P-Set 2 Adv Econometrics

Ethan Grant uni: erg2145

October 21, 2016

1. i) The difference between a finite sample and large sample test for a single regressors is that the finite sample uses the t stat without assuming it converges to the normal distribution while the large sample assumes it converges. For similar reasons the finite sample uses the F stat while the large sample assumes it converges to the Weld stat.

   Finite sample properties:
   E[b—X] = E[b] = β this means the least squares coefficient estimator is unbiased $E[s^2|X] = E[s^2] = \sigma^2$ The disturbance variance estimator is unbiased
   $Var[b|X] = \sigma^2(X'X)^{-1}$ and $Var[b] = \sigma^2 E[(X'X)^{-1}]$
   The MVLUE of $w'\beta$ is $w'b$ for any vector of constants w

   Large sample properties: Large sample estimator is consistent as opposed to unbiased which means that it gets closer to the real value as more data is added
   normality of the least squares estimators

   Asymptotic Properties:
   Asymptotic efficiency. The estimator is consistent, asymptotically normally distributed, and has an asymptotic covariance matrix that is not larger than the asymptotic covariance matrix of any other consistent, asymptotically normally distributed estimator.

   ii) see code

   iii) estimates for beta:
   beta_hat
   [,1]
   intercept -3.5265028
   log_Q 0.7203941
   log_PL 0.4363412
   log_PK 0.4265170
   log_PF -0.2198884

centered r squared = .926
uncentered r squared = .924

iv) v_homoskedastic:
intercept log_Q log_PL log_PK log_PF
intercept 456.5147020 -0.63344719 -20.43035532 1.03617608 -85.67561184
log_Q -0.6334472 0.04423607 -0.06594069 0.04564141 0.04678118
log_PL -20.4303553 -0.06594069 12.28276319 -1.58647862 3.43306676
log_PK 1.0361761 0.04564141 -1.58647862 1.46072497 -0.96127964
log_PF -85.6756118 0.04678118 3.43306676 -0.96127964 16.70570957

v_White:
intercept log_Q log_PL log_PK log_PF
intercept 413.502348 -1.92249509 -20.3585472 2.95030665 -76.9223976
log_Q -1.922495 0.14876383 -0.1937918 -0.03367381 0.2117281
log_PL -20.358547 -0.19379182 8.4471721 -1.20038898 3.9181152
log_PK 2.950307 -0.03367381 -1.2003890 0.79766956 -0.8560857
log_PF -76.922398 0.21172813 3.9181152 -0.85608566 14.6796013

v) The SE's (square root of values reported above) are all very close to each other which indicates that they are consistent as at larger values of N we would imagine they would get even closer.

vi) If $\beta_3 + \beta_4 + \beta_5 = 1$ it means that a 1 unit increase in the log of price of labor, fuels, and capital leads to a 100% increase in Y so it is exact returns to scale.

vii) R = [0,0,1,1,1]
q = [1]
$R\beta = q$:
$[0, 0, 1, 1, 1]\% * \%t([\beta_1, \beta_2, \beta_3, \beta_4, \beta_5]) = 1$

viii) the limiting distribution of the wald statistic is the chi squared distribution with a single degree of freedom. Wald = .0046

ix) fail to reject the null

x) regression becomes $Y = \beta_1 + \beta_2 * ln(Q) + \beta_3 * X_3 + \beta_4 * (X_4 - X_3) + \beta_5(X_5 - X_3)$
$Y = \beta_1 + \beta_2 * ln(Q) + (\beta_3 - \beta_4 - \beta_5)X_3 + \beta_4 * X_4 + \beta_5 * X_5$
Then test that the new beta is =0

xi) abs(t) =.0679 so fail to reject the null

xii) Greater than 1 means increasing returns to scale while less than 1 means decreasing returns to scale.I conclude that you should fail to reject both hypotheses.

xiii) This null hypothesis allows us to test if at least one of the inputs has no effect because if any beta =0 then the whole expersion is equal to zero. Thus, if one is zero we would fail to reject the null. The appropriate test is a wald test which is large sample.

You use the white standard error that was computed above
You also need to compute $\beta_4 * beta_5$, $\beta_3 * beta_5$, $\beta_4 * beta_3$, to make your G matrix

The appropriate test statistic is the chi-squared distribution with 1 degree of freedom

xiv) w_stat = .0036 so don't reject the null

2. i) $E[\epsilon_i|x_i] = E[e^{x_i} * u_i|x_i] = e^{x_i} * E[u_i|x_i] = 0$

ii) $Var[\epsilon_i|x_i] = Var[e^{x_i} * u_i|x_i] = e^{x_i} * Var[u_i|x_i] = e^{x_i} * \sigma_y$
The variance is dependent on the value of x_i which means it is not homoskedastic

iii) The OLS assumptions are not satisified because the variance is a function of x_i meaning that the errors are not spherical and assumption A4 is violated.

iv) homoskedatic SE rejection = 0/1000
heteroskedastic SE rejeciton = 0/1000
The tests seem to have an asymptotic size equal to 0% as in this large sample they reject 0% of the time

v) using $2x_i$:
homoskedatic SE rejection = 0/1000
heteroskedastic SE rejeciton = 0/1000

using $.1x_i$:
homoskedatic SE rejection = 0/1000
heteroskedastic SE rejeciton = 0/1000

using $0x_i$:
homoskedatic SE rejection = 0/1000
heteroskedastic SE rejeciton = 0/1000

I conclude they are all the same due to large sample size

vi) It does not seem important to use white SE if sample size is large

vii) Factor the sigma matrix into P'P and then rotate the data in the following manner

$y* = Py$
$X* = Px$

$e* = Pe$

So you need to first run the regression using the usual OLS method to get the residuals and the sigma matrix. Once that is done you factor the matrix and rotate the data as described above. Then you run the regression again using the usual method, but with the rotated data.