Second Midterm Review Sheet, Part I

Ec240a – Second Half, Fall 2019

In preparing for the exam review your lecture notes, assigned course readings, problem sets and prepare answers to the questions on the review sheet(s). You may bring to the exam a single 8.5×11 inch sheet of paper with notes on it. No calculation aides are allowed in the exam (e.g., calculators, phones, laptops etc.). You may work in pencil or pen, however I advise the use of pencil. Please be sure to bring sufficient blue books with you to the exam if possible. Scratch paper will be provided.

[1] Let $\mathbf{Y} = (Y_1, \ldots, Y_N)'$ be N independent measurements of the same outcome, each distributed

$$Y_i \sim \mathcal{N}\left(\mu, \sigma_i^2\right).$$

Let $\mathbf{c}$ be an $N \times 1$ vector of constants. Consider estimates of μ in the family

$$\hat{\mu} = \mathbf{c}'\mathbf{Y}. \tag{1}$$

[a] Show that the sample mean $\bar{Y} = \frac{1}{N}\sum_{i=1}^{N} Y_i$ is a member of (1).

[b] Show that the mean squared error minimizing choice of $\mathbf{c}$ is

$$\mathbf{c} = \mu^2 \left(\operatorname{diag}\left\{\sigma_1^2, \ldots, \sigma_N^2\right\} + \mu^2 \iota_N \iota_N'\right)^{-1} \iota_N$$

with ι_N an $N \times 1$ vector of ones and $\operatorname{diag}\left\{\sigma_1^2, \ldots, \sigma_N^2\right\}$ denoting a diagonal matrix.

[c] Further show that the i^{th} element of $\mathbf{c}$ is

$$c_i = \frac{\mu^2}{\sigma_i^2} \frac{\left[\sum_{i=1}^{N} \frac{1}{\sigma_i^2}\right]^{-1}}{\left[\sum_{i=1}^{N} \frac{1}{\sigma_i^2}\right]^{-1} + \mu^2}.$$

HINT: For A an invertible matrix, u and v column vectors and b a scalar:

$$(A + buv')^{-1} = A^{-1} - \frac{b}{1 + bv'A^{-1}u} A^{-1}uv'A^{-1}.$$

[d] Assume that $\sigma_i^2 = \sigma^2$ for all $i = 1, \ldots, N$. Show that in this case the mean squared error minimizing estimate of μ is

$$\hat{\mu} = \frac{\mu^2}{\frac{\sigma^2}{N} + \mu^2} \bar{Y}$$

Prove that this estimate converges in mean square to μ (and hence also converges in probability).

[e] The estimate in part [d] is infeasible. Assume that σ^2 is known and consider the feasible estimator

$$\hat{\mu} = \left(1 - \frac{\frac{\sigma^2}{N}}{\bar{Y}^2}\right) \bar{Y}.$$

Provide a justification for this estimate. *Argue* that $\hat{\mu} \xrightarrow{p} \mu$. Do you think its mean squared error will be lower than that of the sample mean's in finite samples? Why?

[f] Rebut the assertion that "the sample mean's day has come and gone".

[2] You observe a simple random sample of size N from the population

$$Y_0 \sim N\left(\mu, \sigma^2\right)$$

as well as a second, independent, simple random sample, also of size N, from the population

$$Y_1 \sim N\left(\mu, 4\sigma^2\right).$$

The value of σ^2 is known. Consider the family of estimates of μ

$$\hat{\mu}\left(c_0, c_1\right) = c_0\bar{Y}_0 + c_1\bar{Y}_1,$$

where $\bar{Y}_0 = \frac{1}{N}\sum_{i=1}^{N} Y_{0i}$ and $\bar{Y}_1 = \frac{1}{N}\sum_{i=1}^{N} Y_{1i}$.

[a] Show that mean squared error equals

$$\mathbb{E}\left[\left(\hat{\mu}\left(c_0, c_1\right) - \mu\right)^2\right] = \frac{c_0^2\sigma^2}{N} + \frac{c_1^2 4\sigma^2}{N} + \left(1 - c_0 - c_1\right)^2 \mu^2. \tag{2}$$

[b] Derive the oracle estimator (within the family) which minimizes (2).

