Comprehensive Examination
Quantitative Methods
Spring, 2018

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

Instruction for Part I : Solve every question.

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

Question 1 (60 points)

In a given population, commuters can be of type A or type B. The time it takes to commuters of type A to arrive at their destinations is distributed $\exp(\lambda)$, and the time it takes those of type B is distributed $\exp(2\lambda)$. Suppose that the probability p of being a type A is 1/2. Let T denote the time it takes to a commuter to arrive at his/her destination.

(a; 15) Provide expressions, in terms of λ and p , for the Mean, Variance, and Moment Generating Function of T.

Suppose that you are given N independent observations $\{t_1, ..., t_N\}$ of the times $\{T_1, ..., T_N\}$ it took to N commuters to arrive at their respective destinations.

(b; 10) Provide a consistent estimator for λ. Show that it is consistent.

(c; 10) Derive the asymptotic distribution of the estimator you derived in (a). Justify each step.

Suppose that in addition to t_i you also observe for each commuter the value, x_i, of a vector of characteristics X. Assume that the support of X is the real line and that X is distributed independently of type, with a density f_X. Let m denote an unknown function and suppose that for commuters of type A the time is distributed $\exp(m(X))$ while for those of type B the time is distributed $\exp(2\ m(X))$.

(d; 5) Provide an expression, in terms of m, for

$$\Pr\left(T_1 > T_2 |\ X_1 = x_1, X_2 = x_2\right).$$

(It is not necessary to solve the expression.)

(e; 5) Provide an expression, in terms of m, for

$$\Pr\left(A | T_1 \geq \widetilde{t} \text{ and } X = x_1\right),$$

the probability that commuter 1 is of type A given that he/she arrived at time $T_1 \geq \widetilde{t}$ and $X = x_1$. (It is not necessary to solve the expression.)

(f; 5) Is the function m identified from the distribution of (T, X)? Justify your answer.

(g; 10) Suppose that the probability, p, of being of type A is unknown. Is this probability identified from the distribution of (T, X)? Explain.

Question 2 (40 points)

Consider a population of consumers and two possible products available for purchase, A and B. The value of product A is $V_A + \varepsilon_A$, where $V_A = \beta\ X_A - P_A$. The value of product B is $V_B + \varepsilon_B$, where $V_B = \beta\ X_B - P_B$. Not purchasing a product has value 0. The support of (X_A, X_B) is R^2 while that of (P_A, P_B) is R^2_+. Let W denote the random vector (X_A, X_B, P_A, P_B). The random vector $(\varepsilon_A, \varepsilon_B)$ is distributed independently of W with a Normal density with mean $(0, 0)$ and variance

$$\begin{pmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{pmatrix}.$$

Denote $Z_A = 1$ if A is purchased and $Z_A = 0$ otherwise. Similarly, denote $Z_B = 1$ if B is purchased and $Z_B = 0$ otherwise. Let $Z_0 = 1$ if neither A nor B is purchased, with $Z_0 = 0$ otherwise.

Each consumer chooses the alternative (A, B, or neither) with the highest value. That is, in this model,

$$Z_A = \begin{cases} 1 \quad if \quad V_A + \varepsilon_A \geq 0 \quad and \quad V_A + \varepsilon_A \geq V_B + \varepsilon_B \\ 0 \qquad\qquad\qquad\qquad otherwise \end{cases}$$

and

$$Z_B = \begin{cases} 1 \quad if \quad V_B + \varepsilon_B \geq 0 \quad and \quad V_B + \varepsilon_B \geq V_A + \varepsilon_A \\ 0 \qquad\qquad\qquad\qquad otherwise \end{cases}$$

(a; 5) Provide an expression, in terms of the given functions and parameters, for

$$\Pr\left(Z_0 = 1 | W = (x_A, x_B, p_A, p_B)\right),$$

the probability that a consumer does not purchase either product, given (x_A, x_B, p_A, p_B).

(b; 10) Provide an expression, in terms of the given functions and parameters, for

$$\Pr\left(Z_A = 1 |\ Z_0 = 0 \text{ and } W = (x_A, x_B, p_A, p_B)\right),$$

the probability that given (x_A, x_B, p_A, p_B) and given that the consumer purchases a product, the purchased product is A.

(c;10) What parameters (if any) are identified from the distribution of (Z_A, Z_B, W)? Provide proofs for your answers.

(d; 10) If ε_A and ε_B were identically and independently distributed, each with marginal unknown density f_ε, and for all t, $f_\varepsilon(t) = f_\varepsilon(-t)$, would f_ε be identified? Provide a proof for your answer.

