Exam I– with Solutions

Statistics 509 – Winter 2022

March 2, 2022

Instructions: Do all 6 problems – there are 8 pages, and you can use the back of the pages if you need more room. Write your answers on the exam and show your work. You may use a calculator and you may refer to your book and your notes. There are a total of 100 points. Be sure to write your name on the examination.

1. (16pts) Suppose U has the uniform distribution over interval of $(-2, 2)$, and $X \sim \text{DExp}(0, 1)$.

(a) Find the .30 quantile of X.

(b) Find the .70 quantile of $Y = U^2$.

(c) Suppose $W = U + X$ and U, X are independent. Derive the correlation between W and X. *Hint:* You can use the fact that if V is uniform (a, b), then $\text{Var}(V) = \frac{1}{12}(b-a)^2$.

(a) By work in class, we want

$$.3 = F(x_{.3}) = \frac{1}{2}e^{x_{.3}}$$

so want

$$x_{.3} = \log(2 * .3) = -0.5108$$

(b) Now want

$$.7 = P(U^2 \leq \pi_{.7}) = P(-\sqrt{\pi_{.7}} \leq U \leq \sqrt{\pi_{.7}}) = 2 * P(0 \leq U \leq \sqrt{\pi_{.7}})$$

or

$$\frac{\sqrt{\pi_{.7}} - 0}{4} = .35, \quad \text{or} \quad \pi_{.7} = (4 * .35)^2 = 1.4^2 = 1.96$$

(c)

$$\text{Cov}(W, X) = \text{Cov}(U, X) + \text{Cov}(X, X) = 0 + 2 = 2$$

and

$$\text{Var}(W) = \text{Var}(U) + \text{Var}(X) = \frac{4^2}{12} + 2 = \frac{10}{3}, \quad \text{Var}(X) = 2$$

so

$$\text{Corr}(W, X) = \frac{2}{\sqrt{\frac{10}{3} \cdot 2}} = .775$$

2. (16pts) Suppose (R_1, R_2) represent the returns of 2 stocks and

$$E(R_1) = .03,\ E(R_2) = .02,\ \mathrm{Var}(R_1) = (.05)^2,\ \mathrm{Var}(R_2) = (.03)^2,\ \mathrm{Corr}(R_1, R_2) = .4$$

(a) Determine the minimum volatility portfolio for the two assets, i.e., what proportion w should be in asset 1 and proportion $(1 - w)$ be in asset 2?

(b) Determine the mean and standard derivation of the return corresponding to a portfolio consisting of 70% in asset 1 and 30% in asset 2.

(c) Suppose there is another "different" portfolio of these two risky assets that has the same standard deviation as the portfolio in **(b)**. What can you say about whether short-selling was involved for this other portfolio? What can you say about its expected return relative to the portfolio in **(b)**? Justify your answers to both.

(a) The minimum volatility portfolio corresponds to a

$$\begin{aligned} w_* &= \frac{\sigma_2^2 - \rho * \sigma_1 \sigma_2}{\sigma_1^2 + \sigma_2^2 - 2\rho * \sigma_1 \sigma_2} \\ &= \frac{(.03)^2 - .4 * .05 * .03}{(.03)^2 + (.05)^2 - 2 * .4 * .05 * .03} = 0.136 \end{aligned}$$

