

## Stata Survival Commands for Moore/McCabe/Craig

- To save a work session's commands (without creating a command file or log file): right-click the Review window, then click either 'Save All' or 'Save Selected.'
- To use the Command window's calculator, e.g.:  
display 5 - 1 [press 'enter']  
di 6\*2 [press 'enter']  
di (2+5)/(1+1) [press 'enter']
- To inspect the spreadsheet: browse [type this in Command window, press 'enter']
- findit placevar [type in the Command window and press 'enter']: click the blue-link to download, e.g., the user-created command placevar.
- To make a data folder from within Stata: mkdir c:\hwdata [type this in Command window, press 'enter']
- To change to the data folder from within Stata: cd: c:\hwdata [type this in Command window, press 'enter']
- missing observations = . Stata treats '.' as the highest value, so, e.g.: su x if x>65 & x<. (so that Stata doesn't include '.' in the results).

### Ex. 1.5

```
help input [type in Command window, press 'enter']
input exam [type in Command window, press 'enter']
80 [press 'enter']
73 [press 'enter'] [continue entering all scores]
.
.
end [press 'enter']
```

*Note: If an error message displays after you type 'enter', re-do the procedure by clicking its steps in the Review window, making sure there's an observation number on the left-hand side in the Results window before you type 'enter'.*

```
help list
help stem
```

```
list exam
list exam in 1/4 [1/4: first to fourth observations]
list exam in 5/l [5/l: fifth to last observations]
list exam if exam==45
list exam if exam < 45
list exam if exam <=45
list exam if exam > 45 & exam <. [<.: exclude missing observations]
list exam if exam >= 45 & exam <.
list exam if exam ~=45 [or: !=45]
list exam if exam==. [list missing observations for exam]
```

sort exam [sort: lowest to highest values]  
 list exam  
 gsort -exam [sort: highest to lowest values]  
 list exam  
  
 browse [to inspect spreadsheet]  
 edit [to edit spreadsheet: but this way is not recommended]

summarize exam if exam==45  
 su exam if exam < 60 [abbreviated summarize: su]  
 su exam if exam <= 60  
 su exam if exam >60 & exam<.  
 su exam if exam >=69 & exam<.  
 su exam if exam==65  
 su exam if exam ~= 65 [or: !=65]

stem exam  
 stem exam if exam < 70  
 stem exam if exam >= 45 & exam <.

**Ex 1.19**

use ex01\_019, clear [type in Command window. Abbreviated: u]  
 help describe  
 d [describe: abbreviated]  
 l spam percent [list: abbreviated]

help graph  
 gr bar percent, over(spam, sort(1)) [graph: abbreviated]  
 gr save ex1\_19, replace  
 Right-click graph: Copy graph > paste to Word document

*Try hands-on graph formatting*

gr bar percent, over(spam)  
 Right-click the graph > Start Graph Editor > right-click in bar region > Bar Region  
 Properties: Sort bars > OK  
 Right-click graph: Save graph > ex1\_19, replace  
 Right-click graph: Copy graph > paste to Word document

*Helpful alternative for some kinds of bar graph:*

findit catplot [then click blue link to download catplot]

**Ex 1.20**

u ex01\_020, clear  
 d  
 l  
  
 gr bar female, over(degree)  
 Right-click the graph > Start Graph Editor > right-click in bar region > Bar Region  
 Properties: Sort bars > OK  
 Right-click graph: Save graph > ex1\_19, replace  
 Right-click graph: Copy graph > paste to Word document

**Ex 1.20**

u ex01\_021, clear

d

label variable over65 "66+ years old"

|

help sort

sort over65

[to reverse sort: gsort -over65]

|

stem over65

stem over65 if over65>5.7 & over65<17.6

stem over65 if over65~=5.7 & over65~=17.6

**Ex 1.32**

u ta01\_007, clear

d

|

help encode

encode pasadena, gen(pasa)

