Name: ______________________________

ECON 21020: Econometrics

The University of Chicago, Spring 2022

Instructor: Thomas Wiemann

Midterm

Date: May 2, 2022

1. **Write your name on the exam.**
2. The exam is closed book and closed notes.
3. No calculators are allowed.
4. There are a total of 100 possible points (+ 5 bonus points).
5. Answer as many questions as you can. **You do not need to answer questions in order.** Try to answer later parts of a question even if you struggle with earlier parts.
6. Please write your answers in the space provided on the exam paper. There is additional scratch paper at the end of the exam.
7. Label your final answers clearly where appropriate.
8. Do not take this exam with you.
9. Any students caught cheating will fail the course.
10. **Good luck!**

Problem 1 16 Points

The following are "True or false?"-questions. If the statement is true, provide a brief proof ($\approx$ 3 lines). If the statement is false, provide a counter example. There are no points awarded for answers without a proof or counter example.

a) 4 Points

True or false? Let X and Y be random variables.[1] If $X \perp\!\!\!\perp h(Y)$ for all functions h, then $X \perp\!\!\!\perp Y$.

b) 4 Points

True or false? Let X and U be random variables. If $E[U|X] = 0$, then $E[g(U)|X] = 0$ for all continuous functions g.

[1] Recall from the lectures that $X \perp\!\!\!\perp Y \Rightarrow X \perp\!\!\!\perp h(Y)$, for all function h. Does the reverse hold, too?

c) 4 Points

True or false? Let X be a continuous random variable. If $f(x)$ is the probability density function (pdf) of X, then $f(x) \in [0, 1], \forall x \in \operatorname{supp} X$.

d) 4 Points

True or false? Let X be a discrete random variable. If $f(x)$ is the probability mass function (pmf) of X, then $f(x) \in [0, 1], \forall x \in \operatorname{supp} X$.

Problem 2 24 Points

Let X be a random variable such that $Var(X) > 0$. Consider a sample $X_1, \ldots, X_n \overset{iid}{\sim} X$.

This exercise studies an alternative estimator

$$\hat{\mu}_n^{(\delta)} \equiv \frac{1}{n + n^\delta} \sum_{i=1}^{n} X_i \tag{1}$$

for the sample mean $\mu \equiv E[X]$, where $\delta \in \mathbb{R}$.

a) 4 Points

Show that

$$\frac{1}{n + n^\delta} \sum_{i=1}^{n} X_i = \arg\min_{\mu \in \mathbb{R}} \sum_{i=1}^{n} (X_i - \mu)^2 + n^\delta \mu^2. \tag{2}$$

b) 4 Points

Show that

$$\text{Bias}(\hat{\mu}_n^{(\delta)}) = \frac{-n^\delta}{n + n^\delta} \mu. \tag{3}$$

c) 6 Points

Show that

$$\hat{\mu}_n^{(\delta)} \xrightarrow{p} \mu, \tag{4}$$

whenever $\delta < 1$.

d) 10 Points

Give all values of δ for which it holds that

$$\sqrt{n}\frac{\left(\hat{\mu}_n^{(\delta)} - \mu\right)}{sd(X)} \xrightarrow{d} N(0,1). \tag{5}$$

(Hint: Use Slutsky's Theorem.)

Problem 3 24 Points

Suppose an econometrician wants to analyze the causal effect of a scalar-valued policy variable W on a scalar-valued outcome Y. For this purpose, she considers an all causes model, where (Y, W, U) is a random vector with joint distribution characterized by

$$Y = g(W, U), \tag{6}$$

and $g : \operatorname{supp} W \times \operatorname{supp} U \to \operatorname{supp} Y$ denotes an economic model.

The econometrician observes a sample $(Y_1, W_1), \ldots, (Y_n, W_n) \overset{iid}{\sim} (Y, W)$.

a) 2 Points

Briefly explain the purpose of U in the all causes model.

b) 2 Points

Consider defining the parameters

$$\tau_{w',w}(U) \equiv g(w', U) - g(w, U), \tag{7}$$

and

$$\mathrm{ATE}_{w',w} \equiv E[g(w', U) - g(w, U)], \tag{8}$$

for two values $w', w \in \operatorname{supp} W$.

Provide a brief interpretation of $\tau_{w',w}(U)$ and $\mathrm{ATE}_{w',w}$.

c) 2 Points

Briefly explain why the econometrician cannot construct an estimator for $\text{ATE}_{w',w}$ using the observed sample alone.

d) 4 Points

Suppose for the remainder of this exercise that random assignment of W holds – i.e., $W \perp\!\!\!\perp U$. Show that under random assignment

$$E[g(w,U)] = E[Y|W=w], \quad \forall w \in \operatorname{supp} W. \tag{9}$$

e) 4 Points

Suppose that W is discrete. State a sample analogue estimator for $\text{ATE}_{w',w}$.

(Hint: Recall from Problem Set 2 that for binary X, $E[Y|X=1] = E[YX]/E[X]$.)

f) 2 Points

Briefly explain why your estimator from Part e) would be ill-suited when W is continuous.

g) 4 Points

Suppose for the remainder of this exercise that W is continuous with $Var(W) > 0$. To simplify her analysis, the econometrician considers assuming linearity of the economic model g. In particular, she considers assuming

$$g(W, U) = \tilde{\alpha} + \tilde{\beta}W + U, \tag{10}$$

for two unknown scalars $\tilde{\alpha}, \tilde{\beta} \in \mathbb{R}$ and $E[U] = 0$.

Show that under the linearity assumption $(\tilde{\alpha}, \tilde{\beta})$ are the BLP$(Y|W)$-coefficients.

h) 4 Points

Show that under the linearity assumption,

$$\text{ATE}_{w',w} = \tau_{w',w}(U) = \tilde{\beta}(w' - w). \tag{11}$$

and conclude that $\tau_{w',w}(U)$ is non-random in this case.

