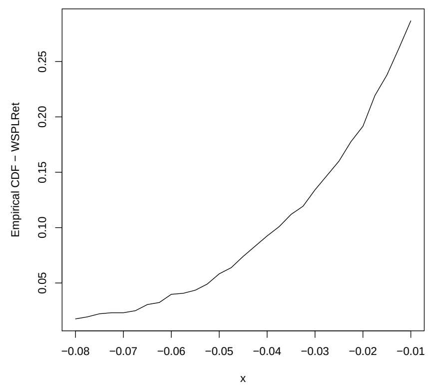
(b) Could use the Jarque-Bera test which utilizes the skewness and kurtosis to do a test for normality.

(c) It does not – almost any goodness-of-fit test for normality would reject at $\alpha = .01$ level with 2500 data points. With this much data, any slight deviation will be detected as being statistically significant, even the model may actually be a good approximation.

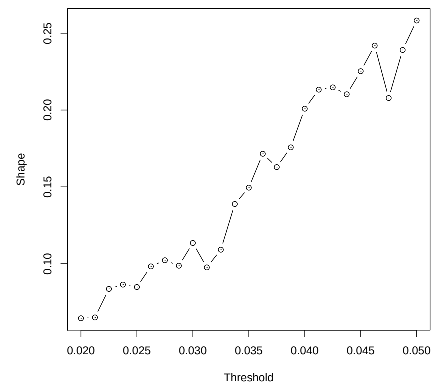
10. Suppose for a GPD analysis for Nasdaq data, we carried out the shape plot on lower tails and generated shape plots and this also shown below.



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.

(a) I would pick threshold of -.03 to -.04, probably -.035, and the resultant shape plot would be approximately .14. It is not a great model, as the shape plot is increasing, and not stable as you would like, but -.035 allows us to be estimating with approximately 10% of the data, which is important to have enough for reliable estimates.

(b) The model is that the log-returns are independent and identically distributed with the Generalized Pareto distribution being appropriate for the tails.

(c) The shape plot does not provide a good justification, as it is not constant, but is increasing as the threshold increases. Thus suggest that the tail distribution is tending to get heavier (or more volatile) as the threshold increases, i.e., as the losses get larger.

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

$\alpha = .005$.

(a) The mean and variance are

$$E(X) = E(X) + \mu = \frac{1}{\lambda} + \mu$$

$$V(X) = \mathrm{Var}(X) = \frac{1}{\lambda^2}$$

(b) The cdf of Y is given by

$$F_Y(y) = P(X + \mu \leq y) = P(X \leq y - \mu) = F_X(y - \mu)$$

Based on the properties of the exponential cdf, since $x - \mu \geq 0$ if and only if $x \geq \mu$, the final form of the cdf is

$$F_Y(y) = \begin{cases} 1 - e^{-\lambda(y-\mu)} & y \geq \mu \\ 0 & y < \mu \end{cases}$$

(c) Now the semi-parametric estimate of the relative $\tilde{\text{VaR}}$, when considering the negative returns, Y, is

$$.005 = \alpha = P(Y \geq \tilde{\text{VaR}}) = (1 - \hat{F}_n(\mu))(1 - F_{\hat{\lambda},\hat{\mu}}(\tilde{\text{VaR}}))$$

with $\hat{F}_n$ being the empirical cdf and $F_{\lambda,\mu}$ being the cdf of the shifted exponential. So with $\mu = .02$,

$$1 - F_{\hat{\lambda},.02}(\tilde{\text{VaR}}) = \frac{.005}{(1 - \hat{F}_n(.02))} = \frac{.005}{.20} = .025.$$

