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

First name: **Last name:**

BOSTON UNIVERSITY QUESTROM SCHOOL OF BUSINESS
MF 840 - Spring 2019
Eric Jacquier

First Partial Exam
Thursday, March 7^{th}

- **Write your name on the first page now.**
- You are allowed writing instruments, and a non-networkable no-functional-memory calculator.
- All bags, coats, phones, must be in the front of the room . We will not start until this is done.
- The exam is 2 hours long and has 106 points.
- Use only the space provided on the front of the pages. The back of the exam will not be read and will not count even if there is correct material.
- Be neat and show your work: Answers without work or motivation receive no credit. Wrong final answers with partially correct work receive partial credit.
- Be concise: Incorrect statements cost points even if you also write the correct answer. Correct statements unrelated to the question get no credit.

ACADEMIC INTEGRITY: I understand that:
- I can not communicate with anyone, in or out of the room. Any communication is considered cheating.
- I can not consult any documentation or use any networkable devices, in or out of the room.
- I can not leave the room unless allowed to and accompanied.
- If I witness cheating, I will report it otherwise I am complicit.
- I can not leave the room with my exam under any circumstance.
- Breaking these rules will cause a disciplinary hearing, a possible 0 at the exam, F at the course or expulsion.

The following tables will help

Cut-off points for N(0,1) and Student-t(35)

p(z < x)	0.001	0.0025	0.005	0.00625	0.01	0.025
z ~ t(35): x=	-3.34	-3.00	-2.72	-2.63	-2.44	-2.03
z ~ N(0,1): x=	-3.09	-2.81	-2.58	-2.50	-2.33	-1.96

95^{th} percentile of the χ^2 distribution

ν	1	2	3	4	5	10	20
$\chi^2_{0.95}$	3.84	6	7.8	9.5	11.1	18	31

95^{th} percentile of F(ν_1, 36-ν_1) distribution

ν_1	1	2	4	5	10
$F_{0.95}$	4.12	3.28	2.67	2.52	2.22

Problem 1: 4 quick questions. 8 points each. You should spend no more than 5 minutes on each question.

a) Write the exact distribution of the Hotelling T^2 for a m dimensional vector of normal sample means. Prove that it is asymptotically a $\chi^2(m)$.

$$T^2 \sim \underbrace{\frac{(T-1)}{(T-m)}}_{\to 1} \underbrace{F(m, T-m) \times m}_{\frac{m\chi^2(m)/m}{\chi^2(T-m)/T-m} \to \frac{\chi^2(m)}{1}}$$

As $T \to \infty$

$$E\left(\frac{\chi^2(T-m)}{T-m}\right) = \frac{T-m}{T-m} = 1$$

$$V\left(\frac{\chi^2(T-m)}{T-m}\right) = \frac{2(T-m)}{(T-m)^2} = \frac{1}{(T-m)} \to 0$$

Explain the logic behind using the square-root of (a 1- α % point on) that distribution to help correct for data mining biases.

- Consider finding combinations $\ell'\bar{x}$ to maximize the resulting t^2. It can be shown that the maximum is equal to $(\bar{X}-\mu)'\left(\frac{S}{N}\right)^{-1}(\bar{X}-\mu)$, the hotelling T^2.
- So using the square root of its cutoff value accounts for all possible data mining over the components of $\bar{X}$.

b) You use $\alpha = 0.05$ as size to test $\beta = 1$. The sample size is very large, so the estimate is very precise. You plotted the power curve for $\beta \in [0.8, 1.2]$: The power at β= 1.025 is 0.99.
What is the probability of type II error at $\beta = 1.025$, ..., and for β larger than 1.025?

P(Type II error) = 1 − Power = 0.01 at $\beta = 1.025$
The power is higher and P(type II error) is lower for βs further away from 1, such as > 1.025

You have no preference between the two types of errors. And you see no practical difference between βs of 1 or 1.025. How would you modify your test and how would it affect the type I and type II errors?

We chose P(type I) = .05 It results in P(type II) < .01 for values of β meaningfully different from 1.