[c] Show that

$$\hat{R}\left(c_0, c_1\right) = \frac{c_0^2\sigma^2}{N} + \frac{c_1^2 4\sigma^2}{N} + \left(1 - c_0 - c_1\right)^2 \frac{1}{2}\left\{\bar{Y}_0^2 + \bar{Y}_1^2 - \frac{5\sigma^2}{N}\right\} \tag{3}$$

is an unbiased estimate of of (3). Can you propose another unbiased risk estimate? Why would you prefer one unbiased risk estimate over another?

[d] Describe in *words* how one might use (3) to construct an implementable estimator of μ.

[3] Let Y denote log-earnings and X years of completed schooling for a cohort of workers. Assume a random sample of size N is available from this population. Let $D_x = 1$ if $X = x$ and zero otherwise. Assume that $X \in \{0, \ldots, 16\}$ with positive probability attached to each support point.

[a] Let

$$\mathbb{E}^*\left[Y \mid D_1, \ldots, D_L\right] = \alpha_0 + \sum_{l=1}^{16} \gamma_{0l} D_l.$$

What is the relationship between this linear predictor and $\mathbb{E}\left[Y \mid X = x\right]$?

[b] Assume that $\Pr\left(X = 6\right) = 0$. Is the linear predictor defined in part [a] still well-defined? Why or why not?

[c] You hypothesize that $\mathbb{E}\left[Y \mid X = x\right]$ is linear in x. Consider the linear predictor in part [a] and let $\beta = \left(\alpha, \gamma_1, \ldots, \gamma_{16}\right)'$. Show how your hypothesis may be equivalently expressed as set of linear restrictions of the form $C\beta_0 = c$. Provide explicit expressions for C and c. Describe how you would construct a test statistic for your hypothesis. What is the asymptotic sampling distribution of your statistic under the null? Assume that your have a consistent estimate $\hat{\Lambda}$ of the asymptotic variance-covariance matrix of $\sqrt{N}\left(\hat{\beta} - \beta\right)$, with $\hat{\beta}$ the least squares estimate.

[4] Let $X \in \{0, 1, 2\}$ and $Y \in \{0, 1, 2\}$. The probability of the event $X = x$ and $Y = y$ for all possible combinations of x and y is given in the following table:

X\Y	0	1	2
0	$\frac{1}{18}$	$\frac{1}{9}$	$\frac{3}{18}$
1	$\frac{3}{18}$	$\frac{1}{18}$	$\frac{1}{9}$
2	$\frac{1}{9}$	$\frac{3}{18}$	$\frac{1}{18}$

[a] Calculate $\mathbb{E}[Y]$ and $\mathbb{E}[Y|X=1]$. Are X and Y independent?

[b] Calculate $\mathbb{E}[X]$, $\mathbb{E}\left[X^2\right]$ and hence $\mathbb{V}(X)$.

[c] Calculate $\mathbb{C}(X,Y)$ and also the coefficient on X in $\mathbb{E}^*[Y|X]$.

[d] Calculate the intercept of $\mathbb{E}^*[Y|X]$.

[e] Repeat [a] to [d] above for the following joint distribution

X\Y	0	1	2
0	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$
1	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$
2	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$

[5] Let $X_1, ..., X_K$ be a set of regressors with the property that $\mathbb{C}(X_k, X_l) = 0$ for all $k \neq l$. We will show that

$$\mathbb{E}^*[Y|X_1,\ldots,X_K] = \sum_{k=1}^{K} \mathbb{E}^*[Y|X_k] - (K-1)\,\mathbb{E}[Y].$$

[a] First show that

$$\mathbb{E}^*[\mathbb{E}^*[Y|X_k]|X_l] = \mathbb{E}[Y]$$

for every $k \neq l$.

[b] Second verify the orthogonality conditions

$$\mathbb{E}[UX_l] = 0$$

for $U = \left(Y - \sum_{k=1}^{K} \mathbb{E}^*[Y|X_k] + (K-1)\,\mathbb{E}[Y]\right)$ and $l = 1,\ldots,K$.

[6] Let Y be a scalar random variable, X a K vector of covariates (which includes a constant), and W a vector of additional covariates (which excludes a constant). Consider the long (linear) regression

$$\mathbb{E}^*[Y|W,X] = X'\beta_0 + W'\gamma_0. \tag{4}$$

Next define the short and auxiliary regressions

$$\mathbb{E}^*[Y|X] = X'b_0 \tag{5}$$
$$\mathbb{E}^*[W|X] = \Pi_0 X. \tag{6}$$

[a] Let $V = W - \mathbb{E}^*[W|X]$ be the projection error associated with the auxiliary regression. Show that

$$\begin{aligned}\mathbb{E}^*[Y|V,X] &= \mathbb{E}^*[Y|X] + \mathbb{E}^*[Y|1,V] - \mathbb{E}[Y]\\ &= \mathbb{E}^*[Y|X] + \mathbb{E}^*[Y|V]\end{aligned}$$

where $\mathbb{E}^*[Y|1,V]$ denotes the linear regression of Y onto a constant and V, while $\mathbb{E}^*[Y|V]$ denotes the corresponding regression without a constant (HINT: Observe that $\mathbb{C}(X,V)=0$).

