Second Midterm Review Sheet, Part II

Ec240a – Second Half, Fall 2017

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 $\{X_i\}_{i=1}^N$ be a simple random sample from some population with density function $f(x)$. Consider the following uniform kernel density estimate of $f(x)$:

$$\hat{f}(x) = \frac{1}{2Nh} \sum_{i=1}^{N} \mathbf{1}(x - h \leq X_i \leq x + h).$$

[a] Show that

$$2Nh\hat{f}(x) \sim \text{binomial}(N, p_h(x))$$

for $p_h(x) = h \int \mathbf{1}(-1 \leq u \leq 1) f(x + uh) \, \mathrm{d}u$ and hence that

$$\mathbb{E}\left[\hat{f}(x)\right] = \frac{1}{2} \int \mathbf{1}(-1 \leq u \leq 1) f(x + uh) \, \mathrm{d}u$$

and further that

$$\mathbb{V}\left(\hat{f}(x)\right) = \frac{p_h(x)[1 - p_h(x)]}{4Nh^2}.$$

[b] Use the Central Limit Theorem introduced in class to prove that

$$\sqrt{N}\left(\hat{f}(x) - \mathbb{E}\left[\hat{f}(x)\right]\right) \xrightarrow{D} \mathcal{N}\left(0, \frac{p_h(x)[1 - p_h(x)]}{4h^2}\right).$$

[c] Use a Taylor series expansion to show that the bias of $\hat{f}(x)$ is, up to order h^2,

$$\mathbb{E}\left[\hat{f}(x) - f(x)\right] = \frac{h^2 f''(x)}{2} + O\left(h^2\right).$$

Further show that the variance is

$$\mathbb{V}\left(\hat{f}(x)\right) = \frac{f(x)}{2Nh} + O\left(\frac{1}{N}\right)$$

and consequently mean square error is

$$\mathbb{E}\left[\left(\hat{f}(x) - f(x)\right)^2\right] = \frac{h^4 \left[f''(x)\right]^2}{4} + \frac{f(x)}{2Nh} + O\left(\frac{1}{N}\right) + O\left(h^4\right).$$

[d] Derive an approximate mean squared error minimizing choice of h. Discuss how this choice reflects a "bias-variance trade-off". Can you suggest a feasible way of choosing h in practice?

[2] Consider the statistical model

$$\mathbf{Z} \sim N\left(\theta, \frac{\sigma^2}{N} I_K\right)$$

and estimate of $\theta = (\theta_1, \ldots, \theta_K)'$

$$\hat{\theta} = \arg\min_{\theta} \frac{1}{2} \sum_{k=1}^{K} (Z_k - \theta_k)^2 + \lambda \sum_{k=1}^{K} |\theta_k| .$$

[a] Show that

$$\hat{\theta}_{k,\lambda} = \operatorname{sgn}(Z_k) (|Z_k| - \lambda)_+ .$$

Discuss why this estimator is called the soft-threshold estimator. In what settings will this estimator exhibit low risk?

[b] Show that Stein's Unbiased Estimate of Risk (SURE) for this procedure is

$$\hat{R}_{\text{SURE}}(\mathbf{Z}) = \sum_{k=1}^{K} \left\{ \frac{\sigma^2}{N} - \frac{2\sigma^2}{N} \mathbf{1}(|Z_k| \leq \lambda) + \min\left(Z_k^2, \lambda^2\right) \right\} .$$

Describe how to use $\hat{R}_{\text{SURE}}(\mathbf{Z})$ to choose λ in practice.

[c] Now consider the following curved soft-threshold estimator

$$\hat{\theta}_{k,\lambda} = \begin{cases} -(Z_k + \lambda)^2 & Z_k < -\lambda \\ 0 & -\lambda \leq Z_k \leq \lambda \\ (Z_k - \lambda)^2 & Z_k > \lambda \end{cases}$$

Can you constructed a penalized least squares representation of this estimator?

[d] Calculate $\hat{R}_{\text{SURE}}(\mathbf{Z})$ for the curved soft-threshold estimator.

[e] In what settings will this estimator exhibit low risk?

[3] Let $X \sim \text{Uniform}[-1, 1]$ and assume that

$$Y = -\frac{2}{3} + X^2 + V, \;\; V|\, X \sim \mathcal{N}\left(0, \sigma^2\right) \tag{1}$$

[a] Calculate $\mathbb{E}[Y|\, X]$

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

[c] Calculate $\mathbb{E}[Y]$

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

[e] Let $U = Y - \mathbb{E}^*[Y|\, X]$. Show that $\mathbb{C}(U, X) = 0$. Give an intuitive explanation for this result.