By choosing <u>a lower</u> α, we will <u>increase</u> a bit P(type II) error so they will be more balanced

c) $r_t = Log(1+R_t) \sim i.i.d.N(\mu, \sigma)$. You know $\sigma = 0.2$. You estimate $\mu = 0.1$ per year with T= 50 years of data. Compute the expected total wealth in 20 years V_{20} for \$ 1 invested today given by the arithmetic, geometric, unbiased, and *most precised* methods. (4 numbers, \$ answers).

$$A = e^{20(.1+\frac{1}{2}.2^2)} = \$ \quad 11.02$$

$$G = e^{20(.1)} = \$7.4$$

$$U = e^{20(.1+\frac{1}{2}(1-\frac{20}{50}).2^2)} = \$9.4$$

$$M = e^{20(.1+\frac{1}{2}(1-\frac{20}{50}\times 3).2^2)} = \$6.8$$

d) You realize now that r_t above is autocorrelated, still constant variance σ^2. The standard deviation of $\widehat{\mu}$ is not $\sigma/\sqrt{T}$ anymore! Write the correct standard deviation of $\widehat{\mu}$ as a function of C, the TxT correlation matrix, σ and T. Then write C when the autocorrelation is an MA(1) with parameter θ, and write the corrected standard deviation of $\widehat{\mu}$ as a function of T, σ, θ.

- $\hat{\mu} = \frac{1}{T}\sum r_t, \quad r_t = \mu + \sigma\varepsilon_t \quad \varepsilon_t \sim N(0,1) \quad \varepsilon \sim N(0, C)$

$$= \frac{1}{T}(\sum \mu + \sigma\varepsilon_t) = \mu + \frac{\sigma}{T}\sum\varepsilon_t \quad \sum\varepsilon_t = \iota'\varepsilon \quad V(\iota'\varepsilon) = \iota' C \iota$$

$$V(\hat{\mu}) = \frac{\sigma^2}{T^2}\iota' C \iota \qquad \boxed{\sigma_{\hat{\mu}} = \frac{\sigma}{T}\sqrt{\iota' C \iota}}$$

- $\text{MA}(1): \quad C = \begin{pmatrix} 1 & \theta & & 0 \\ \theta & \ddots & \ddots & \\ & \ddots & \ddots & \theta \\ 0 & & \theta & 1 \end{pmatrix} \qquad \iota' C \iota = T + 2(T-1)\theta$

$$\boxed{\sigma_{\hat{\mu}} = \frac{\sigma}{T}\sqrt{T+2(T-1)\theta}} = \frac{\sigma}{\sqrt{T}}\sqrt{1+2(1-\frac{1}{T})\theta}$$

Problem 2: 12 points

Consider the multivariate normal **m**-dimensional random vector $R \sim N(\mu, \Sigma)$.
Write the joint multivariate normal density of R. Then prove that the quantity in the exponential is a $\chi^2(m)$.
You can use the result that a sum of m squared i.i.d N(0,1)s is $\chi^2(m)$

$$p(R) = \frac{1}{|\Sigma|^{1/2}} \exp\left\{-\frac{1}{2}(R-\mu)'\Sigma^{-1}(R-\mu)\right\}$$

We know that $\Sigma = P^* P^{*\prime}$ and $\Sigma^{-1} = P^{*\prime -1} P^{*-1}$

Then $$(R-\mu)'\Sigma^{-1}(R-\mu) = \underbrace{(R-\mu)'P^{*\prime -1}}\ \underbrace{P^{*-1}(R-\mu)}$$
$$= (Y-\mu)'(Y-\mu) \qquad [1]$$

and $Y-\mu \sim N(0, I)$

because $$E[(Y-\mu)(Y-\mu)'] = P^{*-1}\underline{E(R-\mu)(R-\mu)'}P^{*\prime -1}$$
$$= P^{*-1}\Sigma P^{*\prime -1}$$
$$= P^{*-1}P^* P^{*\prime}P^{*\prime -1} = I$$

So $(Y-\mu)'(Y-\mu)$ is a sum of m iid $N(0,1)$ normals squared and is $\chi^2(m)$

So $[1] \sim \chi^2(m)$

Problem 3: 26 points Consider the regression $Y = X\beta + \epsilon$, with $\epsilon \sim i.i.dN(0, vI)$, where v is the variance. You collect T observations of X, Y.

