Session 8: Computer-based Exercise

Question 1

The following commands generate 1000 values of random variable A (denoted `A` in Stata syntax below), which has a $N(0,1)$ distribution.

```
set obs 1000
set seed 16743
gen A = rnormal()
```

[Note that by not including anything in the brackets in the `rnormal()` command we tell Stata to use the default values, which are a mean of 0 and standard deviation of 1.]

Using similar commands, generate 1000 values each of four random variables, A, B, C, D, where each variable has a $N(0,1)$ distribution.

[You only need to set the seed once: this is to start off your random sampling so as to agree with others using this seed.]

(a) Compare the distribution of the square of the first random variable, A^2, to a χ_1^2 distribution: for example, look at a histogram, and compare the sample mean, variance and 50th, 90th and 95th percentiles with those of the theoretical distribution.

[You may find the `invchi2tail()` function and `centile` commands useful for this: use Stata `help` to find out more about these commands.

You will find, for example, that `.display invchi2tail(1, 0.05)` displays the value 3.84...; and `.centile  A, c(95)` displays the 95th centile of variable `A`.]

(b) Similarly compare the distribution of $E = B^2 + C^2 + D^2$ to a χ_3^2 distribution.

(c) What theoretical distribution would you expect the random variable $A^2/(E/3)$ to have? Compare the 50th, 90th and 95th percentiles calculated from your sample with those from the theoretical distribution.

[You may find the `invFtail()` function useful here.]

What would be the problem with using $B^2/(E/3)$ instead of $A^2/(E/3)$?

(d) What theoretical distribution would you expect the random variable $+\sqrt{A^2/(E/3)}$ to have? Compare the 50th, 90th and 95th percentiles calculated from your sample with those from the theoretical distribution.

Question 2
Generate 1000 values of a random variable which has a $\chi^2_{10}/10$ distribution, using:

```
set seed 909693
gen xc = rchi2(10)
gen x10=xc/10
```

(a) Look at the distribution of `x10` , and compare the sample mean and variance to those from the theoretical distribution.

(b) Do the same for random variables which have $\chi^2_{50}/50$ and $\chi^2_{100}/100$ distributions.

(c) Based on these results, what happens to χ^2_n/n as n increases? Why does this help to explain why a t_n distribution tends to a $N(0,1)$ distribution as n increases?