Comprehensive Examination
Quantitative Methods
Spring, 2023

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.

1. (4 pts.) *No derivation is required for the questions below; your derivation will not be read anyway.* Consider $X_1, \ldots, X_n$ *i.i.d.* $N(\mu, \sigma^2)$. We assume that σ^2 is known to be equal to 1. We have $H_0 : \mu = 0$ vs. $H_1 : \mu \neq 0$. We plan to reject H_0 if $\left|\sqrt{n}\overline{X}\right| > 1.96$. Let $P_\mu\left[\left|\sqrt{n}\overline{X}\right| > 1.96\right]$ denote the probability of rejecting H_0, where the subscript μ makes its dependence on μ explicit. For $\mu > 0$, calculate $\lim_{n\to\infty} P_\mu\left[\left|\sqrt{n}\overline{X}\right| > 1.96\right]$. Your answer should be a number.

2. (6 pts.) *No derivation is required; your derivation will not be read anyway.* Suppose that $X_1, \ldots, X_n$ are *i.i.d.* and their common distribution is uniform on $(0, \theta)$, i.e., their common PDF $f(x)$ is equal to
$$\frac{1(0 < x < \theta)}{\theta}.$$
We have $H_0 : \theta = 1$ vs. $H_1 : \theta > 1$.

   (a) (3 pts.) Suppose that you decided to reject H_0 if $\max(X_1, \ldots, X_n) > 1 - \frac{t}{n}$. Suppose that the t is chosen such that the asymptotic size of the test is 5%, i.e., $\lim_{n\to\infty} \Pr\left[\max(X_1, \ldots, X_n) > 1 - \frac{t}{n}\right] = 0.05$ when $\theta = 1$.What is t? Your answer should be a number or a mathematical formula that can be evaluated by a scientific calculator. Hint:
$$\Pr\left[\max(X_1, \ldots, X_n) \leq 1 - \frac{t}{n}\right] = \Pr\left[X_1 \leq 1 - \frac{t}{n}, \ldots,\ X_n \leq 1 - \frac{t}{n}\right].$$

(b) (3 pts.) Now, suppose that you decided to reject H_0 if $\max(X_1, \ldots, X_n) > 1$. Consider the power of the test at $\theta = 1 + \frac{s}{n}$ for some $s > 0$. What is its limit as $n \to \infty$? Your answer should be a mathematical function in s.

3. (4 pts.) *No derivation is required for the questions below; your derivation will not be read anyway.* Suppose that the conditional distribution of X given $Y = y$ is binomial$(1, y)$. Also suppose that the (marginal) distribution of Y is *beta* $(1, 2)$. Hint: You may want to recall that the pdf of *beta* (α, β) is:

$$f(x) = \begin{cases} \dfrac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\,\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1} & 0 < x < 1 \\ 0 & \text{elsewhere} \end{cases}$$

You may also want to recall that the mean and variance of the *beta* (α, β) distribution are

$$\frac{\alpha}{\alpha+\beta}, \quad \frac{\alpha\beta}{(\alpha+\beta+1)(\alpha+\beta)^2}.$$

(a) (2 pts.) What is $E[X]$? Your answer should be a number.

(b) (2 pts.) What is $\mathrm{Var}(X)$? Your answer should be a number.

4. (4 pts.) *No derivation is required for the questions below; your derivation will not be read anyway.* Suppose that $X_1, X_2, \ldots$ are independent and identically distributed such that their common MGF is $\exp[e^t - 1]$. Let $F_n(x) \equiv \Pr\left(n^{-1}\sum_{i=1}^n X_i \le x\right)$. What is $\lim_{n\to\infty} F_n(0.9)$? Your answer should be a number.

5. (4 pts.) *No derivation is required for the questions below; your derivation will not be read anyway.* Suppose that the joint PDF of (X, Y) is given by the formula $f_{X,Y}(x, y) = (x + y)\,1(0 < x < 1)\,1(0 < y < 1)$. What is $E[Y|X = x]$? Your answer should be a mathematical function in x.