(b) Now

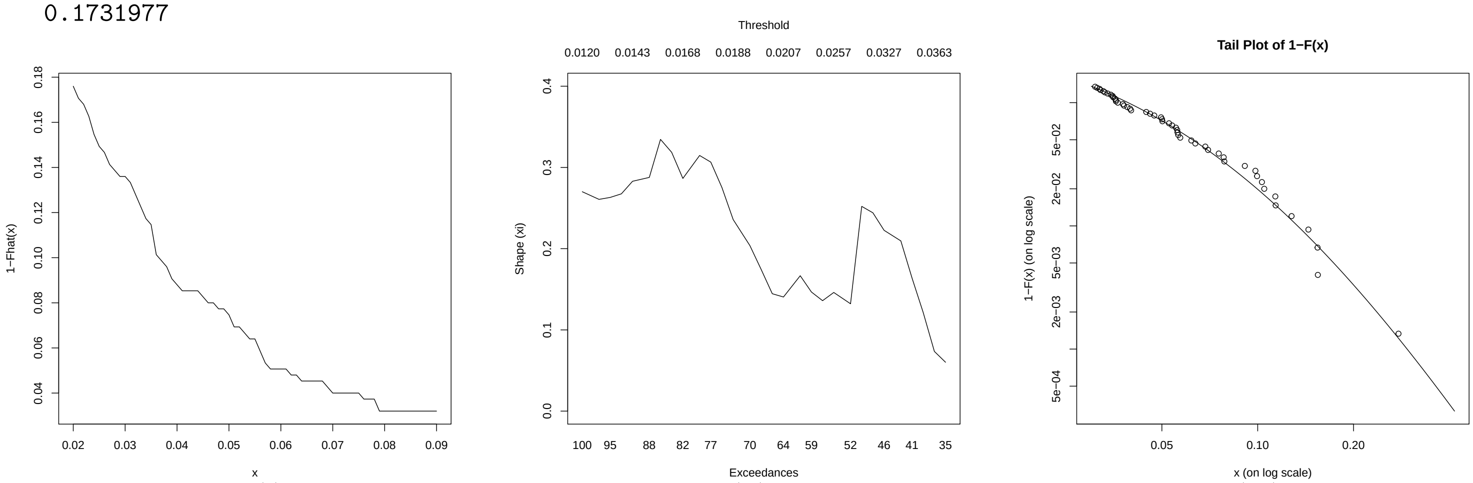
$$\begin{aligned} E(R) &= .7E(R_1) + .3E(R_2) = .7 * .03 + .3 * .02 = .027 \\ \sigma_R &= \sqrt{(.7)^2 * (.05)^2 + (.3)^2 * (.03)^2 + 2 * (.4) * (.7) * (.3) * .05 * .03} = 0.0395 \end{aligned}$$

(c) The weight for asset 1 for the portfolio in **(b)** is larger than the weight in **(a)** corresponding to the minimum variance portfolio. So for this "different" portfolio to have the same standard deviation, it corresponds to a portfolio that is on the inefficient frontier, and so must be have a lower expected return than the portfolio in **(b)** which is on the efficient frontier. Also, since it's standard deviation of .0395 is larger than the standard deviation of .03 relative to asset 2, it must be the case that the weight in front of asset 2 is bigger than 1, and this implies there is short-selling.

3. (18pts) Suppose have the following R-analysis/plots for a semi-parametric estimation of relative VaR for daily return data from Crypto 200 Index from Jan-2019 to June-2020 .

```
> library(evir)
> library(fExtremes)
> X = read.csv("..\\Data\\CMC200_Jan1_2019-Jun30_2020.csv",header=TRUE)
> Cry200_ret = exp(diff(log(X$Adj.Close)))-1
> eecdf = ecdf(-Cry200_ret)
> uv = seq(from = .02, to = .09, by = .001 )
> plot(uv,1-eecdf(uv),type='l',xlab = 'x', ylab = '1-Fhat(x)')
> shape(-Cry200_ret, models = 30, start = 35, end = 100,reverse = TRUE,
+ ci=FALSE,auto.scale='FALSE', ylim=c(0,.4),labels = TRUE)
> mu = .030
> eecdf(mu)
[1] 0.864
> gpd_est = gpd(-Cry200_ret, threshold = mu, method = c("ml"),information = c("observed"))
> tp = tailplot(gpd_est,main="Tail Plot of 1-F(x)")

> scale = gpd_est$par.ests[2]
> xi = gpd_est$par.ests[1]
> show(c(xi,scale))
        xi       beta
0.23345539 0.02876318
> alphat = 1-.005/(1-eecdf(mu))
> qgpd(alphat,xi,mu,scale)
0.1731977
```



(i) (ii) (iii)

(a) What model is being used for the tail distribution of the negative returns, and what are all the relevant parameters/estimated parameters? Based on the output from R, what is the VaR for an investment of 10 million dollars, and what is the q value being utilized?

(b) Do the plots in (i)-(iii) suggests that the semi-parametric model being used for the tail distribution is reasonable? Explain.

(c) What other plot might you want to generate, and what would it show? Be specific.

(d) Based on the model utilized for VaR in **(a)**, what is the expected shortfall given an investment of 10 million dollars? In your own words, specify what the expected shortfall is.

