Comprehensive Examination
Quantitative Methods
Spring, 2019

Instruction: This exam consists of three parts. You are required to answer all the questions in all the parts.

Grading policy:

1. Each part will be graded separately, and there are four possible results in each part: H (honor pass), P (PhD pass), M (master pass), and F (fail).

2. Each part contains a precise grade determining algorithm.

3. The grades from the three parts will be summarized in the descending order, after which the overall grade will be determined using the algorithm summarized in the table below:

Highest	Middle	Lowest	**Overall**
H	H	H	**H**
H	H	P	**H**
H	H	M	**P**
H	H	F	**M**
H	P	P	**P**
H	P	M	**P**
H	P	F	**M**
H	M	M	**M**
H	M	F	**M**
H	F	F	**F**
P	P	P	**P**
P	P	M	**P**
P	P	F	**M**
P	M	M	**M**
P	M	F	**M**
P	F	F	**F**
M	M	M	**M**
M	M	F	**F**
M	F	F	**F**
F	F	F	**F**

Part I - 203A

Instructions for Part I: Solve every question. For every question in this part, your answer should be numerical; an abstract formula will not be accepted as an answer. Also, you are required to be very specific about the dimension of the a zero vector or zero matrix. If you intend to write a 1×3 zero vector, you should write it $\begin{bmatrix} 0 & 0 & 0 \end{bmatrix}$ or $\underset{1\times 3}{0}$. If you simply write "0", it shall be understood to be a scalar.

Grading policy for Part I: Your grade in this part of the exam is based on the total points that you earn. Below is how the grade is determined. In order to avoid any confusion arising from the difference between strict and weak inequalities, the grade assignment algorithm is presented in terms of mathematical inequalities. Note that $<$ denotes a strict inequality, and $\leq$ denotes a weak inequality. Let T denote the total number of points.

1. If $T \geq 25$, you will get H.
2. If $20 \leq T < 25$, you will get P.
3. If $15 \leq T < 20$, you will get M.
4. If $T < 15$, you will get F.

Question 1 (5 pts.) *No derivation is required for this question; your derivation will not be read anyway.* Let $Z_1, Z_2, \ldots$ denote an i.i.d. sequence of uniform$(-1, 1)$ random variables, i.e., their common PDF is given by $f(z) = \frac{1}{2} 1(|z| \leq 1)$. Let

$$X_i = \begin{bmatrix} Z_i \\ Z_i^3 \end{bmatrix}.$$

What is the asymptotic distribution of $n^{-1/2} \sum_{i=1}^n X_i$ as $n \to \infty$?

Question 2 (4 pts.) *No derivation is required for this question; your derivation will not be read anyway.* Suppose that $X \sim N(0, 1)$, and let

$$f_n(x) = x^2 + \frac{1}{n} \sin(x).$$

What is $\lim_{n\to\infty} E[f_n(X)]$?

Question 3 *No derivation is required for the questions below; your derivation will not be read anyway.* Suppose that

$$Y = X + \varepsilon$$

where X and ε are independent random variables such that $\varepsilon \sim N(0, 1)$, and X has a $\Gamma(4, 2)$ distribution. Answer the two sub-questions below. You may want to recall that if Z has a $\Gamma(\alpha, \beta)$ distribution, then $E\left[e^{tZ}\right] = \left(\frac{1}{1 - \beta t}\right)^{\alpha}$ for $t < \frac{1}{\beta}$.

(a) (3 pts.) What is $E\left[Y^2 \middle| X\right]$?

(b) (3 pts.) What is $E\left[Y^2\right]$?

Question 4 (3 pts.) *No derivation is required for the question below; your derivation will not be read anyway.* Suppose that X has the MGF equal to

$$M(t) = E\left[\exp(tX)\right] = \exp\left(\frac{t^2}{2}\right).$$

Let A_n and B_n denote the events defined by

$$A_n = \left\{X \le -\frac{1}{n}\right\},$$
$$B_b = \left\{X < \frac{1}{n}\right\}.$$

Furthermore, let A and B denote the events defined by

$$A = \lim_{n\to\infty} A_n,$$
$$B = \lim_{n\to\infty} B_n.$$

What is $P(B) - P(A)$?