[f] Find a function $g(X)$ such that that $\mathbb{C}(U, g(X)) = \mathbb{V}\left(X^2\right)$. Give an intuitive explanation for your answer.

[g] Describe how your answers in (d) to (f) would change if (1) held but now $X \sim \text{Uniform}[0, 2]$. You may find it helpful to sketch a figure.

[4] Consider two large populations of households and firms, both of which may costlessly migrate across $i = 1, \ldots, N$ different localities. Conditional on choosing to reside in location i, households $t = 1, \ldots, M_i$, choose the amount of unimproved land, l, and composite commodity, x, to consume in order to maximize utility:

$$\max_{l,x} Q_i l^{\gamma} x^{1-\gamma} \ s.t.\ H_{it}W_i \leq R_i l + x, \tag{2}$$

where Q_i, W_i and R_i respectively denote the "quality of life", wage rate per efficiency unit of labor, and rent per unit of unimproved land, in locality i. The price of the composite commodity is fixed on world markets and normalized to one. A household's endowment of efficiency units of labor is given by H_{it} so that their total budget conditional on residence in locality i is $H_{it}W_i$. Let L_{it} and X_{it} denote the household's utility-maximizing land and composite commodity consumption.

Equilibrium requires that households are indifferent between locations, or that the indirect utility associated with one efficiency unit of labor is constant across communities. Under (2) this condition takes the form

$$V(W,R;Q) = \psi \frac{QW}{R^{\gamma}} = \bar{u},\ \psi = \gamma^{\gamma}(1-\gamma)^{1-\gamma}, \tag{3}$$

where $\bar{u}$ it the common, nationwide, utility level (available to an owner of an efficiency unit of labor).

Firm $f = 1, \ldots, L_i$ in locality i produces X_{if} units of the composite commodity using a constant returns to scale production technology requiring three inputs: (i) efficiency units of labor, n ; (ii) units of improved land, l , and (iii) capital, k. Labor and land prices in a locality are determined in equilibrium, while the cost of capital is constant and equal to σ. The firm's problem is to minimize costs:

$$\min_{n,l,k} W_i n + R_i l + \sigma k,\ s.t.\ A_i n^{\alpha} l^{\gamma} k^{1-\alpha-\beta} = X_{if}, \tag{4}$$

where A_i is locality-specific total factor productivity. Free entry ensures that profits are zero in equilibrium. The zero profit condition requires that unit costs equal the normalized price of the composite good:

$$C(W,R,A) = \xi \frac{W^{\alpha}R^{\beta}}{A} = 1,\ \xi = \left(\frac{1}{\alpha}\right)^{\alpha}\left(\frac{1}{\beta}\right)^{\beta}\left(\frac{1}{1-\alpha-\beta}\right)^{1-\alpha-\beta}\sigma^{1-\alpha-\beta} \tag{5}$$

[a] Use (3) and (5) to show that the equilibrium wage and rent in locality i is given by

$$\begin{aligned} \ln W_i &= \kappa_W + \frac{\gamma}{\alpha\gamma+\beta}\ln A_i - \frac{\beta}{\alpha\gamma+\beta}\ln Q_i \\ \ln R_i &= \kappa_R + \frac{1}{\alpha\gamma+\beta}\ln A_i + \frac{\alpha}{\alpha\gamma+\beta}\ln Q_i, \end{aligned} \tag{6}$$

with

$$\begin{aligned} \kappa_W &= \frac{1}{\alpha\gamma+\beta}(\beta\ln\bar{u} - \beta\ln\psi - \gamma\ln\xi) \\ \kappa_R &= \frac{1}{\alpha\gamma+\beta}(\alpha\ln\bar{u} - \alpha\ln\psi - \ln\xi). \end{aligned}$$

HINT: You do not need to verify the form of the intercepts in (6).

[b] Assume that A_i and Q_i vary independently. Let $\phi = \frac{\gamma^2\mathbb{V}(\ln A_i)}{\gamma^2\mathbb{V}(\ln A_i)+\beta^2\mathbb{V}(\ln Q_i)}$. Show that the coefficient on $\ln W_i$ in the (mean squared error minimizing) linear predictor of $\ln R_i$ onto a constant and $\ln W_i$ equals

$$b = \frac{1}{\gamma}\phi - \frac{\alpha}{\beta}(1-\phi).$$

Interpret this expression. Under what conditions is b positive? Negative? Why?