6. (4 pts.) *No derivation is required for the questions below; your derivation will not be read anyway.* Suppose that X is Bernoulli(1/2); that is $X = 1$ with probability 1/2 and $X = 0$ with probability 1/2. Suppose also that the distribution of Y conditional on $X = 0$ is exponential with mean 1, and the distribution of Y conditional on $X = 1$ is exponential with mean 2. In other words, when $X = 0$, the conditional PDF of Y is

$$f_{Y|X=0} = \begin{cases} e^{-y} & \text{if } y > 0 \\ 0 & \text{otherwise} \end{cases}$$

and when $X = 1$, the conditional PDF of Y is

$$f_{Y|X=1} = \begin{cases} \frac{1}{2}e^{-y/2} & \text{if } y > 0 \\ 0 & \text{otherwise} \end{cases}$$

What is $\mathrm{Cov}(Y, X)$? Your answer must be a number. Hint: If a random variable X has the PDF $\lambda e^{-\lambda x} 1(x > 0)$, we have $E[X] = 1/\lambda$ and $\mathrm{Var}(X) = 1/\lambda^2$.

7. (4 pts.) *No derivation is required for the questions below; 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

$$\begin{aligned} E\left[Y^2 - \left(\alpha + \beta X^2\right)\right] &= 0 \\ E\left[X^2\left(Y^2 - \left(\alpha + \beta X^2\right)\right)\right] &= 0 \end{aligned}$$

(a) (2 pts.) What is the numerical value of α?

(b) (2 pts.) What is the numerical value of β?

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 randomized controlled trial (RCT) with a binary treatment $D_i \in \{0, 1\}$, a binary covariate $X_i \in \{0, 1\}$, and potential outcomes $(Y_i(0), Y_i(1))$. We observe an i.i.d. sample $\{Y_i, D_i, X_i\}_{i=1}^n$ in which Y_i is determined according to

$$Y_i = Y_i(0) + D_i(Y_i(1) - Y_i(0)).$$

Suppose that in this experiment the probability of assigning an individual to treatment depends on the covariate X (e.g., women are more likely to be assigned to treatment than men) and assume D is independent of $(Y(0), Y(1))$ *conditionally* on X. In what follows it will be useful to note that $Y(d)$ independent of D conditionally on X implies

$$E[Y(d)|D = d, X = x] = E[Y(d)|X = x]$$

(a) (10 points) Show that D being independent of $Y(1)$ conditionally on X implies

$$\begin{aligned} &E[Y(1)|D = 1] \\ &\qquad = E[Y(1)|X = 0]P(X = 0|D = 1) + E[Y(1)|X = 1]P(X = 1|D = 1). \end{aligned}$$

(b) (10 points) Show that D is not necessarily independent of $(Y(0), Y(1))$ in this application. (**Hint:** Compare part (a) to $E[Y(1)]$).

(c) (10 points) A researcher decides to estimate the following regression

$$Y = \alpha + D\beta + X\gamma + DX\pi + \varepsilon$$

Use that the regression is saturated to argue that the OLS estimands satisfy

$$E[Y(1) - Y(0)|X = 0] = \beta \qquad E[Y(1) - Y(0)|X = 1] = \beta + \pi$$

(d) (10 points) Propose an estimator for the average treatment effect $E[Y(1) - Y(0)]$ and rigorously establish its consistency. Provide intuition on why your estimator is consistent despite D not being fully independent of $(Y(0), Y(1))$.

(e) (10 points) Rigorously derive the asymptotic distribution of the estimator you proposed in part (d).

Question 2 (50 points) Consider a panel data model in which we observe an outcome $Y_{it} \in \mathbf{R}$ and covariates $X_{it} \in \mathbf{R}$ for $1 \leq i \leq n$ individuals and $1 \leq t \leq T$ time periods. Throughout assume observations are i.i.d. across individuals and for $\bar{X}_i \equiv \sum_{t=1}^{T} X_{it}/T$ suppose

$$Y_{it} = X_{it}\beta + \varepsilon_{it} \qquad\qquad \varepsilon_{it} = \bar{X}_i\gamma + \eta_{it} \tag{1}$$

with $(\eta_{i1}, \ldots, \eta_{iT})$ independent of $(X_{i1}, \ldots, X_{iT})$ and satisfying $E[\eta_{it}] = 0$.