This corresponds to

$$e^{-60\cdot(\tilde{\text{VaR}} - .02)} = .025$$

or

$$\tilde{\text{VaR}} = -\frac{1}{60}\log(.025) + .02 = 0.0814$$

For expected shortfall, this is just the expected value of the the exponential random variable Y, conditional on it being greater than the $\tilde{\text{VaR}}$. Thus, letting $f_{\lambda,\mu}$ being the pdf of the shifted exponential, the expected shortfall is

$$\begin{aligned}
\text{ES} &= E(Y|Y \geq .0814) \\
&= \frac{1}{P(Y \geq .0814)} \int_{.0814}^{\infty} y f_{60,.02}(y)\, dy \\
&= \frac{1}{.025} \cdot \int_{.0814}^{\infty} y \cdot 60 \cdot e^{-60\cdot(y-.02)}\, dy \\
&= \frac{e^{1.2}}{.025} \cdot \int_{.0814}^{\infty} y \cdot 60 \cdot e^{-60\cdot y}\, dy \\
&= \frac{e^{1.2}}{.025} \left\{ \left[-y e^{-60\cdot y}\right|_{.0814}^{\infty} + \int_{.0814}^{\infty} e^{-60\cdot y}\, dy \right\} \qquad \text{int. by parts} \\
&= \frac{e^{1.2}}{.025} \left\{ (.0814) \cdot e^{-60\cdot .0814} + \frac{1}{60} \cdot e^{-60\cdot .0814} \right\} \\
&= 0.0985
\end{aligned}$$

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

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

(a), Problem 1-(a) For this we want weight value w such that

$$\mu_R = w * (2.3) + (1 - w) * 4.5 = 4.5 - w(4.5 - 2.3) = 3$$

or

$$w = \frac{1.5}{2.2} = .682$$

Problem 1-(b) Now want

$$\begin{aligned} \sigma_w^2 - 5.5 &= w^2 * 6 + (1 - w)^2 * 11 + 2w(1 - w) * .17\sqrt{6 * 11} - 5.5 \\ &= (6 + 11 - 2 * .17\sqrt{6 * 11})w^2 + (-2 * 11 + 2 * .17\sqrt{6 * 11})w + 11 - 5.5 \\ &= aw^2 + bw + c = 0 \end{aligned}$$

or

$$w = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

There are two solutions – $w = .411$ and $w = .940$, and it is clear that the former value of $w = .411$ provides the portfolio with the higher return as it weights more towards asset 2, which is has the higher expected return.

(b) Now want to have a portfolio being the tangent portfolio with the fixed asset, i.e.,

$$R_p = wR_T + (1 - w)R_f$$

Now want

$$\sigma_{R_p} = w\sigma_{R_T} = w * 7 = 5$$

implies that

$$w = \frac{5}{7}$$

so want

$$\frac{5}{7} * .65 = .464 \text{ in asset C}$$

$$\frac{5}{7} * .35 = .250 \text{ in asset D}$$

$$\frac{2}{7} = .286 \text{ in risk-free asset}$$

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

(a) The mean for the portfolio $R_P = wR_1 + (1 - w)R_2$ is given

$$E(R_P) = w \cdot .0047 + (1 - w) \cdot .0065 = .0065 - w \cdot .0018$$

and setting that equal to .005, we have the weight for this (efficient) portfolio

$$w = \frac{.0065 - .005}{.0018} = .833$$

Note that it is the only weight that works.

(b) The estimated variance is given by

$$\sigma_P^2 = .833^2 \cdot .0125 + .167^2 \cdot .023 + 2 \cdot .833 \cdot .167 \cdot .0058 = 0.0109$$

(c) Now the actual expected return and variance, with the original means, variances, and covariance is

$$\mu_P = .833 \cdot .0032 + .167 \cdot .0074 = 0.0039$$

and

$$\sigma_P^2 = .833^2 \cdot .017 + .167^2 \cdot .025 + 2 \cdot .833 \cdot .167 \cdot .0059 = 0.0141$$

Note that these values are quite a bit different than predicted by the bootstrap values.

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.

(a) The covariance between the returns is

$$\text{Cov}(R_1, R_2) = \rho_{12} * \sigma_1 * \sigma_2 = .35 * .03 * .04 = 0.00042$$

so the covariance matrix for $[R_1 R_2]^\top$ is given

$$\mathbf{\Sigma} = \begin{bmatrix} .0009 & 0.00042 \\ 0.00042 & .0016 \end{bmatrix}.$$

The weight w for stock 1 (and hence $(1-w)$ for stock 2) is given by

$$w = \frac{100 * 200}{100 * 200 + 100 * 125} = .615$$

and so the mean return is

$$\mu_P = .615 * .001 + .385 * .0015 = 0.00119$$

and the standard deviation is given by

$$\sigma_P = \sqrt{.615^2 * .03^2 + .385^2 * .04^2 + 2 * .615 * .385 * .35 * .03 * .04} = .0279$$

