Computer Exercise 5. t and F Distributions in Econometrics

Read in the data set *birth.dta* from the course website
The data set (adapted from Wooldridge) contains information on the, birth weight of children and the smoking habits and other characteristics of their parents

Run a regression of birthweight (bwgt) on father's education, mothers education, sex of child, birth order and the number of cigarettes smoked by the mother during pregnancy

a) for the 1st 20 observations in the data set
b) all the 1191 observations in the data set

Test the hypothesis that the number of cigarettes smoked by the mother has no effect on child birth weight in both samples

Graph the t distributions in both cases and show the critical values at the 95% level and the estimated t value on both graphs
(use the "rnd" command from the previous exercise you will have to work out what the degrees of freedom are in both cases. To ensure you can replicate your answers type "set seed 1" before you do any estimation)

```
reg bwght fathed moth order male numcigs if _n<21
```

```
      Source |       SS       df       MS              Number of obs =      20
-------------+------------------------------           F(  5,    14) =    1.24
       Model |  1.75707495     5   .35141499           Prob > F      =  0.3411
    Residual |  3.95847031    14  .282747879           R-squared     =  0.3074
-------------+------------------------------           Adj R-squared =  0.0601
       Total |  5.71554526    19  .300818172           Root MSE      =  .53174
------------------------------------------------------------------------------
       bwght |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    fatheduc |   .1138938   .0529348     2.15   0.049     .0003599    .2274278
    motheduc |  -.0285735   .0814309    -0.35   0.731    -.2032254    .1460783
       order |   .0174558   .1118564     0.16   0.878    -.2224524     .257364
        male |   .0490677   .2932761     0.17   0.870    -.5799469    .6780823
     numcigs |   .1213121   .0978241     1.24   0.235    -.0884997    .3311239
       _cons |   2.704827   1.049031     2.58   0.022     .4548801    4.954774
------------------------------------------------------------------------------
```

so cigarettes appear to have no effect on birth weight (t=1.24 not significantly different from zero, so lies in acceptance region of null hypothesis – that variable has no effect)

```
rndt 10000 14
```

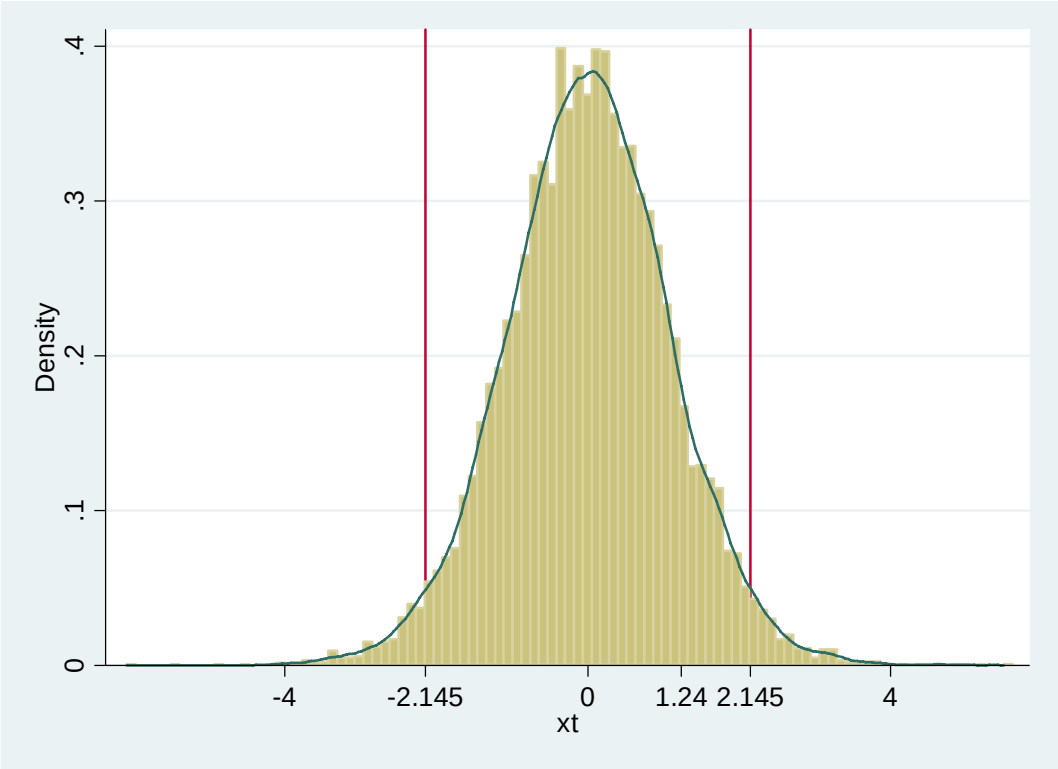
From t tables critical value at 95% level df = N – k = 20 – 6 = 14 is +- 2.145