Question 5 (4 pts.) *No derivation is required for the question below; your derivation will not be read anyway.* Let $f(x_1, x_2) = C x_1^2 x_2^3 1(0 < x_1 < 1)\, 1(0 < x_2 < 2)$ denote the joint PDF of X_1 and X_2, where $C > 0$ is some constant that makes f a valid PDF. What is $E\left[X_1 \middle| X_2\right]$?

Question 6 (4 pts.) *No derivation is required for this question; your derivation will not be read anyway.* Suppose that

$$\begin{bmatrix} X \\ Y \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right),$$

and

$$E\left[Y^2 - \left(\alpha + \beta X^2\right)\right] = 0,$$
$$E\left[X^2\left(Y^2 - \left(\alpha + \beta X^2\right)\right)\right] = 0.$$

What are α and β?

Question 7 *No derivation or explanation is required for the questions below; your derivation or explanation will not be read anyway.* Suppose that the support of Y is $\{0, 1, 2\}$ and

$$\Pr\left[Y = 0 \middle| X, Z\right] = \frac{1}{1 + \exp(\alpha + X\beta) + \exp(\gamma + Z\delta)},$$
$$\Pr\left[Y = 1 \middle| X, Z\right] = \frac{\exp(\alpha + X\beta)}{1 + \exp(\alpha + X\beta) + \exp(\gamma + Z\delta)},$$
$$\Pr\left[Y = 2 \middle| X, Z\right] = \frac{\exp(\gamma + Z\delta)}{1 + \exp(\alpha + X\beta) + \exp(\gamma + Z\delta)},$$

where X and Z are scalar random variables such that the support of (X, Z) is $\{(0,0), (0,1)\}$. We know that

$$\Pr[Y = 0| (X, Z) = (0,0)] = \frac{1}{3},$$
$$\Pr[Y = 1| (X, Z) = (0,0)] = \frac{1}{3},$$
$$\Pr[Y = 2| (X, Z) = (0,0)] = \frac{1}{3},$$

and

$$\Pr[Y = 0| (X, Z) = (0,1)] = \frac{1}{3},$$
$$\Pr[Y = 1| (X, Z) = (0,1)] = \frac{1}{3},$$
$$\Pr[Y = 2| (X, Z) = (0,1)] = \frac{1}{3}.$$

(a) (1 pt.) Is α identified? If so, what is its numerical value?

(b) (1 pt.) Is β identified? If so, what is its numerical value?

(c) (1 pt.) Is γ identified? If so, what is its numerical value?

(d) (1 pt.) Is δ identified? If so, what is its numerical value?

Part II - 203B

Instruction for Part II : Solve every question.

Grading policy for Part II: Your grade in this part of the exam is based on the total points that you earn. Below is how the grade is determined. In order to avoid any confusion arising from the difference between strict and weak inequalities, the grade assignment algorithm is presented in terms of mathematical inequalities. Note that $<$ denotes a strict inequality, and $\leq$ denotes a weak inequality. Let T denote the total number of points.

1. If $T \geq 80$, you will get H.
2. $60 \leq T < 80$, you will get P.
3. $45 \leq T < 60$, you will get M.
4. If $T < 45$, you will get F.

Question 1 (50 points) Consider a two period panel data model. Each individual i at time period $t \in \{1, 2\}$ has potential outcomes $(Y_{it}(0), Y_{it}(1))$. The treatment status of individual i at times t is given by $D_{it} \in \{0, 1\}$ with $D_{it} = 1$ indicating receipt of treatment. For each individual i we only observe $(Y_{it}, D_{it})_{t=1}^2$ with Y_{it} given by

$$Y_{it} = Y_{it}(0) + D_{it} Y_{it}(1).$$

Moreover, assume that at time period $t = 1$ no one is treated (i.e. $D_{i1} = 0$ with probability one). At time period $t = 2$, some individuals are treated while others are not (i.e. $D_{i2} \in \{0, 1\}$). Throughout the problem assume we have an i.i.d. sample $\{(Y_{i1}, Y_{i2}, D_{i1}, D_{i2})\}_{i=1}^n$ but do not assume that D_{it} is independent of $(Y_{it}(0), Y_{it}(1))$.

(a) (10 points) Provide expressions for $E[Y_{i2}|D_{i2} = 1]$ and $E[Y_{i1}|D_{i2} = 1]$.

