Final Exam for Quantitative Methods

Economics 492L, Fall 2003

Maxwell B. Stinchcombe

This take-home exam has nine equally weighted questions. Your answers are due before noon, Monday, December 15, 2003, in the Economics Department main office.

I expect you to work alone. Please keep answers to separate questions on separate pieces of paper.

The only acceptable reference works are the notes and textbooks for this class, your micro class, or your macro class. My guess is that you will want to use Casella and Berger several times.

If a numbered part of a problem makes a statement, your job is to prove it if it is true, and to find a counter-example if it is not.

A: Suppose that we observe n values of y, arranged in an $n \times 1$ vector $\mathbf{Y}$, and n values of x_i, $i = 1, \ldots, k$, arranged in an $n \times k$ matrix $\mathbf{X}$. We assume that $n > k$ and that the columns of $\mathbf{X}$ are independent. The β that makes

$$f(\beta) := (\mathbf{Y} - \mathbf{X}\beta)'(\mathbf{Y} - \mathbf{X}\beta) = e'_\beta e_\beta$$

as small as possible is $\widehat{\beta}_{LS} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$.

1. Let $e_1 = (1, 0, \ldots, 0)'$. Solve the minimization problem

$$\min_\beta (\mathbf{Y} - \mathbf{X}\beta)'(\mathbf{Y} - \mathbf{X}\beta) \text{ subject to } e'_1\beta = 0,$$

   and let β_1 denote the answer. Note that $e'_1\beta = 0$ iff the first component of β is equal to 0. Express β_1 in terms of $\widehat{\beta}_{LS}$. [It should be a simple expression.]
2. Interpret your previous answer in terms of M, the column span of $\mathbf{X}$.
3. Show that $e'_{\beta_1}e_{\beta_1} \geq e'_{\beta_{LS}}e_{\beta_{LS}}$. [If the inequality is too large, we'd reject the hypothesis that $e'_1\beta = 0$.]
4. Explicitly give $e'_{\beta_1}e_{\beta_1}$ and $e'_{\beta_{LS}}e_{\beta_{LS}}$ when $\mathbf{X}$ is an $n \times 1$ column of 1's. Relate this to the problem of finding the mean of collection of numbers.

B: (This problem uses the notation from the previous problem.) We suppose that there are n observations stacked as above, and that rearranging the rows into $\epsilon_i = Y_i - X_i\beta$ gives a collection, ϵ_i, $i = 1, \ldots, n$, of independent, mean 0, variance $\sigma^2 > 0$, random variables. The parameters for the distribution of the ϵ_i is (β, σ^2), and the random vector of ϵ_i's is denote $\boldsymbol{\epsilon}$.

We use $\widehat{\beta}$ for $(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$. This problem asks you to find $E\,\mathbf{e}'\mathbf{e}$, where the least squares residuals (the errors made in fitting the data to a linear relation) are defined as $\mathbf{e} = (\mathbf{Y} - \mathbf{X}\widehat{\beta})$.

Note carefully the distinction between $\mathbf{e}$ and $\boldsymbol{\epsilon}$.

The trace of a square matrix $\mathbf{M} = (\mathbf{M})_{ij} = (m_{ij})$, $i, j = 1, \ldots, n$, is defined as the sum of the diagonal elements, $\mathbf{tr}(\mathbf{M}) = \sum_i m_{ii}$. If $\mathbf{M}$ is a 1×1 matrix, it is a scalar, and the trace of a scalar is the scalar itself.

1. $\mathbf{tr}(\mathbf{M}_1 + \mathbf{M}_2) = \mathbf{tr}(\mathbf{M}_1) + \mathbf{tr}(\mathbf{M}_2)$.
2. When $\mathbf{A}$ is $n \times m$ and $\mathbf{B}$ is $m \times n$, $\mathbf{tr}(\mathbf{AB}) = \mathbf{tr}(\mathbf{BA})$, that is, matrixes can be commuted under the trace operator.
3. $E\,\boldsymbol{\epsilon}\boldsymbol{\epsilon}' = \sigma^2\mathbf{I}_n$ where $\mathbf{I}_n$ is the $n \times n$ identity matrix.
4. For any $n \times n$ matrix $\mathbf{M}$, $\boldsymbol{\epsilon}'\mathbf{M}\boldsymbol{\epsilon}$ is a scalar, and $E\,\boldsymbol{\epsilon}'\mathbf{M}\boldsymbol{\epsilon} = \sigma^2\mathbf{tr}(\mathbf{M})$.
5. $\mathbf{e}'\mathbf{e} = \boldsymbol{\epsilon}'(\mathbf{I}_n - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')\boldsymbol{\epsilon}$.
6. $E\,\mathbf{e}'\mathbf{e} = \sigma^2(n - k)$ so that $s^2 = \frac{\mathbf{e}'\mathbf{e}}{(n-k)}$ is an unbiased estimator of σ^2.
7. Explicitly give the matrix $(\mathbf{I}_n - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')$ and its trace when $\mathbf{X}$ is an $n \times 1$ column of 1's.

C: The exponential(β) density function is $f(y|\beta) = \frac{1}{\beta}e^{\frac{-y}{\beta}}1_{[0,\infty)}(y)$. This density is often used as a model for the length of life of memoryless physical systems. Suppose that Y has the exponential distribution just given so that $E\,Y = \beta$.

1. Show that Y is memoryless — that is, show that for $a, b > 0$, $P(Y > a + b|Y > a) = P(Y > b)$.
2. Consider the random variable $U = \sqrt{Y}$. Give the density for U and find $E\,U$.
3. The memorylessness of Y implies that $E\,(Y|Y > a) = a + \beta = a + E\,Y$. For $a > 0$, is $E\,(U|U > a) < a + E\,U$ or $E\,(U|U > a) > a + E\,U$? Explain.

