ECON 581 Startz/ Tsang
Econometrics II Winter 2007

Homework 4 – Some Geometry and Some Hypothesis Testing

Due 2nd Feb 2007

Analytical Exercises

1. (From Davidson and MacKinnon Exercise 2.19, a rather tricky question) Show that $\mathbf{P_X} - \mathbf{P}_1 = \mathbf{P}_{\mathbf{M}_1\mathbf{X}_2}$, where $\mathbf{P}_{\mathbf{M}_1\mathbf{X}_2}$ is the projection on to the span of $\mathbf{M}_1\mathbf{X}_2$. This can be done most easily by showing that any vector $\mathbf{M}_1\mathbf{X}_2\mathbf{a} \in S\left(\mathbf{M}_1\mathbf{X}_2\right)$ is invariant under the action of $\mathbf{P_X} - \mathbf{P}_1$ (i.e. $\left(\mathbf{P_X} - \mathbf{P}_1\right)\mathbf{M}_1\mathbf{X}_2\mathbf{a} = \mathbf{M}_1\mathbf{X}_2\mathbf{a}$), and that any vector orthogonal to this span $\mathbf{b} \in S^{\perp}\left(\mathbf{M}_1\mathbf{X}_2\right)$ is annihilated by $\mathbf{P_X} - \mathbf{P}_1$.

2. Fun with chi-squared distribution (Trust me, it is fun):

a) Given a standard normal PDF $\phi\left(z\right) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}z^2\right)$, find for the derivatives $\phi'\left(z\right)$, $\phi''\left(z\right)$, $\phi'''\left(z\right)$ and $\phi''''\left(z\right)$ all as functions of $\phi\left(z\right)$ and z.

b) Use the results in a) to find the first, second, third and fourth moments for z, where the k-th moment is defined by $\int_{-\infty}^{\infty} z^k f\left(z\right)dz$.

c) Let $z_1, z_2, ..., z_m$ be mutually independent standard normal random variables, then $y \equiv \sum_{i=1}^{m} z_i^2$ follows the chi-squared distribution with degrees of freedom m, or $y \sim \chi^2\left(m\right)$. Show that $E\left(y\right) = m$ and $Var\left(y\right) = 2m$ (You need the results in b)).

Matlab Exercise

Consider the following DGP:

$y_i = 1 + 2x_i + u_i$, with $u_i \sim N(0,1)$ and $x_i \sim U(0,1)$ (standard uniform)

Now do an experiment with 10,000 runs, with a sample size of 1000. In each iteration, run the regression $y_i = \alpha + \beta x_i + u_i$, and another regression $y_i = \alpha + \beta x_i^* + u_i$, where $x_i^* = x_i + v_i$, with $v_i \sim N(0, 0.0025)$ being a rather tiny random noise. For each regression, perform the *t*-test for the null hypothesis that $\beta = 2$ at 5% level (i.e. the critical value is 1.96).

For each regression, out of what percentage of the 10,000 runs is the null hypothesis wrongly rejected? Is it close to 5%? Why and why not?