[b] Next show that $\mathbb{E}^*[Y|V,X] = \mathbb{E}^*[Y|W,X]$ and hence that the coefficient on V in $\mathbb{E}^*[Y|V,X]$ coincides with that on W in $\mathbb{E}^*[Y|W,X]$.

[c] Let $U = Y - \mathbb{E}^*[Y|X]$ be the projection error associated with the short regression. Derive the coefficient on V in the linear regression of U onto V (excluding a constant).

[d] Discuss the possible practical value of the results shown in [b] and [c] above.

[7] Let $m(Z) = \mathbb{E}[X|Z]$ and consider the linear regression

$$\mathbb{E}^*[Y|X, m(Z), A] = \alpha_0 + \beta_0 X + \gamma_0 m(Z) + A.$$

[a] Show that

$$\mathbb{E}^*[m(Z)|X] = \delta_0 + \xi_0 X$$

with

$$\begin{aligned} \delta_0 &= (1-\xi_0)\,\mathbb{E}[X] \\ \xi_0 &= \frac{\mathbb{V}(\mathbb{E}[X|Z])}{\mathbb{E}[\mathbb{V}(X|Z)] + \mathbb{V}(\mathbb{E}[X|Z])}. \end{aligned}$$

[b] Assume the population under consideration is working age adults who grew up in the San Francisco Bay Area. Let Y denote a adult log income, let X denote the log income of one's parents as a child and let Z be a vector of dummy variables denoting an individual's neighborhood of residence as a child. Provide an interpretation of ξ_0 as a measure of residential stratification by income.

[c] Establish the notation $\rho = \operatorname{corr}(A,X)$, $\mu_A = \mathbb{E}[A]$, $\mu_X = \mathbb{E}[X]$, $\sigma_A^2 = \mathbb{V}(A)$ and $\sigma_X^2 = \mathbb{V}(X)$. Show that

$$\mathbb{E}^*[Y|X] = \alpha_0 + \gamma_0(1-\xi_0)\mu_X + \left(\mu_A - \rho\frac{\sigma_A}{\sigma_X}\mu_X\right) + \left\{\beta_0 + \gamma_0\xi_0 + \rho\frac{\sigma_A}{\sigma_X}\right\}X.$$

[d] Your research assistant computes an estimate of $\mathbb{E}^*[Y|X]$ using random sample from San Francisco. She computes a separate estimate using a random sample from New York City. Assume that there is more residential stratification by income in New York than in San Francisco. How would you expect the intercept and slope coefficients to differ across the two regression fits?

[8]] Consider the statistical model

$$\begin{aligned} Y &= \alpha_0 + \beta_0 X + U \\ X &= \eta_0 + Z'\pi_0 + V \end{aligned}$$

with

$$\left(\begin{array}{c} U \\ V \end{array}\right)\Bigg| Z \sim \mathcal{N}\left(\left(\begin{array}{c} 0 \\ 0 \end{array}\right), \left(\begin{array}{cc} \sigma_U^2 & \rho\sigma_V\sigma_U \\ \rho\sigma_V\sigma_U & \sigma_V^2 \end{array}\right)\right).$$

[a] Let $\mu^2 = \pi_0'\mathbb{V}(Z)\pi_0/\mathbb{V}(V)$. Consider the (mean squared error minimizing) linear predictor of Y given

X. Let b_0 be the coefficient on X in this linear predictor. Show that

$$b_0 = \beta_0 + \rho \frac{\sigma_U}{\sigma_V} \frac{1}{\mu^2 + 1}.$$

Comment on your result. Why does $b_0 \neq \beta_0$ for $\rho \neq 0$? Discuss the above statistical model in light of the Card and Krueger (1996) model of schooling discussed in the readings.