(e; 5) Suppose N independent consumers face the same value (x_A, x_B, p_A, p_B) of W. Let Y_0 denote the number of them that purchase no product, Y_A denote the number of them that purchase product A, and Y_B denote the number of them that purchase product B. What are the expected values of Y_0, Y_A and Y_B? Conditional on $Y_0 = 0$, are Y_A and Y_B independent? Explain.

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

Question 1 (60 points) Suppose $\{Y_i, X_i\}_{i=1}^n$ is an i.i.d. sample with $Y_i \in \{0,1\}$ (i.e. Y_i is binary) and $X_i \in \mathbf{R}^d$. Throughout the problem, assume that $E[X_i X_i']$ is finite and invertible, and that

$$Y_i = X_i'\beta_0 + \varepsilon_i \qquad E[\varepsilon_i | X_i] = 0. \tag{1}$$

(a) (5 pts) Show that model (1) implies that $P(Y_i = 1|X_i) = X_i'\beta_0$.

(b) (10 pts) Show that model (1) implies the distribution of ε_i conditional on X_i equals

$$\varepsilon_i = \begin{cases} 1 - X_i'\beta_0 & \text{with probability } X_i'\beta_0 \\ -X_i'\beta_0 & \text{with probability } 1 - X_i'\beta_0 \end{cases} . \tag{2}$$

Is model (1) homoskedastic or heteroskedastic? Justify your answer.

(c) (15 pts) Let $\hat{\beta}_n$ denote the ordinary least squares (OLS) estimator of β_0. Formally derive the asymptotic distribution of $\sqrt{n}\{\hat{\beta}_n - \beta_0\}$ under the stated assumptions.

(d) (20 pts) Suppose that for a known $x_0 \in \mathbf{R}^d$ you want to test the following null hypothesis

$$H_0 : x_0'\beta_0 \geq 0 \qquad H_1 : x_0'\beta_0 < 0. \tag{3}$$

Propose a test statistic and a critical value for (3) that result in a level α test.

(e) (10 pts) Explain why upon rejecting (3) we should also reject model (1). Suppose in fact the null hypothesis of interest is whether model (1) is correct. Is the test in part (d) consistent for this null hypothesis? Justify your answer.

Question 2 (40 pts) Let $D_i \in \{0,1\}$ be an indicator for whether individual i receives treatment, let $Y_i(d)$ denote the outcome of individual i when treatment status equals d, and assume $E[Y_i^2(d)] < \infty$ for $d \in \{0,1\}$. We do not observe $(Y_i(0), Y_i(1))$ but instead see Y_i which is equal to

$$Y_i \equiv Y_i(0) + D_i(Y_i(1) - Y_i(0))$$

(a) (10 pts) Suppose individuals decide whether to receive treatment according to the rule

$$D_i = 1\{U_i \leq 0.5\} \tag{4}$$

for U_i independent of $(Y_i(0), Y_i(1))$ and uniformly distributed on $[0, 1]$. Show that $\alpha_0 = E[Y(0)]$ and $\beta_0 = E[Y(1) - Y(0)]$ is the unique solution to the moment conditions

$$E\Big[(Y - \alpha_0 - D\beta_0)\begin{pmatrix} 1 \\ D \end{pmatrix}\Big] = 0.$$

(Hint: In the identity $Y_i = \alpha_0 + D_i\beta_0 + \varepsilon_i$, what is ε_i?)

(b) (15 pts) Suppose $\{Y_i, D_i\}_{i=1}^n$ is an i.i.d. sample (with D_i satisfying (4)). Let $(\hat{\alpha}_n, \hat{\beta}_n)$ be the ordinary least squares (OLS) estimators defined from

$$(\hat{\alpha}_n, \hat{\beta}_n) = \arg\min_{(\alpha,\beta)} \frac{1}{n}\sum_{i=1}^n (Y_i - \alpha - D_i\beta)^2.$$

Show that $\hat{\beta}_n$ is consistent for the average treatment effect.

(c) (15 pts) For simplicity next assume $Y(0) = 0$ with probability one. Suppose $0 < E[D] < 1$, and that (instead of (4)) individuals now select into treatment according to the rule

$$D_i = 1\{Y_i(1) > Y_i(0)\}.$$

Formally show that even if the average treatment effect is zero, the OLS estimator $\hat{\beta}_n$ (as in part (b)) converges to a positive number. What is the intuition for this result?

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

Question 1 (20 points) Suppose we have data $\{X_t\}_{t=1}^n$ from a linear process

$$X_t = \mu + \sum_{k=0}^{\infty} \varphi_k u_{t-k}$$

where μ is a finite constant, $\sum_{k=0}^{\infty} k\,|\varphi_k| < \infty$ and $\{u_t\}_t \sim iid(0, \sigma^2)$.

(a) (5 points) Derive the asymptotic distribution of $\overline{X}_n$, where $\overline{X}_n = n^{-1}\sum_{t=1}^n X_t$.