D: Suppose that $(X_1, \ldots, X_n)$ is iid $f(x|\theta) = \theta x^{\theta-1}1_{[0,1]}(x)$, $\theta > 0$.

1. The density given above has a name, what is it?
2. Show that the density is a one-parameter exponential family.
3. Show that $\sum_i \ln(X_i)$ is sufficient for θ.
4. Find the maximum likelihood estimator of θ.
5. Find the method of moments estimator of θ.
6. Find the distribution of $Y_i := \ln(X_i)$, and the distribution of $\sum_i Y_i$.
7. Find an unbiased estimator with minimal variance. [This one is difficult.]

E: There are n stocks, with prices $p \in \mathbb{R}^n_{++}$. Their random rates of return are the vector $\mathbf{X} \in \mathbb{R}^n$ (so that d dollars in stock i gives a return $d \cdot X_i$). $E\,\mathbf{X} = \mu \gg 0$ and $\mathbf{Var}(\mathbf{X}) = \Sigma$ where Σ is the positive definite $n \times n$ matrix with ij'th entry $\Sigma_{ij} = \mathbf{Cov}(X_i, X_j)$. The choice problem is how much to invest in stock i. Let $q \in \mathbb{R}^n$ denote a vector of investment levels, that is, a portfolio. $E\,q \cdot \mathbf{X} = q'\mu$ is the mean return of the portfolio q. We assume that a portfolio having mean r and variance v is valued using the utility function $U = r - \frac{\beta}{2}v$, $\beta > 0$.

1. Mean-variance utility functions with $\beta > 0$ violate first order stochastic dominance — there are random variables R and S with R first order stochastically dominating S having $U(R) < U(S)$. [Despite this drawback, mean-variance utility functions have been widely used, and the rest of the problem asks you to use them.]
2. $\mathbf{Var}(q'\mathbf{X}) = q'\Sigma q$.
3. Solve the unconstrained investment problem $\max_q \mu' q - \frac{\beta}{2} q'\Sigma q$ in terms of μ, Σ, and β. Give an expression for $\partial q_i^*/\partial \mu_i$ and give the intuition for its sign and why this derivative depends as it does on β and σ_i^2.
4. Suppose now that there is a budget b. Solve the constrained optimization problem $\max_q \mu' q - \frac{\beta}{2} q'\Sigma q$ subject to $p'q \leq b$ assuming the constaint is binding. Give an expression for $\partial q_i^*/\partial \mu_i$ in this constrained problem and give an intuition for its sign.

F: One definition of X being riskier than Y is that $X = Y + R$ where R is a random variable satisfying $E(R|Y) = 0$.

1. Show that $\mathbf{Var}(X) \geq \mathbf{Var}(Y)$.
2. Show that $E\,u(X) \leq E\,u(Y)$ for any concave u.

G: Suppose that $X_1, \ldots, X_n$ are i.i.d. with the uniform distribution on $(-\theta, \theta)$ for some unknown $\theta \in \Theta = \mathbb{R}_{++}$.

1. Give the likelihood function $L(X_1, \ldots, X_n : \theta)$ and the maximum likelihood estimator of θ.
2. Find the 95% confidence interval for $\widehat{\theta}_{MLE}$.
3. Calculate the bias of the MLE.
4. Describe the $\alpha = 0.01$ one-sided test for the null hypothesis $\theta \geq 7$.
5. Calculate the power of the test as a function of $\theta < 7$.
6. Give a method of moments estimator of θ.
7. Give a sufficient statistic for θ and an unbiased estimator with minimal variance.

H: Suppose that $\mathbf{X} = (X_1, \ldots, X_n)$ is iid with density $\frac{1}{2}e^{-|x-\theta|}$, $\theta \in \mathbb{R}$.

1. The density given above has a name, what is it?
2. Describe the likelihood ratio test for the null hypothesis $H_0 : \theta = 0$. [It is possible to express this in terms of the averages of some data-defined absolute values.]
3. Give the associated power function.

I: In the mid 1700's, British Spice Importers regularly received shipments of barrels peppers and other tropical spices meant to make British food minimally palatable. (Historical wisdom has it that the British conquered the world looking for a good meal. Economists naysay these ideas, pointing at the early development of a stock market and banking system as the main British advantages, but what do they know anyway?) Over one 18 month period, 2% of these barrels have been substandard (read, too full of dead bugs and rat droppings for even the notoriously insensitive British public to stomach). After the 18 month period, the supplier switched production to a new island, and BSI is concerned that quality might have deteriorated (even further). Quality control spice tasters (with the unenviable job of) randomly sampling $n = 500$ of the barrels in the next shipments find that 21 of them are substandard.

1. Calculate the 95% confidence interval and the 99% confidence interval for the proportion of substandard barrels from the new island.
2. Is the null hypothesis that quality has not deteriorated rejected at $\alpha = 0.05$? At $\alpha = 0.01$?
3. In this example, the p-value is the probability that the sample value would be as large as it is if H_0 is true. For data giving a p-value less than α, we reject the null hypothesis with confidence $(1 - \alpha)$, and for giving a p-value greater than α, we accept the null hypothesis with confidence $(1 - \alpha)$. What is the p-value for this data?
4. Since statistics had not been invented in the 1700's, give a verbal summary of the evidence, without using any modern statistical jargon, that the new island is a worse source, including an argument about the kinds of mistakes one might be making.