(a) The model being fit is that of a Generalized Pareto distribution, and the threshold value is .030, and the estimated values for ξ and σ are 0.2335 and 0.0288. The relative VaR is determined as 0.1732, and so the VaR is

$$\text{VaR} = 10^7 * .1732 = 1,732,000$$

and this corresponds to $q = .005$.

(b) The shape plot in (ii) is moderately constant, but they are somewhat high up to threshold of .018, and then the shape parameter goes down till the threshold of .0327. Now we want to be out in the tail, and we want to stay in the stable (higher region) and we are pretty good in that you are using a threshold .030 where from plot in (a), about 14% of the data is above, which is good. The shape parameter estimate as a function of threshold, as shown in (b) appears to be fairly constant till the very end. With the threshold of $\mu = .030$ and the estimated parameters, the tail probability seems to match very well as shown in the tailplot in (iii).

(c) One other plot would be to show the quantile/VaR as a function of the threshold value, as done by the quant function from Evir. This this would be used to see how stable our estimate of relative VaR is as a function of threshold and that would help us to assess (lessen or confirm) our confidence for our final estimate.

(d) Based on results in class, and the fact that we are using a generalized pareto distribution for the tail, the relative expected shortfall is given as

$$\begin{aligned}\tilde{\text{ES}} &= \tilde{\text{VaR}} + (\sigma + \xi * (\tilde{\text{VaR}} - 0.030))/(1 - \xi) \\ &= .1732 + (0.0288 + 0.2335 * (.1732 - .030))/(1 - 0.2335) = 0.2544\end{aligned}$$

so the expected shortfall is 10 million times this, and that is 2,544,000. In our own words, this is the expected loss given that our loss exceeds the threshold of the VaR, which we calculated to be 1,732,000.

4. (16pts) Suppose X, Y are two random variables corresponding to the returns of two securities, and suppose that X and Y both correspond to t-distributions with location paramters $\mu_X = \mu_Y = 0$, scale parameters $\lambda_X = .02, \lambda = .03$ and degrees of freedom parameters $\nu_X = 5, \nu_Y = 8$, and suppose (X, Y) has a Clayton copula model with $\theta = 3$, i.e., the bivariate cdf is

$$C_3(u, v) = (u^{-3} + v^{-3} - 1)^{-\frac{1}{3}}.$$

(a) Derive the probability that both securities will have returns below -5% (below -.05 at the same time). Note that you can use the following results from R.

```
> pt(-2.5,5)
[1] 0.02724505
> pt(-2,8)
[1] 0.04025812
```

(b) What is the probability in **(a)** if the securities were assumed independent?

(c) Derive the probability that both securities have returns above 5% (above .05 at the same time).

(a) The probability is given by

$$P(X \le -.05, Y \le -.05) = C_3(F_X(-.05), F_Y(-.05))$$

But with $F_t(x, \nu)$ being the cdf of the standard t with ν degrees of freedom, we have

$$F_X(-.05) = F_t\left(\frac{-.05}{.02}, 5\right) = 0.0272$$

and

$$F_Y(-.05) = F_t\left(\frac{-.05}{.025}, 8\right) = 0.0403$$

by the output from R. So

$$\begin{aligned} P(X \le -.05, Y \le -.05) &= C_3(0.0272, 0.0403) \\ &= \left[(.0272)^{-3} + (.0403)^{-3} - 1\right]^{-1/3} \\ &= 0.0249 \end{aligned}$$

(b) In the case of independence it is simply the product of the probabilities, i.e.,

$$P(X \le -.05, Y \le -.05) = (.0272) * (.0403) = 0.0011$$

(c) Now this probability is

$$\begin{aligned} P(X \ge .05, Y \ge .05) &= 1 - P(X \le .05) - P(Y \le .05) + P(X \le .05, Y \le .05) \\ &= 1 - F_X(.05) - F_Y(.05) + C_3(F_X(.05), F_Y(.05)) \\ &= 1 - .9728 - .9597 + \left[(.9728)^{-3} + (.9597)^{-3} - 1\right]^{-1/3} \\ &= 0.0040 \end{aligned}$$