[c] Assume that the $\gamma = 1/3$, $\alpha = 3/5$, and $\beta = 1/5$. A macro economist, with considerable central banking experience, asserts that the "quality-of-life" differences across cities are "overblown". An econometrician, with almost no "real world" experience, asserts that "firms are equally unproductive in all places". An urban economist claims that "actually both are equally important".

Using a random sample of US cities you compute an estimate of b equal to -3. Who is correct? The macro-economist, econometrician or urban economist?

[d] It turns out that you made a Python coding error and that the correct estimate of b is zero. Who do you believe to be correct now?

[e] Use (6) to solve for $\ln A_i$ and $\ln Q_i$ (again you may ignore constants). Use the parameter values given in (c) to simplify your expressions.

[f] You observe mean January temperature for each city in your sample. Can you compute consistent estimates of the semi-elasticity of total factor productivity and quality of life with respect to temperature? Describe your procedure.

[5] 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

$$\mathbb{E}\left[\left(\hat{\mu}\left(c_0, c_1\right) - \mu\right)^2\right] = \frac{c_0^2\sigma^2}{N} + \frac{c_1^2 4\sigma^2}{N} + \left(1 - c_0 - c_1\right)^2 \mu^2. \quad (7)$$

[b] Derive the oracle estimator (within the family) which minimizes (7).

[c] Show that

$$\hat{R}\left(c_0, c_1\right) = \frac{c_0^2\sigma^2}{N} + \frac{c_1^2 4\sigma^2}{N} + \left(1 - c_0 - c_1\right)^2 \frac{1}{2}\left\{\bar{Y}_0^2 + \bar{Y}_1^2 - \frac{5\sigma^2}{N}\right\} \quad (8)$$

is an unbiased estimate of of (8). 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 (8) to construct an implementable estimator of μ.

[6] 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

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

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

[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 $\mathbb{E}\left[Y|\, X = x\right]$ 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 statistics 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.

[d] After attending the labor lunch you now believe that $\mathbb{E}\left[Y|\, X = x\right]$ is linear in x but with discrete jumps at $X = 12$ and $X = 16$. Describe, in detail, how you would evaluate this new hypothesis?

[7] You are given a random sample from South Africa in the late 1980s. Each record in this sample includes, Y, an individual's log income at age 40, X the log permanent income of their parents, and D a binary indicator equaling 1 if the respondent is White and zero if they are Black. Let the best linear predictor of own log income at age forty given parents' log permanent income and own race be

$$\mathbb{E}^*\left[Y|\, X, D\right] = \alpha_0 + \beta_0 X + \gamma_0 D.$$

[a] Let $Q = \Pr\left(D = 1\right)$, assume that $\mathbb{V}\left(X|\, D = 1\right) = \mathbb{V}\left(X|\, D = 0\right) = \sigma^2$ and recall the analysis of variance formula $\mathbb{V}\left(X\right) = \mathbb{V}\left(\mathbb{E}\left[X|\, D\right]\right) + \mathbb{E}\left[\mathbb{V}\left(X|\, D\right)\right]$. Show that

$$\mathbb{V}\left(X\right) = Q\left(1 - Q\right)\left\{\mathbb{E}\left[X|\, D = 1\right] - \mathbb{E}\left[X|\, D = 0\right]\right\}^2 + \sigma^2.$$

[b] Let $\mathbb{E}^*\left[D|\, X\right] = \kappa + \lambda X$. Show that

$$\lambda = \frac{Q\left(1 - Q\right)\left\{\mathbb{E}\left[X|\, D = 1\right] - \mathbb{E}\left[X|\, D = 0\right]\right\}}{Q\left(1 - Q\right)\left\{\mathbb{E}\left[X|\, D = 1\right] - \mathbb{E}\left[X|\, D = 0\right]\right\}^2 + \sigma^2}.$$

[c] Assume that $\beta_0 = 0$. Show that in this case $\gamma_0 = \mathbb{E}\left[Y|\, D = 1\right] - \mathbb{E}\left[Y|\, D = 0\right]$.