[b] Assume that $\pi_0 \neq 0$ and show that

$$\mathbb{E}^* \left[Y \vert X, V \right] = \alpha_0 + \beta_0 X + \rho \frac{\sigma_U}{\sigma_V} V.$$

Provide an intuitive explanation for why additionally conditioning on V ensures that the LP coefficient on X is equal to β_0. Explain why this linear predictor is not well-defined if $\pi_0 = 0$.

[c] Using a sample of $N = 32,587$ Honduran males aged 45 to 51 in 1988 we compute the least squares fit of the logarithm of monthly earnings (LogEarnings) onto a constant and years of completed schooling (YrsSch):

$$\text{LogEarnings} = \underset{(0.0064)}{5.2067} + \underset{(0.0010)}{0.1324} \ \text{YrsSch} . \tag{7}$$

In Honduras compulsory schooling for the above cohorts began in the first February after turning seven (i.e., the school year begins in February and you must attend school if you are seven years old). Let D_{FMA} be a dummy variable taking a value of one if an individual was born in February, March or April and zero otherwise, D_{MJJ} a dummy for being born in May, June or July and D_{ASO} a dummy for being born in August, September or October. A least squares of fit of YrsSch onto a constant and these three dummy variables using the same sample yields

$$\text{YrsSch} = \underset{(0.0539)}{4.1127} - \underset{(0.0742)}{0.3372} \ \mathrm{D_{FMA}} - \underset{(0.0734)}{0.2711} \ D_{MJJ} + \underset{(0.0769)}{0.0406} \ \mathrm{D_{ASO}} . \tag{8}$$

Let $\hat{V}$ be the fitted least squares residual associated with (8) above. The least squares fit of LogEarnings onto a constant, YrsSch of $\hat{V}$ is

$$\text{LogEarnings} = \underset{(0.1197)}{5.1765} + \underset{(0.0302)}{0.1400} \ \text{YrsSch} - \underset{(0.0302)}{0.0076} \ \hat{V} . \tag{9}$$

[i] The estimated asymptotic variance-covariance matrix (divided by the sample size) of the least squares coefficient estimates reported in (8) above is

$$\begin{pmatrix} 0.0029 & & & \\ -0.0029 & 0.0055 & & \\ -0.0029 & 0.0029 & 0.0054 & \\ -0.0029 & 0.0029 & 0.0029 & 0.0059 \end{pmatrix}.$$

Test the hypothesis, at the $\alpha = 0.05$ level, that years of completed schooling cannot be predicted by quarter of birth. Provide a precise statement of this hypothesis in terms of the population analogs of the estimated coefficients reported in (8), construct a test statistic, and compare it to the appropriate critical value. For your reference the 0.95 quantiles of χ^2 random variables with parameters 1, 2 and 3 are, respectively, 3.84, 5.99 and 7.81. Why might quarter-of-birth predict years of completed schooling?

[ii] Test the hypothesis, at the $\alpha = 0.05$ level, that the coefficients on YrsSch in (7) and (9) coincide. Explain yourself.

[9] Consider the following model of supply and demand:

$$\begin{aligned} \ln Q_i^D(p) &= \alpha_1 + \alpha_2 \ln(p) + U_i^D \\ \ln Q_i^S(p) &= \beta_1 + \beta_2 \ln(p) + U_i^S, \end{aligned}$$

with i indexing a generic random draw from a population of 'markets' ; U_i^D and U_i^S are market-specific demand and supply shocks. We assume that $\left(U_i^S, U_i^D\right) \overset{i.i.d}{\sim} F$ for $i = 1, 2, \ldots, N$. In each market the observed price and quantity pair (P_i, Q_i) coincides with the solution to market clearing condition

$$Q_i^D(P_i) = Q_i^S(P_i) = Q_i.$$

[a] Provide an economic interpretation of the parameters α_2 and β_2. What signs do you expect them to take? Why?

[b] Depict the market equilibrium graphically. Solve for the equilibrium values of $\ln Q_i$ and $\ln P_i$ algebraically. How is the market price and quantity related to the demand and supply shocks, U_i^D and U_i^S? Provide some economic content for your answer. Can you use a figure to illustrate it?

[c] Calculate $\mathbb{E}^*[\ln Q | \ln P]$. You may assume that $\mathbb{C}\left(U^D, U^S\right) = 0$. Evaluate the coefficient on $\ln(P)$, does it coincide with an economically interpretable parameter? Assume that $\mathbb{V}\left(U_i^S\right) / \left(\mathbb{V}\left(U_i^S\right) + \mathbb{V}\left(U_i^D\right)\right) \approx 1$, does your answer change? Why?