The University of Chicago
Department of Economics
ECONOMICS 209
HONORS ECONOMETRICS
PROBLEM SET #1

Instructions: This problem set contains 15 questions. Some of them review material that you should already know, others are designed to help you understand the concepts we review in class. The more difficult questions extend the theory that we study. So far we have covered matrix algebra and a review of statisitics including maximum likelihood estimation. The problem set will focus on these two topics. You have one week to complete the problem set: it is due on Wednesday, January 23. You are encouraged to work in groups and to discuss the problems, especially the harder ones. Have fun! Please be neat. Clarity helps me to give you partial credit. Your answers should be clearly labelled.

Matrix Algebra

Question MA1: Computation using matrices.

Let X denote a 5x2 matrix and y denote a 5x1 vector. Let these be given by:

$$X = \begin{bmatrix} 1 & 2 \\ 1 & 4 \\ 1 & 3 \\ 1 & 5 \\ 1 & 2 \end{bmatrix}$$

and

$$y = \begin{bmatrix} 14 \\ 17 \\ 8 \\ 16 \\ 3 \end{bmatrix}$$

respectively. Do the following computations, please keep the fractions.

a. $Q = X'X$; $\det(Q)$; Q^{-1}.

b. $A = (X'X)^{-1} X'$

c. $N = X(X'X)^{-1} X'$

d. Is N full rank?

e. Show that N is idempotent.

f. $M = I - N$

g. Is M full rank?

Question MA2: Manipulating regression matrices.

Let X denote an (nxk) matrix, whose rank is k and let

$$\begin{aligned} Q &= X'X \\ A &= (Q)^{-1} X' \\ N &= XA \\ M &= I - N \end{aligned}$$

Also, let

$$\begin{aligned} b &= Ay \\ \widehat{y} &= Ny \\ e &= My \end{aligned}$$

Show as concisely as possible that:

a. $AN = A$, $AM = 0$, $MN = 0$, $NM = 0$.

b. $NX = X$, $MX = 0$.

c. $N\widehat{y} = \widehat{y}$, $Ne = e$.

d. $M\widehat{y} = 0$, $Me = 0$.

e. $X'\widehat{y} = X'y$

f. $y'\widehat{y} = y'Xb = b'X'y = b'Qb = \widehat{y}'\widehat{y}$.

g. $e'e = y'My = y'y = \widehat{y}'\widehat{y}$.

Question MA3: Constructing data matrices.

Suppose that we interview $n = 100$ people and we ask them whether they are males or females. Suppose that n_F is the number of female responses and n_M is the number of male responses. Obviously, $n_F + n_M = n$. We create three variables. Let x_1 denote a variable that is recorded as a 1 for everybody. Let x_2 denote a variable that is recorded as 1 if the respondent is male and 0 if the respondent is female. Let x_3 denote a variable that is recorded as 1 if the respondent is female and 0 if the respondent is male.

a. Write out a $(1x3)$ row vector that records the typical male response.

b. Write out a $(1x3)$ row vector that records the typical female response.

c. Write out the entire matrix in partitioned form, if we partition across variables. Be sure to be very explicit. Label it X_1.

d. Write out the entire matrix in partitioned form, if we partition across observations (we group males and females). Be sure to be very explicit. Label it X_2.

e. Compute $X'X$ in both cases.
f. Is X a full column rank matrix?

Statistics

Question S1: Bias of the sample standard deviation.
Show that:

$$s^2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2$$

is a biased estimator of the population variance σ^2 and provide an unbiased estimator. That is, show that:

$$E\left[s^2\right] \neq \sigma^2.$$

Question S2: Manipulating variance and expectation operators.
Let X and Y denote continuous random variables over an infinite support and let a, b, c, d denote constants. Using the definitions of expectation, variance and covariance, show that the following equalities hold:

$$\begin{aligned}
E[a] &= a \\
E[aX] &= aE[X] \\
E[X+b] &= E[X]+b \\
Var[a] &= 0 \\
Var[aX] &= a^2 Var[X] \\
Var[X+b] &= Var[X] \\
Cov[aX+b, cY+d] &= acCov[X,Y] \\
Cov[X,X] &= Var[X]
\end{aligned}$$

Question S3: Computing moments.
Consider the continuous bivariate probability distribution whose joint probability density function is given by:

$$f(x,y) = \frac{3}{11}\left(x^2 + y\right)$$

for $x \in [0,2]$ and $y \in [0,1]$.
a. Derive the marginal pdf of X.
b. Derive $f_Y(y)$, the marginal pdf of y.
c. For $x \in [0,2]$ derive $g_Y(y|x)$ the conditional pdf of Y given X.

d. Compute $E[X], E[Y], E[X^2], E[Y^2], E[XY], Var[X], Var[Y], Cov[X,Y]$.

e. Find the conditional expectation function: $E[Y|X]$.