a) 2 points Write the joint density of the error vector $p(\epsilon|v)$

$$p(\varepsilon) = \frac{1}{\sqrt{2\pi}^T} \frac{1}{v^{T/2}} \exp\left\{-\frac{\varepsilon'\varepsilon}{2v}\right\}$$

b) 6 points Now write (prove) the joint density of the data $p(Y|v, \beta)$.

$$p(Y|v,\beta) = p(\varepsilon(Y)) \left|\frac{d\varepsilon}{dY}\right| \qquad \varepsilon = Y - X\beta \qquad \frac{d\varepsilon}{dY} = I$$

$$p(Y|v,\beta) \propto \frac{1}{v^{T/2}} \exp\left\{-\frac{(Y-X\beta)'(Y-X\beta)}{2v}\right\}$$

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

$$\operatorname{Log} \ell \propto -\frac{T}{2} \operatorname{Log} v - \frac{1}{2v}(Y-X\beta)'(Y-X\beta)$$

$$\frac{\partial \operatorname{Log} \ell}{\partial \beta} = -\frac{1}{2v}\frac{\partial}{\partial \beta}\left(Y'Y - 2\beta'X'Y + \beta'X'X\beta\right)$$

$$= -\frac{1}{2v}\left(-2X'Y + 2X'X\beta\right) = 0 \Rightarrow \boxed{\hat{\beta}_{MLE} = (X'X)^{-1}X'Y}$$

$$\frac{\partial \operatorname{Log} \ell}{\partial v} = -\frac{T}{2v} + \frac{1}{2v^2}(Y-X\beta)'(Y-X\beta)$$

$$\boxed{\hat{v}_{MLE} = \frac{1}{T}(Y-X\hat{\beta})'(Y-X\hat{\beta})}$$

d) 6 points Write the general result for the asymptotic covariance matrix of an MLE estimator $\widehat{\theta}$ of a parameter vector θ. Then compute the variance of $\hat{v}_{MLE}$. Assume without proof that $\frac{\partial \ell}{\partial \beta \partial v} = 0$ to save time. But explain why it helps !

- Asymptotically $V(\hat{\theta}) = -\left(E\left[\frac{\partial^2 \log L}{\partial \theta \partial \theta'}\right]\right)^{-1} = I(\theta)^{-1}$

- $E\left[\frac{\partial^2 \log \ell}{\partial v^2}\right] = E\frac{\partial}{\partial v}\left[-\frac{T}{2v} + \frac{1}{2v^2}\varepsilon'\varepsilon\right] = E\left[\frac{T}{2v^2} - \frac{2}{2v^3}\varepsilon'\varepsilon\right]$

$$= \frac{T}{2v^2} - \frac{2}{2v^3}E(\varepsilon'\varepsilon) = \frac{T}{2v^2} - \frac{Tv}{v^3} = -\frac{T}{2v^2}$$

$$V(\hat{\theta}) = \begin{bmatrix} \boxed{*} & 0 \\ 0 & +\frac{T}{2v^2} \end{bmatrix}^{-1}$$ Since $I(\hat{\theta})$ is block diagonal we know that the inverse will be

and $V(\hat{v}_{MLE}) = \boxed{\frac{2v^2}{T}}$

e) 6 points Write the general formula for the Δ-method. Use it to prove the variance of the MLE estimator of precision, h = 1 / v.

- $V(f(\hat{\theta})) \doteq \frac{\partial f}{\partial \theta}' V(\hat{\theta}) \frac{\partial f}{\partial \theta}$

$h = \frac{1}{v} \qquad \frac{dh}{dv} = -\frac{1}{v^2}$

$V(\hat{h}) \doteq \left(-\frac{1}{v^2}\right)\frac{2v^2}{T}\left(-\frac{1}{v^2}\right) = \boxed{\frac{2h^2}{T}}$

Problem 4: 22 points

You estimated market regressions on several industries, assuming (it's OK) i.i.d normally distributed noise for each. The goal is to test the null hypothesis $H_0 : \beta = 1$ at the 5% level (two-sided). The sample size is 3 years of monthly returns.