(Alternatively use the stata command

```
di invttail(14, .025)
```

2.145

```
histogram xt, bin(100) kdensity xscale(range(. .))  xline(-2.064 2.064) xlabel(-
4 -2.145 0 1.24 2.145 4)
```

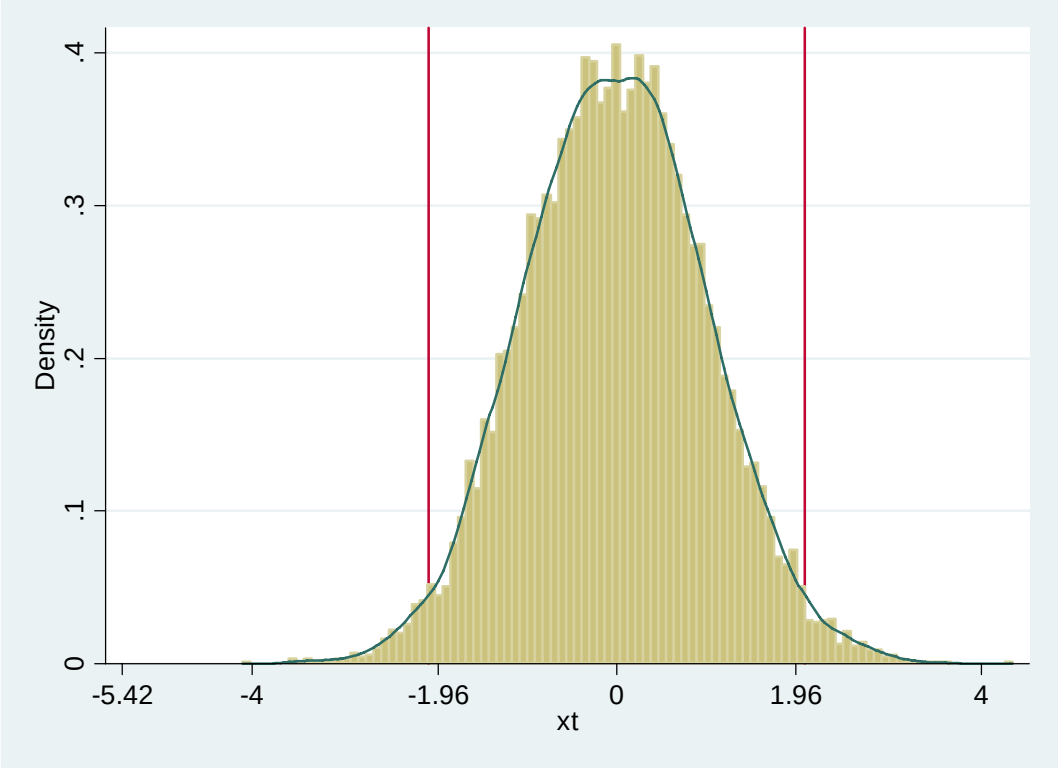



```
. reg bwght fathed moth order male numcigs

      Source |       SS       df       MS              Number of obs =    1191
-------------+------------------------------           F(  5,  1185) =   11.12
       Model |  22.0081423     5  4.40162846           Prob > F      =  0.0000
    Residual |  468.862655  1185  .395664688           R-squared     =  0.0448
-------------+------------------------------           Adj R-squared =  0.0408
       Total |  490.870798  1190  .412496469           Root MSE      =  .62902

------------------------------------------------------------------------------
       bwght |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    fatheduc |   .0182366   .0087055     2.09   0.036     .0011567    .0353166
    motheduc |  -.0084008   .0099792    -0.84   0.400    -.0279797    .0111781
       order |   .0576975   .0209599     2.75   0.006     .0165749      .09882
        male |   .1149613   .0365081     3.15   0.002     .0433336    .1865889
     numcigs |  -.0189771   .0035032    -5.42   0.000    -.0258502   -.0121039
       _cons |   3.562041   .1191713    29.89   0.000      3.32823    3.795851
------------------------------------------------------------------------------
```