(b) (5 points) Let $X_t^* = X_t - \mu$. Define $B_{X^*,n}(r) = n^{-1/2}\sum_{t=1}^{[nr]} X_t^*$ for any $r \in [1/n, 1]$ and $B_{X^*,n}(r) = 0$ for any $r \in [0, 1/r)$, where $[nr]$ denotes the largest integer which is smaller than or equal to nr. Show that

$$\int_0^1 (B_{X^*,n}(r))^2 dr \to_d \int_0^1 (B_{X^*}(r))^2 dr$$

where $B_{X^*}(\cdot)$ is a Brownian motion on $[0, 1]$.

(c) (10) Suppose we are interested in testing H_0: $\mu = 0$ using the following test statistic

$$T_n = \frac{n^{1/2}\overline{X}_n}{\left(\int_0^1 (B_{X,n}(r))^2 dr\right)^{1/2}}$$

where $B_{X,n}(r) = n^{-1/2}\sum_{t=1}^{[nr]} X_t$ for any $r \in [1/n, 1]$ and $B_{X,n}(r) = 0$ for any $r \in [0, 1/r)$. Show that the asymptotic distribution of T_n under H_0 does not depend on any unknown parameters. Briefly discuss how to obtain critical values for the test based on T_n.

Question 2 (30 points) Let $\{X_t\}_{t \in \mathbb{Z}}$ be a time series which satisfies:

(i) $E[X_t] = 0$ for any t;

(ii) $\sup_t E[X_t^r] < \infty$ for some $r \geq 2$;

(iii) $|E[X_t X_{t-j}]| \leq 2\phi_j^{1-1/r}(E[X_t^2])^{1/2}(E[|X_t|^r])^{1/r}$ for any t and any positive integer j, where $r \geq 2$ and ϕ_j is a positive constant which satisfies

$$\sup_{j\geq 0}\left\{\phi_j \times j^{\frac{r}{r-1}+\delta}\right\} \leq C_\phi < \infty$$

where $\delta > 0$ is a finite constant.

(a) (10 points) show that the average autocovariance of order j of $\{X_t\}_{t\in\mathbb{Z}}$ can be bounded by $C/j^{1+\varepsilon}$ for some positive constants C and ε, i.e.,

$$\left| n^{-1}\sum_{t=j+1}^{n} Cov(X_t, X_{t-j})\right| \leq \frac{C}{j^{1+\varepsilon}}.$$

Find the constants C and ε.

(b) (10 points) Using the result in (a), show that the variance of $n^{-1/2}\sum_{t=1}^{n} X_t$ is bounded from above by some fixed constant.

(c) (10 points) Show that $n^{-1}\sum_{t=1}^{n} X_t = O_p(n^{-1/2})$.

Question 3 (20 points) The trend regression equation

$$Y_t = \theta_o\rho^t + u_t,\ t = 1, 2, ..., n \tag{5}$$

models the scalar observed time series Y_t in terms of the exponential trend t where $\rho > 1$ is known, where θ_o is an unknown coefficient to be estimated, and where $\{u_t\}_{t\in\mathbb{Z}}$ is i.i.d. with normal distribution $u_t \sim N(0, \sigma^2)$. Let $\widehat{\theta}_n$ be the OLS estimator of θ_o, i.e.,

$$\widehat{\theta}_n = \left(\sum_{t=1}^{n}\rho^{2t}\right)^{-1}\left(\sum_{t=1}^{n}\rho^t Y_t\right).$$

Show that $\widehat{\theta}_n$ is consistent and find its limit distribution.

Question 4 (30 points) Suppose that we have data $\{X_t\}_{t=1}^n$ from the following ARMA(1,1) model

$$X_t = \theta_o X_{t-1} + \varepsilon_t + \varepsilon_{t-1}$$

where θ_o is a real value with $|\theta_o| < 1$ and $\varepsilon_t \sim iid\ (0, \sigma_\varepsilon^2)$ with $E[\varepsilon_t^4] < \infty$.

(a) (10 points) Let $\widehat{\theta}_n$ denote the OLS estimator of θ_o, i.e.,

$$\widehat{\theta}_n = \frac{\sum_{t=1}^{n} X_{t-1}X_t}{\sum_{t=1}^{n} X_{t-1}^2}.$$

Derive the probability limit of $\widehat{\theta}_n$. Is it a consistent estimator?

(b) (10 points) Show that $E\left[(X_t - \theta_o X_{t-1})X_{t-2}\right] = 0$. Provide a consistent estimator of θ_o based on the above moment condition. Derive the limiting distribution of your estimator.

(c) (10 points) Using the results in (b), provide a test with asymptotic size $\alpha \in [0, 1]$ for the $H_0 : \theta_o = 0$.