(a) (10 points) A researcher ignores that the dependence of ε_{it} on $\bar{X}_i$ and computes

$$\hat{\beta}_n^{(\mathrm{i})} = \arg\min_{b \in \mathbf{R}} \frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} (Y_{it} - X_{it}b)^2.$$

Formally show this estimator is inconsistent and that its probability limit equals

$$\hat{\beta}_n^{(\mathrm{i})} \xrightarrow{p} \beta + (\frac{1}{T}\sum_{t=1}^{T} E[X_{it}^2])^{-1} E[(\bar{X}_i)^2]\gamma.$$

(b) (10 points)A second researcher argues that the estimator in part (a) has an omitted variable bias. Motivated by this observation she instead computes

$$(\hat{\beta}_n^{(\mathrm{ii})}, \hat{\gamma}_n) \equiv \arg\min_{b,g \in \mathbf{R}} \frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} (Y_{it} - X_{it}b - \bar{X}_i g)^2.$$

Is it possible to estimate this regression if X_{it} is constant through time at the individual level (i.e. $X_{i\tilde{t}} = X_{it}$ when $\tilde{t} \neq t$)? Justify your answer.

(c) (10 points) Formally show that $\hat{\beta}_n^{(\mathrm{ii})}$ is consistent for β.

(d) (10 points) Establish that the estimator $\hat{\beta}_n^{(\mathrm{ii})}$ is in fact numerically equivalent to the fixed effects estimator (**Hint:** Frisch-Waugh-Lovell).

(e) (10 points) Suppose that in (1) we have $\varepsilon_{it} = F(\bar{X}_i)\gamma + \eta_{it}$ for some function F (e.g., $F(\bar{X}_i) = (\bar{X}_i)^2$) instead of $\varepsilon_{it} = \bar{X}_i\gamma + \eta_{it}$. Would the estimator $\hat{\beta}_n^{(\mathrm{ii})}$ remain consistent for β in this case? Why or why not?

Part III - 203C

Instruction for Part III: Solve every question. The total number of points is 100. Allocate your time wisely!

Let T denote the total number of points.

1. If $T \geq 85$, you will get H.
2. $60 \leq T < 85$, you will get P.
3. $45 \leq T < 60$, you will get M.
4. If $T < 45$, you will get F.

Question 1 (40 points)

Consider a model specified as

$$Y_1 = m\left(Y_2\right) + \varepsilon_1$$

$$Y_2 = s\left(X, \varepsilon_2\right)$$

where (Y_1, Y_2, X) is observed, $(\varepsilon_1, \varepsilon_2)$ is unobserved and distributed independently of X, the functions m and s are continuously differentiable, and s is strictly *decreasing* in ε_2. Denote the cumulative distribution of $(\varepsilon_1, \varepsilon_2)$ by $F_{\varepsilon_1,\varepsilon_2}.$ and denote the cumulative distribution of (Y_1, Y_2, X) by $F_{Y_1,Y_2,X}$. Assume that both are continuously differentiable and that $F_{Y_1,Y_2,X}$ is known while $F_{\varepsilon_1,\varepsilon_2}$ is unknown. Assume also that the functions m and s are unknown.

(a; 10) Obtain an expression, using only the known distribution and the known properties of the functions, for the derivative of s with respect to x when $Y_2 = y_2$, $X = x$, and when the value of ε_2 is t such that $y_2 = s\left(x, t\right)$. Justify your steps.

(b; 10) Obtain an expression, using only the known distribution and the known properties of the functions, for the derivative of m with respect to y_2 when $Y_2 = y_2$. Justify your steps. If you need additional assumptions, impose them after justifying them.

