```
clear all
set more off, perm

* set working directory
global dir "C:\Users\Nmath_000\Documents\MI_school\Second Year\675 Applied Econometrics\hw\hw1"

*import data
import delimited using "$dir\LaLonde_1986.csv"

**************
* question 2 *
**************


* create needed variables
gen educ_sq = educ^2
gen black_earn74 = black*earn74
gen const = 1

* store needed variables in locals
*local y earn76
*local x const treat black age educ educ_sq earn74 black_earn74 u74 u75

* use mata
mata:


y = st_data(., "earn78")
x = st_data(., ("const", "treat", "black", "age", "educ", "educ_sq", "earn74","black_earn74", "u74","u75"))

n_row = rows(x)
n_col = cols(x)

b = invsym(cross(x,x))*cross(x,y)

bc = cholinv(cross(x,x))*cross(x,y)

diff = b-bc

diff

my_resid = y - x*b
d = diag(my_resid:*my_resid:*(n_row/(n_row-n_col)))

v = invsym(cross(x, x))*(x' * d * x) * invsym(cross(x, x))

se = sqrt(diagonal(v))

tstat = b :/ se

p_value = 2*ttail(n_row-n_col, abs(tstat))

CI_L = b - (se) * invt(n_row-n_col, .975 )
CI_U = b + (se) * invt(n_row-n_col, .975 )

all_data = b, se, tstat, p_value, CI_L, CI_U
all_data
end

// now run regression
reg earn78 treat black age educ educ_sq earn74 black_earn74 u74 u75, robust

// nice, they match

**************
* question 3 *
**************

```

```
**********
* neyman *
**********

sum earn78 if treat==0
local N0 = r(N)
local mu0 = r(mean)
local sd0 = r(sd)
local V0 = r(Var)/r(N)
local sig_sq0 = r(Var)

sum earn78 if treat==1
local N1 = r(N)
local mu1 = r(mean)
local sd1 = r(sd)
local V1 = r(Var)/r(N)
local sig_sq1 = r(Var)

local tau = `mu1'-`mu0'
local v = sqrt(`V1'+`V0')
local T = `tau'/`v'
local pval = 2*normal(-abs(`T'))

local mu0 = round(`mu0', .01)
local mu1 = round(`mu1', .0001)
local sd0 = round(`sd0', .01)
local sd1 = round(`sd1', .0001)

di "`tau'"


local CIlower = `tau' - invnormal(0.975)*`v'
local CIupper = `tau' + invnormal(0.975)*`v'

di "`CIlower'"
di "`CIupper'"

**********
* fisher *
**********

* Using difference in means estimator
permute treat diffmean=(r(mu_2)-r(mu_1)), reps(1999) nowarn: ttest earn78, by(treat)
matrix pval = r(p)
display "p-val = " pval[1,1]

* Using KS statistic
permute treat ks=r(D), reps(1999) nowarn: ksmirnov earn78, by(treat)
matrix pval = r(p)
display "p-val = " pval[1,1]

*************************
* 95% confidence interval*
*************************


* Infer missing values under the null of constant treatment effect
gen     Y1_imputed = earn78
replace Y1_imputed = earn78 + `tau' if treat==0

gen     Y0_imputed = earn78
replace Y0_imputed = earn78 - `tau' if treat==1

bootstrap treat diffmean=(r(mu_2)-r(mu_1)), reps(1999) nowarn: ttest earn78, by(treat)

*****************
*power funciton *
*****************

twoway function y= 1 - normal(invnormal(0.975)-x/`v') + normal(-invnormal(0.975)-x/`v'),
```

```
      range(-5000 5000)
139
140
141   mata: mata clear
142   mata:
143
144
145    function myfunc(N, s0, s1, p, tau){
146
147      return(1 - normal(invnormal(0.975)-tau/sqrt(1/N*s1*(1/p)+1/N*s0*(1/(1-p)))) +
148          normal(-invnormal(0.975)-tau/sqrt(1/N*s1*(1/p)+1/N*s0*(1/(1-p)))) -0.8)
149
150    }
151    s0 =  30072466.58373794
152    s1 =  61896056.06715253
153       p     = 2/3
154      tau   = 1000
155      p
156      tau
157    s0
158    s1
159
160
161     mm_root(x=., &myfunc(), 1000, 1500, 0, 10000, s0,s1, p ,tau)
162
163        x
164
165   end
166
167
168
169
```