[d] Let $\mathbb{E}^*\left[Y|\, X\right] = a + bX$. Maintaining the assumption that $\beta_0 = 0$ show that

$$b = \frac{Q\left(1 - Q\right)\left\{\mathbb{E}\left[Y|\, D = 1\right] - \mathbb{E}\left[Y|\, D = 0\right]\right\}\left\{\mathbb{E}\left[X|\, D = 1\right] - \mathbb{E}\left[X|\, D = 0\right]\right\}}{Q\left(1 - Q\right)\left\{\mathbb{E}\left[X|\, D = 1\right] - \mathbb{E}\left[X|\, D = 0\right]\right\}^2 + \sigma^2}.$$

[e] Let $Q\left(1 - Q\right) = 1/10$, $\sigma^2 = 3/10$ and $\mathbb{E}\left[Y|\, D = 1\right] - \mathbb{E}\left[Y|\, D = 0\right] = \mathbb{E}\left[X|\, D = 1\right] - \mathbb{E}\left[X|\, D = 0\right] = 3$. Provide a numerical value for $\mathbb{V}\left(X\right)$ and b.

[f] On the basis of β_0 a member of the National Party argues that South Africa is a highly mobile society. One the basis of b a member of the African National Congress argues that it is a highly immobile one. Comment on the relative merits of these two assertions.

[8] Consider the following joint probability density function for x and y

$$f_{X,Y}(x,y) = \begin{cases} \frac{3}{11}\left(x^2+y\right), & (x,y) \in [0,2]\times[0,1] \\ 0 & (x,y) \notin [0,2]\times[0,1] \end{cases}$$

[a] Show that $f_{X,Y}(x,y)$ is a valid probability density function.

[b] Compute the conditional density $f_{Y|X}(y|x)$.

[c] Compute the conditional expectation function $\mathbb{E}[Y|X=x]$.

[d] Compute the linear predictor $\mathbb{E}^*[Y|X=x]$.

[e] Consider a joint distribution with a conditional that coincides with the one derived in part [b] above, but where the marginal distribution of X is uniform on $[0,2]$. Explain, qualitatively, how this change would affect your answers in parts [c] and [d] above?

[9] For a random draw from the population of US workers, let Y equal log earnings and X be a binary indicator taking a value of one if the worker is black and zero otherwise. Let $\{(Y_i,X_i)\}_{i=1}^N$ be a random sample of size N. Let N_1 denote the number of sampled units that are black (i.e., $X=1$) and $N_0 = N - N_1$ the number that are non-black. Assume that

$$Y_i = \alpha_{\tau 0} + \beta_{\tau 0} X_i + U_i$$

with

$$Q_{U|X}(\tau|X) = 0.$$

Let a and b be a candidate values for 'the truth' (i.e., $\alpha_{\tau 0}$ and $\beta_{\tau 0}$). Let $u_\tau^1(a,b)$ be the τ^{th} quantile of $U(a,b) = Y - a - bX$ given $X=1$. Let $u_\tau^0(a,b)$ be the corresponding τ^{th} quantile given $X = 0$. Let $R_1(a,b),\ldots,R_{N_1}(a,b)$ denote the N_1 order statistics of $U(a,b)$ in the $X_i = 1$ subsample. Let $S_1(a,b),\ldots,S_{N_0}(a,b)$ denote the N_0 corresponding statistics from the $X_i = 0$ subsample.

[a] Let $\tau = 0.25$. Interpret (in words) the parameters $\alpha_{\tau 0}$ and $\beta_{\tau 0}$. What is true about the distribution of black versus non-black earnings if $\beta_{\tau 0} = 0$? Let $\tau = 0.75$. How does your interpretation of $\alpha_{\tau 0}$ and $\beta_{\tau 0}$ change or not change?

[b] What is the τ^{th} quantile of $U(\alpha_{\tau 0},\beta_{\tau 0})$ given, respectively, $X=1$ and $X=0$?

[c] Assume that $a = \alpha_{\tau 0}$ and $b = \beta_{\tau 0}$. Let $j/(N_1+1) < \tau \leq (j+1)/(N_1+1)$. Before looking at your sample you are asked to guess the value of $(R_j(a,b) + R_{j+1}(a,b))/2$. What is your guess? Justify your answer.

[d] Let $N_1 = 3$ and $N_0 = 3$ and $\tau = 0.25$ (for this part of the problem only). Consider the order statistic intervals $[R_1(a,b), R_3(a,b)]$ and $[S_1(a,b), S_3(a,b)]$. Assume $a = \alpha_{\tau 0}$ and $b = \beta_{\tau 0}$; what is the ex ante probability that each of these intervals contain zero? What is the ex post probability they contain zero? Be sure to explain your work.

