Midterm Exam 2010, Econometrics, Questions from Mark Watson

Problem 1 (20 points): The density of X is $f(x) = 4x^3$ for $0 < x < 1$ and zero elsewhere.

(a) (5 points) Find $P(0.5 < x < 1)$.

(b) (5 points) Find $E(X)$.

(c) (5 points) Find the variance of X.

(d) (5 points) Let $Y = ln(X)$. Derive the probability density of Y.

Problem 2 (20 points): Y_i, $i = 1, \dots, n$ are *iid* Bernoulli random variables with $P(Y_i = 1) = p$.

(a) (10 points) Show that $\bar{Y}(1-\bar{Y}) \xrightarrow{p} Var(Y)$, where $Var(Y)$ denotes the variance of Y.

(b) (10 points) I am interested in H_o: $p = 0.5$ vs. H_a: $p > 0.5$. The sample size is $n = 100$, and I decide to reject the null hypothesis if $\sum_{i=1}^{n} Y_i > 55$. Use the central limit theorem to derive the approximate size of the test. (Note: Because you do not have a table of the normal distribution, you may leave your answer in terms of a probability statement about a standard normal random variable.)

Problem 3 (10 points): X, Y, and U are independent random variables: $X \sim N(5,4)$, $Y \sim \chi_1^2$, and U is distributed Bernoulli with $P(U=1)=0.3$. Let $W=UX+(1-U)Y$. Find the mean and variance of W.

Problem 4 (10 points): Suppose Y_i, $i = 1, \dots, 10$ are iid with mean μ and variance σ^2, and let $\bar{Y} = \frac{1}{10}\sum_{i=1}^{10} Y_i$. You want to estimate μ, and suppose that risk is mean squared error, so that $R(\hat{\mu}, \mu) = E[(\hat{\mu} - \mu)^2]$. Let $\hat{\mu} = 2\bar{Y}$. Show that $\hat{\mu}$ is inadmissable.

Honoré Questions

Problem 5 (31 points)

The regression output at the end of this test shows the results of two of wage regressions. The observations are individuals and the dependent variable is log-wage (`lwage`).

In addition to a constant, the explanatory variables in Regression 1 are: the number of years of education (`ed`), a dummy variable for being African American (`black`), a dummy variable for living in a metropolitan area (`smsa`), a dummy variable for living in the South (`so`), the number of years of experience (`exp`) and the square of the number of years of experience (`expsq`).

Regression 2 also includes `exso=exp*so` and `exso2=expsq*so` as explanatory variables.

In answering the questions below, assume that the assumptions in Hayashi Chapter 1 are satisfied.

1. (4 points) Consider Regression 1. Construct a 95% confidence interval for the coefficient on education.

2. (4 points) Consider Regression 1. Test whether the coefficient on experience–squared is equal to 0.001.

3. (5 points) Consider Regression 1. What is the estimated derivative of wages with respect to number of years of education?

4. (5 points) Consider Regression 1. What is the estimated derivative of wages with respect to number of years of experience?

5. (10 points) Test whether both the coefficient on experience and the coefficient on experience–squared are the same in the South and in the North.

6. (3 points) The dependent variable, `lwage`, is defined as the natural logarithm of the wage rate. Briefly discuss how regression 1 would change if you used the logarithm of wages to base 10 as the dependent variable.

Problem 6 (17 points)

Consider the model

$$y_i = z_i\delta + \varepsilon_i, \qquad \text{with} \qquad E\left[\varepsilon_i z_i\right] = -1$$

Moreover, assume that there is another variable x_i such that $E\left[\varepsilon_i x_i\right] = 0$. Both z_i and x_i are one–dimensional. Assume that assumptions (3.1)–(3.5) of Hayashi are satisfied and that $E\left[x_i\right] = 0$, $E\left[x_i^2\right] = 2$, $E\left[x_i^3\right] = 0$, $E\left[x_i^4\right] = 7$, $E\left[x_i z_i\right] = 1$, $E\left[z_i^2\right] = 4$ and $E\left[\varepsilon_i^2 \middle| x_i, z_i\right] = x_i^2$. Also assume that all other moments that show up in your answers exist and are finite.

1. (5 points) Find the probability limit of the OLS estimator in a regression of y_i on z_i.

2. (12 points) Find the asymptotic distribution of the 2SLS estimator of δ that uses x_i as instrument.

Problem 7 (12 points)

Suppose that

$$y_i^* = b_1 + b_2 x_i^* \qquad \text{with} \qquad b_2 > 0$$

but that an econometrician observes a random sample of (y_i, x_i), where

$$y_i = y_i^* + u_i \qquad \text{and} \qquad x_i = x_i^* + v_i$$

with u_i, v_i and (y_i^*, x_i^*) independent of each other.

1. (6 points) What is the probability limit of the OLS estimator (of the slope) in a regression of y_i on x_i and a constant?

2. (6 points) What is the probability limit of the OLS estimator (of the slope) in a regression of x_i on y_i and a constant?

   Assume that all relevant moments exist.

REGRESSION OUTPUT FOR PROBLEM 5.

Regression 1.

```
. reg lwage ed black smsa south exp expsq

      Source |       SS       df       MS              Number of obs =    3010
-------------+------------------------------           F(  6,  3003) =  204.93
       Model |  172.165642     6  28.6942737           Prob > F      =  0.0000
    Residual |  420.476039  3003  .140018661           R-squared     =  0.2905
-------------+------------------------------           Adj R-squared =  0.2891
       Total |  592.641681  3009  .196956358           Root MSE      =  .37419

------------------------------------------------------------------------------
       lwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          ed |    .074009   .0035054    **.**   *.***     ********    ********
       black |  -.1896315   .0176266   -10.76   0.000    -.2241929   -.1550702
        smsa |    .161423   .0155733    10.37   0.000     .1308876    .1919583
       south |  -.1248615   .0151182    -8.26   0.000    -.1545046   -.0952184
         exp |   .0835958   .0066478    12.57   0.000     .0705612    .0966305
       expsq |  -.0022409   .0003178    -7.05   0.000    -.0028641   -.0016177
       _cons |    .128494   .0676026     1.90   0.057     -.004058    .2610461
------------------------------------------------------------------------------
```

Regression 2.

```
. generate exso=exp*so
. generate exso2=expsq*so

. reg lwage ed black smsa south exp expsq exso exso2

      Source |       SS       df       MS              Number of obs =    3010
-------------+------------------------------           F(  8,  3001) =  156.20
       Model |  174.226968     8   21.778371           Prob > F      =  0.0000
    Residual |  418.414713  3001  .139425096           R-squared     =  0.2940
-------------+------------------------------           Adj R-squared =  0.2921
       Total |  592.641681  3009  .196956358           Root MSE      =   .3734

------------------------------------------------------------------------------
       lwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          ed |   .0730965   .0035125    20.81   0.000     .0662093    .0799837
       black |  -.1856423   .0176208   -10.54   0.000    -.2201923   -.1510923
        smsa |   .1589221   .0155589    10.21   0.000     .1284149    .1894292
       south |    .041205   .0625563     0.66   0.510    -.0814525    .1638625
         exp |   .0894204   .0089523     9.99   0.000     .0718671    .1069737
       expsq |  -.0022838   .0004574    -4.99   0.000    -.0031806    -.001387
        exso |  -.0260771    .013428    -1.94   0.052    -.0524061     .000252
       exso2 |    .000671   .0006533     1.03   0.304      -.00061    .0019519
       _cons |   .0968781   .0718049     1.35   0.177    -.0439136    .2376698
------------------------------------------------------------------------------
```