Name:

Student ID:

Midterm ARE256b W2024, February 7, 2024

- You have 80 minutes.
- The maximum number of points is 100.
- Put your name and student ID number on every sheet of paper.
- Do not write outside of the boxes. We will only grade what you write inside the boxes. You can use extra page at the end if some of your answers didn't fit the corresponding box.

Notation. You can use the following functions. $\Lambda(z) = 1/(1+e^{-z})$ is the logistic CDF and $\lambda(z)$ is its pdf. $\Phi(z)$ is the standard normal cdf and $\phi(z)$ is its density.
For variable $Z \sim N(\mu, \sigma^2)$, the density $f(Z) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{Z-\mu}{\sigma}\right)^2}$.

Name:

Student ID:

---

1. **(50 points total) Short theory questions**

   (a) (12.5) In the following list, there is one log-likelihood function that corresponds to a single-variable logit model, and another that corresponds to a single-variable tobit model (where censoring occurs for $Y_i^* \leq 0$). One corresponds to neither. Explain *briefly* which likelihood corresponds to which model and why.

   i. $\sum_{i=1}^{n} \log \left\{ \left(1 - \Phi\left(\frac{\beta_0 + \beta_1 X_i}{\sigma}\right)\right)^{1\{Y_i=0\}} \left(\frac{1}{\sigma}\phi\left(\frac{Y_i - \beta_0 - \beta_1 X_i}{\sigma}\right)\right)^{1\{Y_i>0\}} \right\}$

   ii. $\sum_{i=1}^{n} \log \frac{\phi((Y_i - \beta_0 - \beta_1 X_i)/\sigma)}{\sigma \Phi((\beta_0 + \beta_1 X_i)/\sigma)}$

   iii. $\sum_{i=1}^{n} \log \left\{ (\Lambda(\gamma_0 + \gamma_1 Z_i))^{Y_i} (1 - \Lambda(\gamma_0 + \gamma_1 Z_i))^{1-Y_i} \right\}$

Name:

Student ID:

(b) (12.5) Suppose that Y_i is a binary random variable with conditional distribution $P(Y_i = 1|X_i) = \beta_0 + \beta_1 X_i$ and $P(Y_i = 0|X_i) = 1 - (\beta_0 + \beta_1 X_i)$. Write down the objective function for the conditional maximum likelihood estimator of parameters β_0 and β_1 using an i.i.d. sample $\{Y_i, X)_i\}_{i=1}^n$.

Name:

Student ID:

(c) (12.5) Consider a discrete variable, e.g. schooling, and earnings. Suppose that the schooling variable takes K different values. How can you estimate the relationship between schooling and earnings nonparametrically? Write down the regression equation and explain all the variables.

Name:

Student ID:

(d) (12.5) Give an example of an economic application of regression discontinuity (RDD) analysis. Discuss what endogeneity problem the RDD is addressing in your application. Write down the regression equation and provide an interpretation of the regressors, dependent variables, and the parameter of interest in this example.

Name:

Student ID:

---

2. **(30 points total) Using IV.** Suppose that you are interested in estimating demand function for milk,

$$\log q_i = \alpha + \beta \log p_i + \varepsilon_i. \tag{1}$$

We observe the market price p_i and sales q_i of a farm $i = 1, \ldots, N$. The residuals ε_i are demand shocks. Both shocks are unobserved.

Assume that demand shocks may be correlated with prices (farmers may reduce the price if demand is too low or increase it if demand is high).

(a) (10) How can you estimate the elasticity of the demand parameter β using proxy variables for demand shocks? Write down the properties of good proxy variables and the corresponding regression equation.

We can control for unobservables by adding proxi variables to the regression.

A good proxi variable is highly correlated with ε_i. In this case it explains all the unobservable factors that affect the demand for milk of brand i.

---

we can use level of advertisement b i, brand fixed effects, and demographic variables that explain taste for brand i.

Name:

Student ID:

(b) (10) How can you estimate the elasticity of the demand parameter β using instrumental variables (hint: supply shocks are good instruments)? Write down the properties of good instrumental variables and the corresponding first- and second-stage regression equations.

first stage $p_i = \delta_1 + \delta_2 Z_i + e_i$

second stage $q_i = \theta_1 + \theta_2 \log \hat{p}_i + v_i$

Good IV should satisfy:

1) Relevence: Z_i be correlated with P_i $\quad Cov(Z_i, P_i) \neq 0$

2) Validity: Z_i should be uncorrelated with unobservables $\quad Cov(Z_i, \varepsilon_i) = 0$

we can use price of cow feed (or index of prices of corn, oats, and barley) as an instrument.