[destring pasadena]

d

help label

label variable pasa "Pasadena temps"

la var reading "Reading temps"

help histogram

histogram pasa, norm

hist reading, norm

[abbreviated: histogram]

**Ex 1.33**

help twoway

twoway line pasa reading year

twoway line pasa reading year, ylabel(0(25)100) xlabel(1950(25)2000)

**Ex 1.41**

d

|

help bys

bys sex: stem study

[bys = by, sort]

**Ex 1.47**

*After inputting the data for 'exam':*

d

|

help summarize

su exam

**Ex 1.48**

su exam, d

*Another way:*

findit univar  
help univar  
univar exam

[then click the blue-link to download univar]

**Ex 1.49**

su exam, d

[or: univar exam]

**Ex 1.50**

su exam, d

[or: univar exam]

**Ex 1.51**

su exam, d

[or: univar exam]

help graph  
gr box exam

[abbreviated: graph]

**Ex 1.52**

su exam, d  
gr box exam

[or: univar exam]

*Another way:*

findit iqr  
help iqr  
iqr exam

[then click the blue-link to download iqr]

**Ex 1.57**

u ex01\_057, clear  
d  
|  
su crp, d  
hist crp, norm  
gr box crp

*For Ex. 1.58*

gen lcrp= $\ln(\text{crp} + 1)$   
su crp lcrp, d

[use either 'log' or 'ln' (natural log)]

hist crp, norm  
hist lcrp, norm  
for varlist crp-lcrp: hist X, norm \ more

gr box crp  
gr box lcrp  
for varlist crp-lcrp: gr box X \ more

**Ex 1.97**

```
findit robmean
help robmean
d days
la var days "Number of days"
```

[then click the blue-link to download robmean]  
[‘robust mean’: trimmed mean]

```
su days, d
hist days, norm
gr box days
robmean days, trim(.1)
```

**Ex 1.101**

*Computing z-score:* display (individual score – mean score)/standard deviation  
di (600 – 572)/51

**Ex 1.149**

```
u ex01_080, clear
d
la var trees "Number of trees"
la var group "Group category"
su trees, d
tab group
```

```
hist trees, norm
gr box trees
qnorm trees
kdensity trees, norm
```

```
bys group: su trees
gr box trees, over(group, total)
qnorm trees if group==1
qnorm trees if group==2
qnorm trees if group==3
```

**Ex 2.6**

```
u ex02_006, clear
d
la var first "First exam"
la var final "Final exam"
```

```
su, d
hist first, norm
gr box first
hist final, norm
gr box final
```

[or: univar first final]

```
twoway scatter final first || lfit final first
twoway scatter final first || qfit final first
```

[lfit: linear]  
[qfit: quadratic]

**Ex 2.8**

```
u ex02_007, clear
d
```

```
help generate
gen id = _n
```

[gen: abbreviated 'generate']

```
lab var id "ID number"
lab var final "Final exam"
lab var second "Second exam"
```

```
help move
move id second
```

[move 'id' before 'second']

[findit placevar and download: to move multiple variables together: placevar x y z, before(a)]

[findit renvars and download: to change variables between upper and lower case]

```
d
l
```

```
help expand
expand 2 if id==8
l
```

```
help recode
recode second 153=199 if id==9
recode final 160=187 if id==9
l in 8/9
```

```
su, d
gr box second, marker(1, mlabel(id))
gr box final, marker(1, mlab(id))
```

```
twoway scatter final second, ml(id) || lfit final second
twoway scatter final second, ml(id) || qfit final second
```

**Ex 2.17**

```
input socdis brainact
1. 1.26 -0.055 [press 'enter']
2. [continue inputting the data]
14. end
```

```
gen subject=_n
la var subject "Subject"
move subject socdis
la var socdis "Social distress"
la var brain "Brain activity"
```

```
l
su, d
```

```
hist socdis, norm
gr box socdis, marker(1, mlab(id))
```

```
hist brain
gr box brain, marker(1, mlab(id))
```

```
twoway scatter brain socdis, ml(id) || lfit brain socdis
twoway scatter brain socdis, ml(id) || qfit brain socdis
```