(b) (10 points) The average treatment effect (ATT) on the treated is given by

$$\text{ATT} = E[Y_{i2}(1) - Y_{i2}(0)|D_{i2} = 1].$$

Suppose that potential outcomes are constant through time. Show that

$$\text{ATT} = E[Y_{i2}|D_{i2} = 1] - E[Y_{i1}|D_{i2} = 1].$$

(c) (10 points) Maintaining the assumptions of part (b), propose an estimator for the ATT and establish its consistency. Clearly state any assumptions you need.

(d) (10 points) Establish the asymptotic normality of the estimator for the ATT that you proposed in part (c). Clearly state any assumptions you need.

(e) (10 points) No longer assume that potential outcomes are constant through time. Instead, suppose we are willing to assume that

$$E[Y_{i2}(0) - Y_{i1}(0)|D_{i2} = 1] = E[Y_{i2}(0) - Y_{i1}(0)|D_{i2} = 0],$$

which is called a *parallel trends* assumption. Show that the ATT is identified – i.e. provide an expression for the ATT in terms of conditional moments of observed variables. (Hint: Use the parallel trends assumption to "learn" about the unobserved quantity $E[Y_{i2}(0)|D_{i2} = 1]$.)

Question 2 (50 points) Consider a simple dynamic panel data model in which

$$Y_{it} = \alpha_i + Y_{it-1}\beta_0 + \varepsilon_{it}.$$

Throughout the problem, assume we observe $Y_i \equiv (Y_{i0}, \ldots, Y_{iT})$ for each individual i and that observations $\{Y_i\}_{i=1}^n$ are i.i.d. across individuals i.

(a) (10 points) Is it credible in this model to assume $E[\varepsilon_{it}|Y_i] = 0$? Why or why not?

(b) (10 points) Suppose that $E[\alpha_i|Y_i] = 0$ and $E[\varepsilon_{it}|Y_{it-1}] = 0$ for all t. Show that

$$\beta_0 = \frac{E[Y_{it}Y_{it-1}]}{E[Y_{it-1}^2]}.$$

(c) (10 points) Maintaining the assumptions of part (b), propose an estimator for β_0 and establish its asymptotic normality. Clearly state any assumptions you need.

(d) (10 points) Maintaining the assumptions of part (b), propose an estimator for the asymptotic variance you derived in part (c). You do not need to show this variance estimator is consistent.

(e) (10 points) Drop the assumptions of part (b). Instead suppose $[\varepsilon_{it}|Y_{it-1}, \ldots, Y_{i0}] = 0$ for all $t \geq 1$. Do not assume that $E[\alpha_i|Y_i] = 0$. Propose an estimator for β_0 and establish its consistency. Clearly state any assumptions you need.

Part III - 203C

Instruction for Part III : Solve every question.

Grading policy for Part III: Your grade in this part of the exam is based on the total points that you earn. Below is how the grade is determined. In order to avoid any confusion arising from the difference between strict and weak inequalities, the grade assignment algorithm is presented in terms of mathematical inequalities. Note that $<$ denotes a strict inequality, and $\leq$ denotes a weak inequality. Let T denote the total number of points.

1. If $T \geq 80$, you will get H.
2. $60 \leq T < 80$, you will get P.
3. $45 \leq T < 60$, you will get M.
4. If $T < 45$, you will get F.

Question 1 (18 points) Let a sample X have only one observation X from the Binomial distribution $B(4, \theta)$. That is, X has the probability mass function

$$f(x, \theta) = {}_4C_x \times \theta^x \times \theta^{1-x} \text{ for } x \in \{0, 1, 2, 3, 4\}$$

where ${}_4C_x = \frac{4!}{x! \times (4-x)!}$ and $a! = a \times (a-1) \times \ldots \times 2 \times 1$ for any positive integer a. Consider the hypotheses:

$$H_0 : \theta \leq \frac{1}{2} \text{ against } H_1 : \theta > \frac{1}{2}.$$

We consider two tests for the above hypotheses. The critical region of the first test φ_1 is $\{4\}$, while the critical region of the second test φ_2 is $\{3, 4\}$.

(a) (3 points) Compute the power function of φ_1.

(b) (3 points) What is φ_1's maximal probability of making Type I error?

(c) (3 points) What is φ_1's probability of making Type II error if $\theta = \frac{2}{3}$?

(d) (3 points) Compute the power function of φ_2.