Do you find non-random $\tau_{w',w}(U)$ a desirable implication of the linearity assumption? Explain briefly.

Problem 4   22 Points

Long-term unemployment is defined as being without work for six months or longer while actively looking for a job. As of January 2022, about 26% of unemployed individuals in the United States are long-term unemployed. Unsurprisingly, there is considerable political interest in finding effective solutions to reduce long-term unemployment.

Among the key reasons for long-term unemployment is structural unemployment, which describes the lack of relevant skills for the labor market. Government-funded job training programs (JTP) are efforts to decrease structural unemployment. These are meant to help individuals who are in long-term unemployment acquire new skills (e.g., IT skills) which improve their opportunities in the labor market.

To think carefully about the effect of a JTP on long-term unemployment, consider the all causes model where (Y, W, U) is a random vector with joint distribution characterized by

$$Y = g(W, U), \tag{12}$$

and $g : \operatorname{supp} W \times \operatorname{supp} U \to \operatorname{supp} Y$ denotes an economic model.

Let $\operatorname{supp} Y = \{0, 1\}$ where $Y = 1$ denotes being employed and $Y = 0$ otherwise, and $\operatorname{supp} W = \{0, 1\}$ where $W = 1$ denotes having participated in a JTP and $W = 0$ otherwise.

a)   2 Points

Give two examples for unobserved determinants U of Y.

b)   2 Points

Give an economic interpretation of

$$E[Y|W = 1] - E[Y|W = 0]. \tag{13}$$

c) 2 Points

Give an economic interpretation of

$$E[g(1, U)] - E[g(0, U)]. \tag{14}$$

Briefly explain why your response differs/does not differ from Part b).

d) 4 Points

Suppose that an econometrician calculated a sample analogue estimate $\hat{\mu}_{\Delta,n}$ of

$$\mu_\Delta \equiv E[Y|W = 1] - E[Y|W = 0] \tag{15}$$

based on the realization of an iid sample from the population of Americans in the labor force – that is, a realization of $(Y_1, W_1), \ldots, (Y_n, W_n) \overset{iid}{\sim} (Y, W)$. Her computations result in

$$\hat{\mu}_{\Delta,n} \approx -0.7, \qquad \text{and} \qquad se\left(\hat{\mu}_{\Delta,n}\right) \approx 0.15. \tag{16}$$

Construct an asymptotically-valid two-sided 95% confidence interval for μ_Δ.

e) 4 Points

Consider testing $H_0 : \mu_\Delta = 0$ against $H_1 : \mu_\Delta \neq 0$. Do you reject H_0 at a 5% significance level? Give an economic interpretation of your result.

f) 4 Points

A political party catches wind of the econometrician's estimates and is devastated. An email to party members on the next day reads:

"New data shows: Job training programs are increasing unemployment!"

Explain briefly why the email does not accurately describe the econometrician's results.

g) 4 Points

State an assumption on the joint distribution of (Y, W, U) that would warrant a causal interpretation of μ_Δ. Does it seem plausible here? Explain briefly.

Problem 5 14 Points

Let Y and X be random variables. Suppose that $X \sim U(-1,1)$ and recall that its probability density function (pdf) is

$$f_X(x) = \begin{cases} \frac{1}{2}, & \forall x \in [-1,1], \\ 0, & \text{otherwise}. \end{cases} \tag{17}$$

Recall from class that for $t \in [-1,1]$, we have

$$\begin{aligned} E[X|X \geq t] &= \frac{1}{2}(t+1), & E[X|X < t] &= \frac{1}{2}(t-1), \\ P(X \geq t) &= \frac{1-t}{2}, & Var(X) &= \frac{1}{3}. \end{aligned} \tag{18}$$

Consider the BLP($Y|X$)-coefficient β. By Yitzhaki's Theorem discussed in class, it holds that

$$\beta = \int_{-1}^{1} \left(\frac{\partial}{\partial t} E[Y|X=t] \right) \omega(t) dt, \tag{19}$$

where

$$\omega(t) = \frac{\left(E[X|X \geq t] - E[X|X < t]\right) P(X \geq t) P(X < t)}{Var(X)}. \tag{20}$$

a) 3 Points

Give an expression for the Yitzhaki weights $\omega(t)$ for $t \in [-1,1]$ and sketch the function.

b) 7 Points

Suppose that $E[Y|X] = -3X^9 + X$. Calculate the BLP($Y|X$)-coefficient and the average derivative of the conditional expectation function $E\left[\frac{\partial}{\partial X}E[Y|X]\right]$ to show that

$$E\left[\frac{\partial}{\partial X}E[Y|X]\right] < 0 < \beta. \tag{21}$$

c) 4 Points

Suppose that $E[Y|X]$ is unknown. Consider testing $H_0 : \beta \geq 0$ against $H_1 : \beta < 0$. Would rejection imply that there is evidence to suggest that a marginal change in X is not associated with an expected increase in Y? Explain briefly.

Problem 6 Bonus

This is an optional exercise. Correct solutions are awarded 5 bonus points.

Name the three distinct tasks arising in the analysis of causal questions according to Heckman and Vytlacil (2007).

References

Heckman, J. J. and Vytlacil, E. J. (2007). Econometric evaluation of social programs, part I: Causal models, structural models and econometric policy evaluation. In Heckman, J. J. and Leamer, E., editors, *Handbook of Econometrics*, volume 6, chapter 70, pages 4779–4874. Elsevier, Amsterdam.

(scratch work)

(scratch work)

(scratch work)

(scratch work)