Finally for the VaR(.05), the relative VaR is given as the negative .05-quantile of the distribution on the log-returns with the above mean and standard deviation – that is derived from utilizing the relationship between standard normal quantiles and quantiles for general normal distributions, i.e.,

$$q_{.05} = \mu_P - 1.645 * \sigma_P = 0.00119 - 1.645 * .0279 = -0.0447$$

and so to get the VaR, just multiple the relative VaR by the portfolio value, i.e.,

$$\text{VaR}(.05) = -(-0.0447) * (100 * 200 + 100 * 125) = 1453$$

so \$1453 is the VaR(.05).

(b) The means and covariances all stay the same as in part **(a)**. Note as shown/stated in class, the sum of random variables from a multivariate t-distribution are still t-distributed with same shape parameter. To compute the quantile of the appropriate t-distribution of the weighted sum of returns, recall that the scale parameter in the t-distribution is given as

$$\lambda = \sqrt{\frac{\nu - 2}{\nu}}\sigma_P = \sqrt{\frac{6.5 - 2}{6.5}} * .0279 = .0232$$

and thus the quantile is given by

$$q_{.05} = \mu_P - 1.917 * \lambda = 0.00119 - 1.917 * .0232 = -0.0433$$

and hence the VaR is

$$\text{VaR}(.05) = -(-0.0433) * (100 * 200 + 100 * 125) = 1407$$

so \$1407 is the VaR(.05).

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

(a) The expected return of the tangency portfolio is

$$\mu_T = .6 * .04 + .4 * .06 = .048$$

(b) The standard deviation is given by

$$\sigma_T = \sqrt{.6^2 * .10^2 + .4^2 * .18^2 + 2 * .6 * .4 * .5 * .10 * .18} = .114$$

(c) Now the efficient portfolio is a mixture of the tangency portfolio and the fixed-rate asset, with the weight w on the tangency portfolio being such that

$$w\sigma_T = .03$$

or

$$w = \frac{.03}{\sigma_T} = \frac{.03}{.114} = .263$$

so proportion in the risk-free asset should be .737, i.e., 73.7%. This is the only solution.

(d) To achieve an efficient portfolio with expected return being 4% (recall we changed problem to be 4% instead of the original 7%), want to again take a mixture of the tangency portfolio and the fixed-rate asset with the weight for the tangency portfolio being

$$w\mu_T + (1 - w)\mu_f = .04$$

or

$$w = \frac{.04 - .012}{.048 - .012} = .778$$

i.e., want 22.2% in the risk-free, want .778*60=46.68% in asset C, and .778*40=31.12% in asset D.

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

(a) If have a portfolio equally weighted between the 3 assets, then

$$R_p = \frac{1}{3}R_1 + \frac{1}{3}R_2 + \frac{1}{3}R_3$$

and so

$$\begin{aligned} R_p - \mu_f &= \frac{1}{3}(\beta_1 + \beta_2 + \beta_3) * (R_M - \mu_f) + \frac{1}{3}(\epsilon_1 + \epsilon_2 + \epsilon_3) \\ &= \tilde{\beta} * (R_M - \mu_f) + \tilde{\epsilon} \end{aligned}$$

where

$$\tilde{\beta} = \frac{1}{3}(\beta_1 + \beta_2 + \beta_3) = .7$$

and

$$\tilde{\epsilon} = \frac{1}{3}(\epsilon_1 + \epsilon_2 + \epsilon_3)$$

and so the beta is .7.

(b) The variance of the excess return is given by

$$\begin{aligned} \text{Var}(R_p - \mu_f) &= \text{Var}(.7 * (R_M - \mu_f)) + \text{Var}(\tilde{\epsilon}) \\ &= .7^2 * .02 + \left(\frac{1}{3}\right)^2 * (.010 + .025 + .012) = 0.0150 \end{aligned}$$

(c) The proportion of risk of asset 1 due to market risk is

$$\frac{.7^2 * .02}{.7^2 * .02 + .01} = .495$$

so just less than 50%.