I have read and understand this front page entirely, including the academic integrity section:

First name: E **Last name:** J

BU ID:

BOSTON UNIVERSITY QUESTROM SCHOOL OF BUSINESS
MF 840 - Spring 2021
Eric Jacquier

First Partial Exam
Thursday, March 11^{th}

- **Write your name at the top of every page as LAST, First. Every page**
- We must be in the zoom room, with video on, visible from 8:15am to 10:15am or until you have uploaded your exam. Otherwise, points will be taken off.
- The exam is individual, you can not communicate with anyone during the exam even if you have finished early, in or outside the zoom room. Any communication with anyone will be considered cheating.
- The exam is from 8:15am to 10:15am, plus 15 minutes to upload your answers on Gradescope. The exam is 104 points.
- Upload your answers as you did for MF793.
- Be neat and show your work: Answers without work or motivation get no credit. Numerical answers without first writing the formula used get no credit. Wrong final answers with correct initial work get partial credit.
- Be concise: Incorrect statements cost points even if you also write the correct answer next to them. Correct statements unrelated to the question get no credit.
- Answers to a question in the space of another question get zero credit even if correct.
- **Write your name at the top of every page as LAST, First. Every page**

ACADEMIC INTEGRITY
you understand that you:

- can not communicate with anyone, in or out of the room, in person or with a device during the exam.
- can not leave the zoom room or turn video off during the exam.
- must report cheating if you witness it.
- Breaking these rules may mean a zero at the exam and disciplinary proceedings.

Good Luck!

Problem 1: 4 quick questions. 6 points each. Spend no more than 10 minutes on each question. For True/False, indicate False if any part of the statement is false. Explain clearly what is true and false and why.

a) Consider a sample of T returns, mean μ, a TxT correlation matrix C, and variance σ^2. Write the sample mean $\widehat{\mu}$ as $i'R/T$, where i is a vector of ones and R is the vector of data. Now write the variance of $\widehat{\mu}$ as a function of C, σ, T, i.

$$Var(\hat{\mu}) = \frac{1}{T^2} Var(i'R) = \frac{1}{T^2} i' V(R) i = \frac{\sigma^2}{T^2} i'Ci$$

Note that for $C = I$, we have σ^2/T

Assume that R only as a non-zero first order autocorrelation ρ, all other autocorrelations are 0. Write the matrix C, then rewrite the variance $\widehat{\mu}$ as a function of σ, T, ρ.

$$C = \begin{pmatrix} 1 & \rho & & 0 \\ \rho & \ddots & \ddots & \\ & \ddots & \ddots & \rho \\ 0 & & \rho & 1 \end{pmatrix} \qquad i'Ci = (1 \cdots 1)\begin{pmatrix} 1+\rho \\ 1+2\rho \\ \vdots \\ 1+2\rho \\ 1+\rho \end{pmatrix} = T + 2\rho(T-2) + 2\rho$$

$$= T + \rho 2(T-1)$$

$$Var\,\hat{\mu} = \frac{\sigma^2}{T}\underbrace{\left[1 + 2\rho\left(1-\frac{1}{T}\right)\right]}_{\text{variance inflation due to } (>0) \text{ autocorrelation}}$$

b) Your estimate $\widehat{\beta}$ is 0.8 with a standard deviation of 0.15. You test the null $H_0 : \beta = 1$ at the 5% level. What is the power of your test under the alternative hypothesis $H_1 : \beta = 0.8$. Compute the left and the right power separately before adding them.

- Under H_0: $\frac{\hat{\beta} - 1}{.15} \sim t \simeq N$ Cutoffs are

$1 \mp .15 \times 1.96 \quad [.706, 1.294]$

- Power

Left $Pr[N(0.8, .15) < .706] = .265$

Right $Pr[N(0.8, .15) > 1.294] = 0.0005$

Power at $\beta = 0.8$: $\boxed{0.265}$

c) There is no real difference between the Bonferroni adjustment and the one that assumes the tests are uncorrelated with one another because for one we take the power of 1/m and for the other we divide by m. Since α is small, this comes down to about the same thing. True or False

False
Despite the small diffence in α's, there is a fundamental difference in interpretation.
Because Bonf. makes no assumption on the correlation across tests, it can not say that the adjustment results in a given type I prob.
It can only say that prob(Type I) $\leq \alpha$

d) What fundamental insight practitioners miss when running a cross-sectional regression of returns only on firm characteristics? What right hand side variable should they add to the regression and why?