Name:

Student ID:

(c) (10) Suppose that we observe all farmers for two periods t now,

$$\log q_{it} = \alpha + \beta \log p_{it} + \varepsilon^d_{it}. \tag{2}$$

Group A of the farmers increased their price by 1% in the second period, while Group B of farmers kept the price constant in both periods due to changes in production costs. Assume demand shocks follow common trend assumption,

$$\varepsilon^d_{it} = \alpha_i + \lambda_t + u_{it}, \tag{3}$$

where u_{it} is independent of $\log p_{it}$. Which of the casual inference techniques can you use to identify β without IV or proxy variables? Briefly explain how to implement it and why it works.

We can Diff-in-Diff.

$\mathbb{E}(\log q_{it} | A, t=2) = \alpha + \beta \mathbb{E}[\log p'_{it}] + \alpha_A + \lambda_2$

$\mathbb{E}(\log q_{it} | A, t=1) = \alpha + \beta \mathbb{E}[\log p_{it}] + \alpha_A + \lambda_1$

$\mathbb{E}(\log q_{it} | B, t=2) = \alpha + \beta \mathbb{E}[\log p'_{it}] + \alpha_B + \lambda_2$

$\mathbb{E}(\log q_{it} | B, t=1) = \alpha + \beta \mathbb{E}[\log p_{it}] + \alpha_B + \lambda_1$

Diff-in-Diff Estimator $= \beta \left(\mathbb{E} \log p' - \mathbb{E} \log p\right)$

$= \beta \log 1.1p - \log p$

$= \beta \left(\log \frac{1.1}{1}\right)$

$\approx \beta (0.04)$

So you can estimate DiD estimator and normalize it to identify β.

Name:

Student ID:

---

3. **(20 points total)** Recall that we said that if Y_i is i.i.d. and binary with $P(Y_i = 1) = \pi$, then $Y_i \sim Bernoulli(\pi)$.

   (a) (10) Derive the cumulative distribution function and the probability mass function of Y_i. What is the joint mass function for $\{Y_i\}_{i=1}^n$?

pmf is a function that gives $\mathbb{P}(Y_i = y_i)$

cdf is a function that gives $\mathbb{P}(Y_i \leq y_i)$

$$\text{pmf}(y_i) = \pi^{y_i}(1-\pi)^{1-y_i} \quad \text{for } y_i = 0, 1$$

$$\text{cdf}(y_i) = \begin{cases} 1 & y_i \geq 1 \\ 1-\pi & 0 < y_i < 1 \\ \pi & y_i \leq 0 \end{cases}$$

joint mass function is a function that gives $\mathbb{P}(Y_1 = y_1, \ldots, Y_n = y_n)$.

As Y_i is iid we can write

$$\mathbb{P}(Y_1 = y_1, \ldots, Y_n = y_n) = \prod_{i=1}^{n} \pi^{y_i}(1-\pi)^{1-y_i}$$

Name:

Student ID:

(b) (10) For the Bernoulli example, what is the MLE of $\hat{\pi}$? Does it have a closed-form solution?

$$L = \prod_{i=1}^{n} \pi^{y_i} (1-\pi)^{1-y_i}$$

$$\log L = \sum_i \left(y_i \log \pi + (1-y_i) \log (1-\pi) \right)$$

MLE estimator maximizes the log L.

$$\frac{\partial \log L}{\partial \pi} = 0$$

$$\Rightarrow \frac{\Sigma_i y_i}{\pi} - \frac{1}{1-\pi} \Sigma_i (1-y_i) = 0$$

$$\Rightarrow \frac{1-\pi}{\pi} = \frac{\Sigma_i (1-y_i)}{\Sigma_i y_i} = \frac{1-\bar{y}}{\bar{y}} \Rightarrow \pi = \bar{y}$$

we can show that the log L is concave at $\pi = \bar{y}$.

So $\hat{\pi} = \bar{y} = \frac{\Sigma_i y_i}{n}$ is the MLE.

Name:

Student ID:

---

4. **(0 points total) Extra space. Use it if some of your answers to the previous questions did not fit into the boxes.**