(e) (3 points) What is φ_2's maximal probability of making Type I error?

(f) (3 points) What is φ_2's probability of making Type II error if $\theta = \frac{4}{5}$?

Question 2 (20 points) Suppose that $\{X_t\}_{t \in Z}$ is a second order auto-regressive process, i.e.

$$X_t = \phi_o X_{t-1} + \frac{1}{2} X_{t-2} + u_t,$$

where $u_t \sim i.i.d.(0, \sigma_u^2)$, $\sigma_u^2 > 0$ and u_t has finite 4-th moment. We have n observations on X_t: $\{X_t\}_{t=1}^n$.

(a) (10 points) Suppose $|\phi_o| < \frac{1}{2}$. Is $\{X_t\}$ a causal process? Justify your answer.

(b) (5 points) Suppose that $\phi_o = 0$. Find the auto-covariance function of $\{X_t\}$.

(c) (5 points) Suppose that $\phi_o = 0$. Find the long-run variance (LRV) of $\{X_t\}$.

Question 3 (40 points) Consider the following model

$$X_t = \alpha_o t + u_t \text{ with } u_t = \sqrt{t}\varepsilon_t \tag{1}$$

where $\varepsilon_t \sim i.i.d.(0, \sigma_\varepsilon^2)$ with $\sigma_\varepsilon^2 > 0$, ε_t has finite 4-th moment, and α_o and σ_ε^2 are unknown parameters. We have n observations on X_t: $\{X_t\}_{t=1}^n$.

(a) (15 points) Derive the asymptotic distribution of the LS estimator of α_o:

$$\widehat{\alpha}_n = \frac{\sum_{t=1}^n tX_t}{\sum_{t=1}^n t^2}. \tag{2}$$

(b) (10 points) Construct a consistent estimator $\widehat{\sigma}_{\varepsilon,n}^2$ of σ_ε^2. Show the consistency of your estimator.

(c) (15 points) Using the estimator $\widehat{\sigma}_{\varepsilon,n}^2$ of σ_ε^2, one can construct an estimator of the variance of u_t as $\widehat{\sigma}_{u_t,n}^2 = t\widehat{\sigma}_{\varepsilon,n}^2$. Consider the generalized LS (GLS) estimator of α_o:

$$\widehat{\alpha}_{gls,n} = \frac{\sum_{t=1}^n tX_t\widehat{\sigma}_{u_t,n}^{-2}}{\sum_{t=1}^n t^2\widehat{\sigma}_{u_t,n}^{-2}}. \tag{3}$$

Derive the asymptotic distribution of $\widehat{\alpha}_{gls,n}$. Compare the asymptotic variances of the LS estimator and the GLS estimator and explain your findings.

Question 4 (22 points) An economist has data $\{Y_t\}_{t=1}^n$ which are generated from the following model:

$$Y_t = \theta_0 Y_{t-1} + u_t,$$

where $|\theta_0| < 1$ and u_t is $i.i.d.$ $(0, \sigma_u^2)$. The economist specifies the following AR(2) model

$$Y_t = \theta Y_{t-1} + \gamma Y_{t-2} + v_t,$$

for the data generating process of $\{Y_t\}_{t=1}^n$, and he/she estimates the coefficients θ and γ by the LS estimators:

$$\begin{pmatrix} \widehat{\theta}_n \\ \widehat{\gamma}_n \end{pmatrix} = \begin{pmatrix} \sum_{t=3}^n Y_{t-1}^2 & \sum_{t=3}^n Y_{t-1}Y_{t-2} \\ \sum_{t=3}^n Y_{t-1}Y_{t-2} & \sum_{t=3}^n Y_{t-2}^2 \end{pmatrix}^{-1} \begin{pmatrix} \sum_{t=3}^n Y_{t-1}Y_t \\ \sum_{t=3}^n Y_{t-2}Y_t \end{pmatrix}.$$

Derive the asymptotic distributions of the LS estimators $\widehat{\theta}_n$. Compare the limiting distribution of $\widehat{\theta}_n$ with the limiting distribution of the LS estimator

$$\widehat{\theta}_{*,n} = \sum_{t=2}^n Y_tY_{t-1} / \sum_{t=2}^n Y_{t-1}^2.$$

Which estimator, $\widehat{\theta}_n$ or $\widehat{\theta}_{*,n}$ do you prefer? Justify your answer.