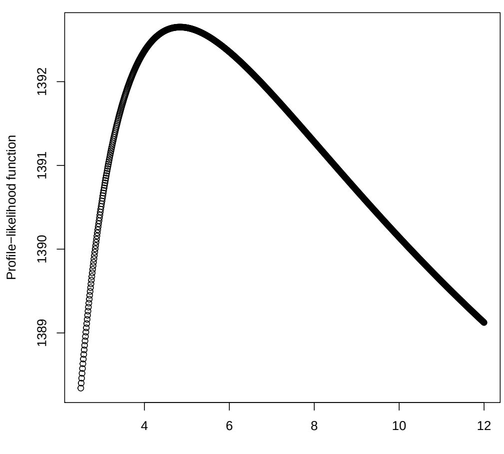
where we have used

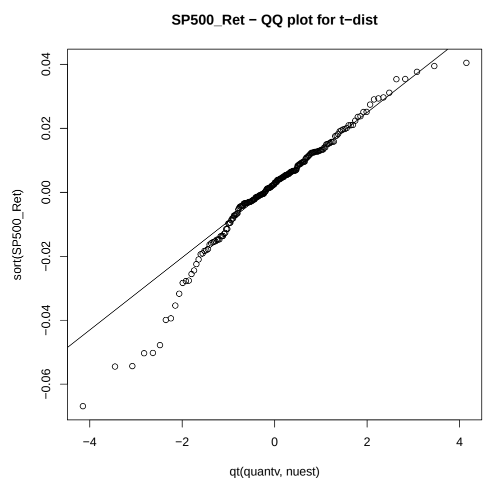
$$F_X(.05) = 1 - F_X(-.05) = .9728, \quad F_Y(.05) = 1 - F_Y(-.05) = .9597$$

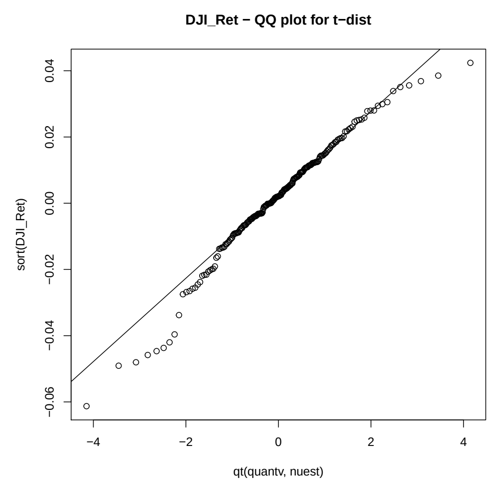
which comes from symmetry of X and Y.

5. (18pts) Suppose 2015-2018 weekly returns of the indices SP500 and DJI are being modeled as multivariate t-distribution, and the following R-code/results were generated.

```
> SP500_DIJ_Ret = cbind(SP500_Ret,DJI_Ret)
> df = seq(2.5,12,.01)
> n = length(df)
> loglik_max = rep(0,n)
> for(i in 1:n){
+   fit = cov.trob(SP500_DIJ_Ret,nu=df[i])
+   mu = as.vector(fit$center)
+   sigma =matrix(fit$cov,nrow=2)
+   loglik_max[i] = sum(log(dmt(SP500_DIJ_Ret,mean=fit$center,
+   S=fit$cov,df=df[i])))}
> plot(df,loglik_max,xlab='nu',ylab='Profile-likelihood function')
> nuest = df[which.max(loglik_max)]
> nuest
[1] 4.84
> fitfinal = cov.trob(SP500_DIJ_Ret,nu=nuest)
> fitfinal$center
  SP500_Ret     DJI_Ret
0.002401763 0.002596477
> fitfinal$cov
             SP500_Ret      DJI_Ret
SP500_Ret 0.0001664135 0.0001642724
DJI_Ret   0.0001642724 0.0001759669
> N = length(SP500_Ret)
> quantv = (1/(N+1))*seq(1,N,1)
> qqplot(qt(quantv,nuest),sort(SP500_Ret),main='SP500_Ret - QQ plot for t-dist')
> qqplot(qt(quantv,nuest),sort(DJI_Ret),main='DJI_Ret - QQ plot for t-dist')
```







(a) What are the Maximum Likelihood estimates for degrees of freedom, location parameters, scale matrix, and correlation? *Hint:* Can assume that "cov.trob" estimates the scale matrix Λ in the multivariate t.