Estimates	Dur.	Chems	Steel	Finance
$\hat{\beta}$	1.48	1.39	1.72	1.38
$s(\hat{\beta})$	0.12	0.08	0.12	0.08
t-statistic	4*	4.88*	6*	4.75*

a) 4points Fill the table with the t-statistics for $H_0 : \beta = 1$. Star the t-stats significant at 5% level.

b) 6 points The 4 regressions residuals are uncorrelated. Under the null that all βs equal 1, what is the probability of finding at least one significant t-stat in question a)?
What size and cut-off should you use on each t-stat to correct for the data-mining (so the test is size 5% overall)?

- $P(\text{at least 1 reject}) = 1 - P(\text{no reject}) \underset{\text{independence}}{=} 1 - \prod(1-\alpha) = 1-(1-\alpha)^4$

$= \boxed{0.185}$

- Pick α so: $(1-\alpha)^4 = 1 - .05$ $\quad \boxed{\alpha = 0.012}$

Pick $|t_{0.012/2}|$ Not exactly in table $|t_{0.006}| = \boxed{2.63}$

c) 6 points You now apply the Bonferroni correction for data mining: What α and cutoff do you use now for each test? How many t-stats are significant? What can you say about the size of the Bonferroni test?

For Data mining through m tests, choose $\frac{0.05}{m} = \frac{0.05}{4} = \boxed{0.0125}$

Again not exactly in Table ≈ 2.63

Still all significant

d) 6 points You also consider the Hotelling correction. What are the exact and the asymptotic cutoffs (numbers) to use? How many statistics are significant now?

$\sqrt{\frac{35}{32} 4 F_{4,32}(.95)} = \sqrt{\frac{35}{32} 4 \times 2.67} = 3.41$

$\sqrt{\chi^2_4(.95)} = 3.08$

All significant

Problem 5: 16 points
You want to test whether small (S) and large (L) firms have the same βs. You estimated regressions of small and large firm portfolios return on the market return. You worried about potential SUR effects and computed the covariance matrix of the residuals. You conclude that they are uncorrelated. Your estimates are $\beta_S =$ 1.06 with standard deviation 0.06, and $\beta_L =$0.96 with standard deviation 0.04.

a) 8 points Prove that for **linear** restrictions $R\theta - r = 0$ the Wald test becomes a simple quadratic form. R represents m restrictions, θ is p-dimensional with p>m. The parameter estimates are $\hat{\theta} \sim N(\theta, V(\hat{\theta}))$. Clearly indicate the vector and covariance matrix in the final quadratic form and their dimension.

$W = f(\theta)' V(f(\theta))^{-1} f(\theta)$ Computed at unrestricted MLE

$f(\theta) = R\theta - r$

$W = (R\hat{\theta} - r)' V(R\hat{\theta})^{-1} (R\hat{\theta} - r)$

The vector of normals is $R\hat{\theta} \sim N(R\theta, RV(\hat{\theta})R')$

It is of dimension m (# of restrictions, rows in R)

$\underset{m\times p}{R}\,\underset{p\times p}{V(\hat{\theta})}\,\underset{p\times m}{R'}$ is an $m \times m$ covariance matrix

b) 8 points Now compute the Wald Test of equality of the βs. Compute the covariance matrix $V(\hat{\beta})$. Write the R restriction matrix (?) and indicate its dimension. Compute the variance of $R\hat{\beta}$. Compute the Wald Test. its approximate cut-off point. Do you accept of reject the null?

- Only 1 restriction $R = (1, -1)$ $r = 0$

- $V(\hat{\beta}) = \begin{pmatrix} .06^2 & 0 \\ 0 & .04^2 \end{pmatrix}$ • $RV(\hat{\beta})R' = (1 \ \ -1)\begin{pmatrix} .06^2 & 0 \\ 0 & .04^2 \end{pmatrix}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$

$= \boxed{.06^2 + .04^2}$

- $R\hat{\beta} = 1.06 - .96 = .1$

- $W = .1^2/(.06^2 + .04^2) = \boxed{1.92} \sim \chi^2(1)$

$\chi^2_1(.95) = 3.84$ Do not reject