Say they run: $R_i = \gamma_0 + \gamma_2 X_i + Noise_i$. They don't know whether γ_2 represents abnomal return or return for risk. If they add a risk measure:

$$R_i = \gamma_0 + \gamma_1 B_i + \gamma_2 X_i + Noise_i$$

, now they can interpret $\gamma_2 X_i$ as abnormal excess return due to X_i.

... And why is this type of cross-sectional regression called *predictive* since it is a .. cross-sectional regression?

The variables B_i (and X_i) are available or estimated 1 period before the returns R_i are measured.

Problem 2: 12 points Log-returns are $r \sim N(\mu, \sigma)$. For a sample of T observations you estimate μ and σ by MLE. Now your boss wants an estimate of the mean of the actual returns $R = e^r - 1$, call it ER (as in Expected Return). You know that $ER = e^{\mu+0.5\sigma^2} - 1$. You also need to estimate the variance of $\widehat{ER}$
Given $\widehat{\mu}, \widehat{\sigma}$, what is the MLE estimate of ER? **(2 points)**

Plug in approach, $\widehat{f(\theta)} = f(\hat{\theta})$

$$\widehat{ER} = \exp\{\hat{\mu} + \frac{1}{2}\hat{\sigma}^2\} - 1$$

ER is a bivariate function of μ, σ. Write the (two by one vector of) first partial derivatives of ER. **(4points)**

$$\frac{\partial ER}{\partial \mu} = \exp\{\mu + \frac{1}{2}\sigma^2\}$$

$$\frac{\partial ER}{\partial \sigma} = \sigma \exp\{\mu + \frac{1}{2}\sigma^2\}$$

You know the covariance matrix of $\widehat{\mu}, \widehat{\sigma}$. Now use the usual approximation method to write an approximate formula for the variance of $\widehat{ER}$ as a function of $\widehat{\mu}, \widehat{\sigma}$. **(4 points)**

$Cov(\hat{\mu}, \hat{\sigma}) = \begin{pmatrix} \sigma^2/T & 0 \\ 0 & \sigma^2/2T \end{pmatrix}$. By Delta method, we have:

$$V(\widehat{ER}) \simeq e^{\mu+\frac{1}{2}\sigma^2} (1 \ \ \sigma) \begin{pmatrix} \sigma^2/T & 0 \\ 0 & \sigma^2/2T \end{pmatrix} \begin{pmatrix} 1 \\ \sigma \end{pmatrix} e^{\mu+\frac{1}{2}\sigma^2}$$

$$\simeq e^{2\mu+\sigma^2} (1 \ \ \sigma) \begin{pmatrix} \sigma^2/T \\ \sigma^3/T \end{pmatrix} = e^{2\mu+\sigma^2} \frac{\sigma^2+\sigma^4}{T}$$

Application: $\mu = 0.1, \sigma = 0.2, T = 60$. Compute the standard deviation of $\widehat{\mu}$ and of $\widehat{ER}$. **(2 points)**

$$s(\hat{\mu}) = \frac{.2}{\sqrt{60}} = 0.0258$$

$$s(\widehat{ER}) = e^{\mu+\frac{1}{2}\hat{\sigma}^2} \hat{\sigma} \sqrt{\frac{1+\sigma^2}{T}} = 0.0297$$

Problem 3: 20 points

In the regression $Y = X\beta + \epsilon, \epsilon \sim N(0, \Omega)$, Ω is a **known** covariance matrix. The sample size is T.

a) 4 points Write (no proof) the joint density of the error vector $p(\epsilon)$

$$p(\varepsilon) = \frac{1}{|\Omega|^{1/2}} \exp\left\{ - \frac{\varepsilon' \Omega^{-1} \varepsilon}{2} \right\}$$

b) 4 points Given $p(\epsilon)$, write the joint density of the data $p(Y)$. Write exactly the known result you use to go from ϵ to Y, and show how you apply it.