For (c) and (d), assume that that $(\varepsilon_1, \varepsilon_2)$ is distributed $N\left(0, \Sigma\right)$ where

$$\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix}$$

and that

$$m(y_2) = \gamma_1 + \gamma_2 \; (y_2)^2$$

$$s(x, \varepsilon_2) = \delta_1 + \delta_2 \; x + \varepsilon_2$$

Assume also that you have available data from N independent observations $\{y_1^i, y_2^i, x^i\}_{i=1,\ldots,N}$

(c; 10) Specify how to use the data to estimate the identified parameters using a Maximum Likelihood estimator. Specify the properties of the estimator and justify them. (You do not need to provide proofs.)

(d; 10) Explain how to test the hypothesis that $\gamma_2 = 0$ versus the hypothesis that $\gamma_2 \neq 0$. Justify your steps and provide expressions only in terms of the observations.

Question 2 (60 points):

Consider a model where each individual faces two products and can choose to purchase either one of them, both, or none. Let

$$(a_1, a_2) = \left\{ \begin{array}{lll} (0,0) & if & \text{no product is purchased} \\ (1,0) & if & \text{only product 1 is purchased} \\ (0,1) & if & \text{only product 2 is purchased} \\ (1,1) & if & \text{both products, 1 and 2, are purchased} \end{array} \right\}$$

Assume that the utility of the individual is given by

$$U(a_1, a_2, W, X_1, X_2)$$

$$= (\alpha_1 + \alpha_2 \; W + \alpha_3 \; X_1 + \varepsilon_1) \; a_1 + (\beta_1 + \beta_2 \; W + \beta_3 \; X_2 + \varepsilon_2) \; a_2 + (\gamma_1 + \gamma_2 \; W + \varepsilon_3) \; a_1 \; a_2$$

where X_1 and X_2 denote two individual specific characteristics and W denotes the income of the individual.

(a; 10) Denoting the cumulative distribution of $(\varepsilon_1, \varepsilon_2, \varepsilon_3)$ by $F_{\varepsilon_1,\varepsilon_2,\varepsilon_3}$, provide expressions for the probability that an individual with $(W, X_1, X_2) = (w, x_1, x_2)$ will choose each of the four alternatives in terms of $F_{\varepsilon_1,\varepsilon_2,\varepsilon_3}$ and the parameters of the function U.

(b; 10) Assume that $(\varepsilon_1, \varepsilon_2, \varepsilon_3)$ is $N(\mu, \Omega)$ where

$$\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{pmatrix} \quad \text{and} \quad \Omega = \begin{pmatrix} \omega_{11} & \omega_{12} & 0 \\ \omega_{21} & \omega_{22} & 0 \\ 0 & 0 & \omega_{33} \end{pmatrix}$$

Obtain an expression for the probability that an individual with $(X_1, X_2, W) = (x_1, x_2, w)$ will choose to purchase only product 1 (that is, $(a_1, a_2) = (1, 0)$) in terms of the parameters of U and of the distribution of $(\varepsilon_1, \varepsilon_2, \varepsilon_3)$.

(c; 10) Specify what parameters, or functions of parameters, are identified. Justify your answer.

(d; 20) Assume that you have a data set of N independent observations $\{(a_1^i, a_2^i), X_1^i, X_2^i, W^i\}_{i=1,\ldots,N}$ generated from the above parametric model. After imposing any necessary normalizations, propose a consistent estimator for the parameters of the utility function. Justify your answer.

(e; 10) Assume now that any typical individual can only purchase either both products or none, and that the utility function is given by

$$U(a_1, a_2, W) = \begin{cases} \gamma_1 + \gamma_2 \, W + \eta & if \quad (a_1, a_2) = (1, 1) \\ \\ 0 & if \quad (a_1, a_2) = (0, 0) \end{cases}$$

with γ_2 being an *unobservable random variable* distributed independently of the unobservable random variable η with a $N(\overline{\gamma}_2, \sigma_2^2)$ distribution and with η distributed $N(0, \sigma_1^2)$. The values of $\gamma_1, \overline{\gamma}_2, \sigma_1^2$ and σ_2^2 are unknown. Determine what parameters, or functions of parameters, can be identified and estimated consistently under this situation, and justify your answer.