Now in full sample cigarettes do appear to make birth weight significantly lower, other things equal. 10 cigarettes a day leads to 10*-.019 = -0.19 kg lower birth weight.

```
di invttail(1185, .025)
1.96
```
So estimated t lies outside acceptance region

```
histogram xt, bin(100) kdensity xscale(range(. .))  xline(-2.064 2.064) xlabel(-5.42 -4 -1.96 0  1.96 4)
```



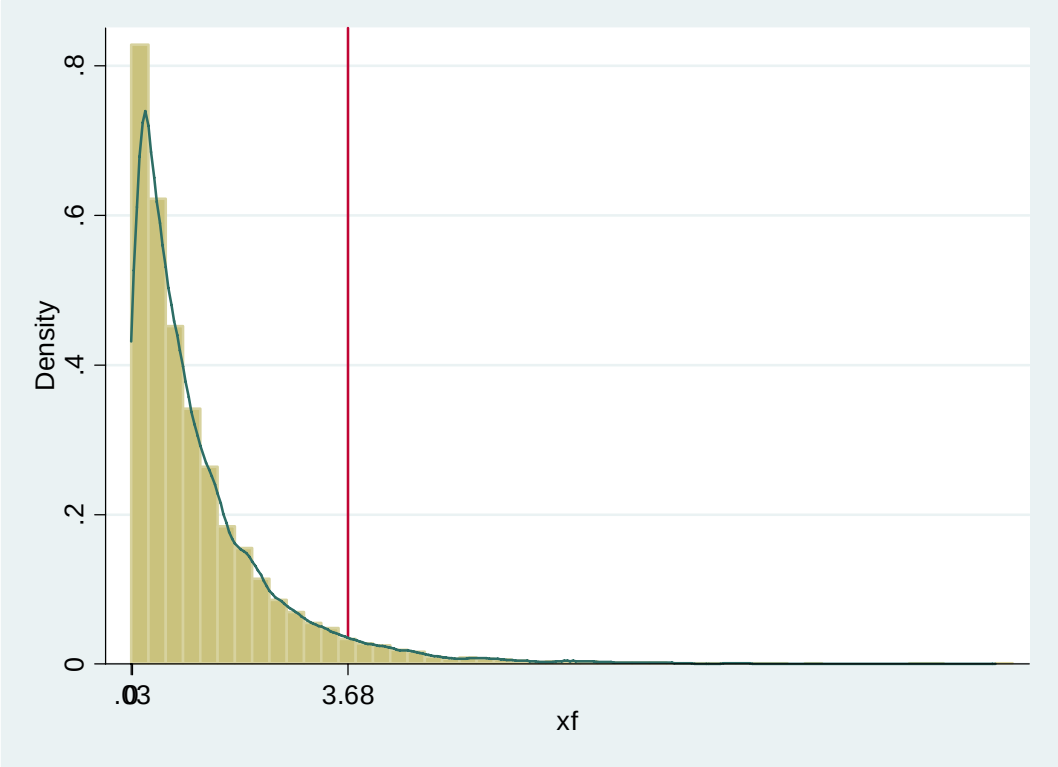
Note first t distribution has fatter tails (more leptokurtic) than second

```
test order male

 ( 1)  order = 0
 ( 2)  male = 0

       F(  2,    14) =    0.03
            Prob > F =    0.9724
```