(b) Give a thorough discussion about the fit of the multivariate t-distribution based on the QQ plots and include specifics about asymmetries in the distribution. Note that the theoretical quantiles are along the X-axis.

(c) Suppose the .95 quantile of the chi-square distribution with 1 degree of freedom is 3.841. Draw in the approximate 95% confidence intervals on the Profile Likelihood Plot (on the left), and give approximate values for lower and upper confidence bounds.

(d) Assuming the estimated multivariate t-distribution, derive the relative VaR for a portfolio evenly split between the two assets at $q = 0.005$. Note the following output from R may be used.

```
> qt(.995,nuest)
[1] 4.101618
```

. **(a)** The maximum likelihood estimates for location and scale matrix is fitfinal$center and fit$final values as shown in the R-code and the correlation value is

$$\frac{0.0001642}{\sqrt{0.0001664 * 0.0001760}} = 0.959$$

.

(b) Both QQ plots show that the t distribution does not fit the returns, and in particular it looks like both distributions are assymetric and the left/lower tail of the returns seems to be particular heavier tailed than the t-distribution. Thus it is clear that the multivariate t-distribution is probably not that good of a model.

(c) The confidence bounds are given by coming down from the max in the plot by $.5 * 3.841 \approx 1.92$, and lower bound is approximately 3 and the upper bound seems to be around 9.5. (Draw in those vertical bars).

(d) Note that the location and scale parameters are given as

$$\mu = .5 * 0.00240 + .5 * 0.00260 = 0.0025$$

and

$$\sigma = \sqrt{(.5)^2 * 0.000166 + (.5)^2 * 0.000176 + 2 * .5 * .5 * 0.000164} = 0.0129$$

where we have invoked the entries in the scale matrix. Thus

$$\tilde{\text{VaR}} = -(0.0025 - 0.0129 * 4.101618) = 0.0504$$

6. (16pts) Suppose have 2 risky assets, A and B, with expected returns of .06 and .03, respectively, standard deviations on the returns being .08 and .05, respectively, and a correlation of .4, and suppose have risk-free asset with fixed return of .01.

(a) Derive the tangency portfolio for the risky assets, and calculate the mean and standard deviation of that portfolio. Also, specify what the tangency portfolio is optimizing.

(b) Derive the optimal portfolio between all 3 assets having standard deviation on its return being at most .04. Also, specify why this portfolio is optimal.

(c) Derive the optimal portfolio between the 3 assets achieving an expected return of .03. Also, specify why this portfolio is optimal.

(a) Now $V_1 = .06 - .01 = .05,\ \ V_2 = .03 - .01 = .02$, and thus

$$\begin{aligned} w_T &= \frac{V_1\sigma_2^2 - V_2 * \rho * \sigma_1\sigma_2}{V_1\sigma_2^2 + V_2\sigma_1^2 - (V_1 + V_2)\rho * \sigma_1\sigma_2} \\ &= \frac{.05 * (.05)^2 - .02 * .4 * .08 * .05}{.05 * (.05)^2 + .02 * (.08)^2 - (.05 + .02) * .4 * .08 * .05} = .660 \end{aligned}$$

The tangent portfolio splits .660 to asset A and .340 to asset B, and this asset is maximizing the so-called Sharpe ratio, which is the ratio of the difference in expected return minus risk-free return relative to the standard deviation. The mean and standard deviation of this portfolio is given as

$$\mu_T = .66 * .06 + .34 * .03 = .0498$$

$$\sigma_T = \sqrt{(.66)^2(.08)^2 + (.34)^2\,(.05)^2 + 2 * .66 * .34 * .4 * .08 * .05} = 0.0616$$

(b) The optimal portfolio is to mix the tangent portfolio with the risk-free asset with the proportion w in the tangent portfolio being

$$w = \frac{.04}{0.0616} = .649$$

so want .351 in fixed asset and $.66 * .649$ in asset A and $.34 * .649$ in asset B. This is optimal since it has maximum expected return of all portfolios with volatility/standard deviation being .04.

(c) Now for this, want

$$w = \frac{.03 - .01}{.0498 - .01} = .503$$

so want .447 in fixed asset and $.66 * .503$ in asset A and $.34 * .503$ in asset B. This is optimal since it has minimum volatility of all portfolios with an expected return being .03.