$$p(Y) = \left|\frac{\partial Y}{\partial \varepsilon}\right| p(\varepsilon(Y)) = |I| \frac{1}{|\Omega|^{1/2}} \exp\left\{ - \frac{(Y - XB)' \Omega^{-1} (Y - XB)}{2} \right\}$$

c) 6 points Find $\widehat{\beta}_{MLE}$

$$\text{Log } \ell \propto \frac{-1}{2} (Y - XB)' \Omega^{-1} (Y - XB)$$

$$\frac{d \text{ Log } \ell}{\partial B} = \frac{d}{\partial B}\left[- \frac{Y'\Omega^{-1} Y - 2B'X'\Omega^{-1} Y + B'X'\Omega^{-1} X B}{2} \right]$$

$$= -\frac{1}{2}(-2X'\Omega^{-1} Y + 2X'\Omega^{-1} X B) = 0$$

$$\hat{B}_{MLE} = (X'\Omega^{-1} X)^{-1} X' \Omega^{-1} Y$$

d) 6 points Use the Cramer-Rao result to get the asymptotic covariance matrix of $\hat{\beta}_{MLE}$.

$$\frac{d^2 \text{ Log } \ell}{\partial B} = -\frac{1}{2} 2X'\Omega^{-1} X = -X'\Omega^{-1} X$$

$$V(\hat{B}) = -\left[E \frac{\partial^2 \text{ Log } \ell}{d B}\right]^{-1} = (X'\Omega^{-1} X)^{-1}$$

Problem 4: 28 points

You estimated market regressions on 12 industries, assuming i.i.d normal noise for each. You are testing the null hypothesis $H_0 : \beta = 1$ at the 5% level (two-sided). The sample size is 1 years of daily returns. $T=252$

a) 6 points You are willing to assume that the covariance matrix of the 12 regression residuals is diagonal. If you use the standard two-sided 5% Student-t cutoff, what is the effective significance level of your *"data mining"* procedure, that is, what is the probability of at least one rejection out of the 12.

$p(\text{at least } 1) = 1-(1-\alpha)^{12} = 1-(.95)^{12}$

$\simeq .46$

b) 6 points Then what significance level and what associated cutoff should you use to flag each t-stat to correct for this data-mining bias, (so your overall test is size 5%)?

Use α such that $1-(1-\alpha)^{12} = .05$ $\quad \alpha = 0.0043$

$t_{\nu = 252-2=250}\left(\alpha = \frac{0.0043}{2}\right) : \mp 2.88$

c) 6 points You now apply the Bonferroni correction for data mining: What α and cutoff do you use now for each test?

$\alpha = \frac{.05}{12} = 0.00416$

$t : \mp 2.89$

d) 6 points You now apply the Hotelling (exact) correction for data mining: What α and cutoff do you use now for each test?

$\sqrt{\frac{12 \times 251}{240} F_{12,240}(.95)} \simeq 4.74$

d) 4 points You boss is shocked how drastic the correction is compared to the others. Briefly explain to him why this is so.

The Hotelling assumes the worst possible case of data mining where the squared linear combination $(\ell'\bar{x})^2$ which is a t^2 for a given value of ℓ, is maximized. It can be shown that $\max_{\ell}(\ell'\bar{x})^2$ is a Hotelling T^2.

Problem 5: 20 points

You estimated market regressions on several industries, assuming (it's OK) i.i.d normally distributed noise for each. The sample size is 5 years of monthly returns. You analyzed the regression residuals and concluded they were about uncorrelated with one another.

Estimates	Dur.	Chems	Finance
$\hat{\beta}$	1.20	0.94	1.12
$s(\hat{\beta})$	0.12	0.08	0.10

You want to test the null that these three βs are equal with a standard Wald test formulation.

a) Write the R matrix and the r vector in the standard Wald constraint $R\theta - r = 0$. **6 points**

b) Write the covariance matrix of the constraint, (the matrix which is inverted in the Wald test quadratic form).**6 points**

c) Compute the Wald test. Feel free to use R to do your computation. **6 points**

d) Do you reject at 5% using the exact distribution? **2 points**

a) $R = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix} \quad r = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

b) $Cov = R\, Cov(\hat{\beta})\, R' = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} 0.12^2 & & \\ & .08^2 & \\ & & .10^2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -1 & 1 \\ 0 & -1 \end{pmatrix}$

$= \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} 0.12^2 & 0 \\ -0.08^2 & 0.08^2 \\ 0 & -.10^2 \end{pmatrix}$

$= \begin{pmatrix} .12^2 + .08^2 & -.08^2 \\ -.08^2 & .08^2 + .10^2 \end{pmatrix}$

c) $W = \begin{pmatrix} .26 & -.18 \end{pmatrix} Cov^{-1} \begin{pmatrix} .26 \\ -.18 \end{pmatrix}$

$= 3.94$

$\frac{2 \times (60-1)}{60-2} F_{.95}(2, 58) = 6.4$

Do not reject