So estimated F value lies inside acceptance region for null hypothesis
Cant reject null of no (joint) effect – Should be obvious from regression output

```
rndf 10000 2 14
di invF(2, 14, .95)
3.7388918
```
gives critical value at 95% level for these degrees of freedom

```
histogram xf, bin(50) kdensity xscale(range(. .))  xline(3.68) xlabel(0 0.03
3.68)
```



For full sample

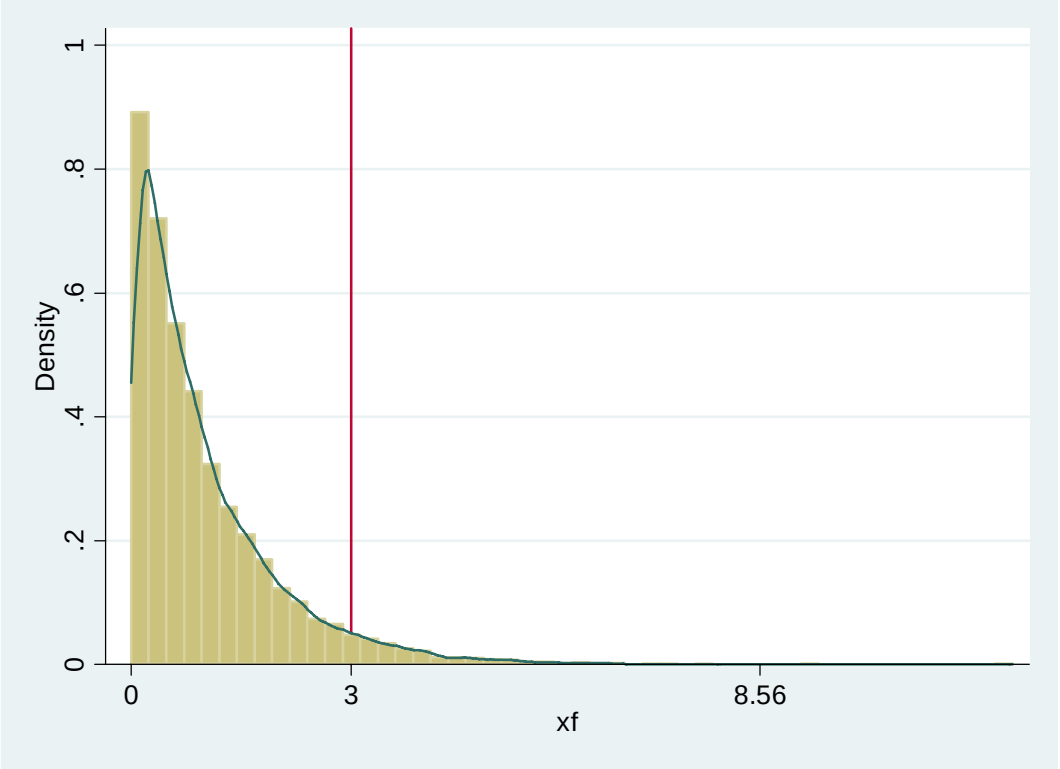
```
test order male

 ( 1)  order = 0
 ( 2)  male = 0

       F(  2,  1185) =    8.56
            Prob > F =    0.00020929
```

So now estimated F value is highly significant – lies outside acceptance region for null

```
rndf 10000 2 1185

di invF(2, 1185, .95)
3.0033184
```

(95% critical value at these degrees of freedom)

```
histogram xf, bin(50) kdensity xscale(range(. .))  xline(3.0) xlabel(0 8.56
3.0)
```


Density
0
.2
.4
.6
.8
1
0
3
8.56
xf