[e] Let a and b be some candidate intercept and slope values. Describe, in detail, an estimate of $u_\tau^0(a,b)$ and $u_\tau^1(a,b)$? Denote these estimates by, respectively, $\hat{u}_\tau^1(a,b)$ and $\hat{u}_\tau^0(a,b)$.

[f] Describe how to construct an approximate 95 percent confidence interval for $u_\tau^0(a,b)$ and $u_\tau^1(a,b)$?

[10] 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	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$

[11] Consider the population of married men. Let Y denote log earnings for a generic random draw from this population, X his years of completed schooling and W the schooling of his spouse. Assume that the conditional mean of own log earnings given own and spouse's schooling is

$$\mathbb{E}[Y|X,W] = \alpha_0 + \beta_0 X + \gamma_0 W,$$

while the best linear predictor of spouse's schooling given own schooling is

$$\mathbb{E}^*[W|X] = \delta_0 + \zeta_0 X.$$

You may assume that the joint distribution of (W,X,Y) is such that these objects are well-defined. You may assume that all the slope coefficients in the two equations above are positive.

[a] Show that $\zeta_0 = \rho_{WX}\frac{\sigma_W}{\sigma_X}$, with ρ_{WX} the correlation of W with X, and σ_W and σ_X respectively the standard deviation of W and X. Further show that $\delta_0 = \mu_W - \rho_{WX}\frac{\sigma_W}{\sigma_X}\mu_X$ with μ_W and μ_X denoting the population means of W and X.

[b] Using your answers in [a] above, as well as the form of $\mathbb{E}[Y|W,X]$, provide an expression for $\mathbb{E}^*[Y|X]$.

[c] Consider another population of married men where $F_{Y|W,X}(y|W=w,X=x)$, $F_W(w)$ and $F_X(x)$ coincide with those for the population described above, but where $F_{W,X}(w,x)$ differs. Assume that in this alternative population $\rho_{WX}=0$. Solve for $\mathbb{E}[Y|X,W]$, $\mathbb{E}^*[W|X]$ and $\mathbb{E}^*[Y|X]$. Use the notation established in parts [a] and [b] to formulate your answer.

[d] Assume that $F_W(w)$ and $F_X(x)$ are identical and that marriage is homogamous in terms of education so that $W=X$ for all couples (i.e., individuals choose partners with identical levels of education). Show that in this world $\rho_{WX}=1$. Solve for $\mathbb{E}[Y|X,W]$, $\mathbb{E}^*[W|X]$ and $\mathbb{E}^*[Y|X]$. Use the notation established in parts [a] and [b] to formulate your answer.

[e] Compare the form of $\mathbb{E}^*[Y|X]$ in the original population with that in the two alternative populations of parts [c] and [d]. In which population does log earnings rise most steeply with years of schooling? Provide some intuition for your answer (5 sentences).

[f] Assume that schooling is binary valued, taking on the values 0,1. Let R_W be a 2×1 vector equal to $(1,0)'$ if $W = 0$ and $(0,1)'$ if $W = 1$. Let S_X be the analogous 2×1 vector defined using X. Let $T_{WX} = (R_W \otimes S_X)$ and

$$\mathbb{E}^* \left[Y | T_{WX} \right] = T'_{WX}\pi,$$

where a constant is *not* included and $\pi = (\pi_{00}, \pi_{01}, \pi_{10}, \pi_{11})'$. Show that

$$\pi_{jk} = \mathbb{E} \left[Y | W = j, X = k \right].$$

[g] Consider the null hypothesis that $\mathbb{E}[Y|W, X] = \alpha_0 + \beta_0 X + \gamma_0 W$. Maintaining this null find an explicit expression for each component of π in terms of α_0, β_0 and γ_0. Express this null in the form $C\pi = c$ for some matrix of constants C and vector of constants c.

[h] Let $W = 1$ if a wife has completed primary school and zero otherwise, let $X = 1$ if a husband has completed primary school and zero otherwise. A least squares fit, loosely based on data from Brazil, of log husband's earnings on T_{WX} as defined in [f] using a random sample of size $N = 50,000$ yields point estimate of

$$\hat{\pi} = \begin{pmatrix} 5.50 \\ 6.00 \\ 5.00 \\ 7.00 \end{pmatrix}$$

with an estimated asymptotic variance-covariance matrix of

$$\hat{\Lambda} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 7 & 0 & 0 \\ 0 & 0 & 6 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

Can you reject the null hypothesis (at the $\alpha = 0.05$ level) formulated in part [g] on the basis of this sample? 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.