Question S4: The Law of Iterated Expectations.

The law of iterated expectations is a very important rule that we will use extensively. In this exercise you are asked to prove it. Let X and Y denote random variables. Let $Z = h(X,Y)$. Show that

$$E[Z] = E_X[E[Y|X]].$$

where E_X denotes the expectation over the marginal of X. Hint: Remember the identity that relates joint pdfs to conditional pdfs.

Question S5: Independence.

Let X and Y denote two random variables. Show that:

$$Var[X+Y] = Var[X] + Var[Y] + 2Cov[X,Y].$$

Show that under the assumption of stochastic independence

$$Cov[X,Y] = 0.$$

In this case what does $Var[X+Y]$ equal?

Question S6: Statistics using matrices.

Now, let us look at some statistical manipulations using vector notation. Suppose that an x is a random $nx1$ vector with expectation: expectation $E[x] = \mu$ and variance $Var[x] = \Sigma$. Let $y = g + Hx$ be a linear function of x. Note that H is an (nxn) matrix of constants.

a. Compute $E[y]$, $Var[y]$, $E[y'y]$, $Cov[y,x]$, and $Cov[x,y]$.

b. Repeat part a, under the assumption that:2

$$\mu = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \Sigma = \begin{bmatrix} 2 & 1 & 2 \\ 1 & 3 & 1 \\ 2 & 1 & 4 \end{bmatrix}, g = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, H = \begin{bmatrix} 1 & 2 & 1 \\ -1 & 2 & 1 \end{bmatrix}$$

Maximum Likelihood Estimation

Question MLE1: The power density.

Let $X_1, ..., X_n$ represent a sample of size n from a population with density

$$f(x_i;\theta) = \theta x_i^{\theta-1}$$

for $x \in (0,1)$. Let the parameter space be $\theta \in (0,\infty)$.

a. What is the likelihood function?
b. Find $\widehat{\theta}^{MLE}$, the maximum likelihood estimator of θ.
c. Find the information matrix.
d. Is $\widehat{\theta}^{MLE}$ an unbiased estimator θ?
e. Find the asymptotic distribution of $\widehat{\theta}^{MLE}$ in terms of θ and n only.
Question MLE2: A simple normal example.
Let $X_1, ..., X_n$ represent a sample of size n from a population
with a normal density with mean θ and variance 1 :

$$f(x_i; \theta) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x_i - \theta)^2}$$

Note that only the mean is unknown. We say that the sample is independently and identically distributed, iid for short, $N(0, 1)$.
a. What is the support of this density?
b. What is the parameter space?
c. What is the likelihood function?
d. Find $\widehat{\theta}^{MLE}$, the maximum likelihood estimator of θ.
e. Find the information matrix.
f. Is $\widehat{\theta}^{MLE}$ an unbiased estimator of θ?
g. Find the asymptotic distribution of $\widehat{\theta}^{MLE}$ in terms of θ and n only.
Question MLE3: The exponential density.
Let $X_1, X_2, ..., X_n$ represent a sample of size n from a population
with density

$$f(x_i; \theta) = \frac{1}{\theta} e^{-\frac{1}{\theta} x_i}$$

for $x \in (0, \infty)$. Let the parameter space be $\theta \in (0, \infty)$.
a. What is the likelihood function?
b. Find $\widehat{\theta}^{MLE}$, the maximum likelihood estimator of θ.
c. Find the information matrix.
d. Is $\widehat{\theta}^{MLE}$ an unbiased estimator of θ?
e. Find the asymptotic distribution of $\widehat{\theta}^{MLE}$ in terms θ and n only.
Question MLE4: Challenging ourselves.
Let $X_1, X_2, ..., X_n$ represent a sample of size n from a population
with density

$$f(x_i; \theta) = \frac{1}{\theta}$$

for $x \in (0, \theta]$. Let the parameter space be $\theta \in (0, \infty)$.

a. What is the likelihood function?

b. Find $\widehat{\theta}^{MLE}$ the maximum likelihood estimator of θ. Hint: Pay close attention to the support and parameter spaces.

c. Is $\widehat{\theta}^{MLE}$ an unbiased estimator of θ?