### Ex 2.29

```
input price pdefor
```

```
la var price "Price in dollars: per pound"
la var pdefor "Percent deforestation"
```

```
su, d [or: univar price pdefor]
hist price, norm
gr box price
hist pdefor, norm
gr box pdefor
```

```
twoway scatter pdefor price || lfit pdefor price
twoway scatter pdefor price || qfit pdefor price
su
```

```
help egen [egen: 'extended generate']
```

```
egen mprice=mean(price)
egen mpdefor=mean(pdefor)
la var mprice "Mean price"
la var mpdefor "Mean percent deforestation"
su price mprice pdefor mpdefor
```

```
egen zprice=std(price)
egen zpdefor=std(pdefor)
la var zprice "Standardized price"
la var zpdefor "Standardized percent deforestation"
su price zprice pdefor zpdefor
[Then compute correlation coefficient step by step]
```

```
help correlate
corr pdefor price
```

### Ex 2.37

```
su price
```

```
gen pricee=price*0.628364 [dollar to euro conversion rate: April 23, 2008]
la var pricee "Price in euros: per pound"
```

```
note pricee: TS gen pricee=price*0.628364. Dollar to euro conversion rate, April 23, 2008.
```

```
su price pricee
twoway lfit price pricee [If the distributions of the two variables are equal, there will
be a 45 degree diagonal.]
```

```
for varlist price-pricee: hist X, norm \ more
for varlist price-pricee: gr box X \ more
```

```
twoway scatter pdefor pricee || lfit pdefor pricee
twoway scatter pdefor pricee || lfit pdefor price
twoway scatter pdefor price || qfit pdefor price
twoway scatter pdefor price || qfit pdefor price
```

```
corr pdefor price
corr pdefor price
```

### Ex 2.58

```
su, d
hist first, norm
gr box first
hist final, norm
gr box final
```

```
twoway scatter final first || lfit final first
twoway scatter final first || qfit final first
```

*Alternative way that doesn't display the observation points:*

```
twoway lfit final first
twoway qfit final first
```

```
corr final first [if no missing obs: pwcorr final first, obs sig star(.05)]
regress final first
```

*Note:* lfit is in fact the graph of the linear regression equation. qfit is the graph of the quadratic regression equation. This is so for regression models with just one explanatory variable. Thus, to graph the linear regression equation in Ex 2.58: twoway lfit final first.

### Ex 2.60

See Ex 2.8.

### Ex 2.67

```
u ex02_017, clear
d
la var brain "Brain activity"
la var distress "Distress level"
su, d
```

```
for varlist brain distress: hist X, norm \ more
for varlist brain distress: gr box X \ norm
```

```
twoway scatter brain distress || lfit brain distress
twoway scatter brain distress || qfit brain distress
```

```
corr brain distress [if no missing obs: pwcorr brain distress, obs sig star(.05)]
```

```
reg brain distress
help lincom
lincom _cons + distress*20
```

```
reg brain distress
predict yhat, xb
la var yhat "Predicted brain activity"
hist yhat, norm
su yhat, d
```

```
twoway scatter yhat brain || lfit yhat brain
```

### **Ex 2.87**

```
u ex02_087, clear
d
la var age "Age in months"
la var weight "Weight in kilograms"
la var resid "Residuals"
su, d
```

```
for varlist weight age: hist X, norm \ more
for varlist weight age: gr box \ norm
```

```
twoway scatter weight age || lfit weight age
twoway scatter weight age || qfit weight age
```

```
corr weight age [if no missing obs: pwcorr weight age, obs sig star(.05)]
```

```
reg weight age
predict yhat, xb
la var yhat "Predicted weight"
hist yhat, norm
su yhat, d
```

*How to compute residuals:*

```
gen resc=weight - yhat
la var resc "Residuals: computed"
```

*How to automatically obtain residuals:*

```
reg weight age
predict res, resid [also standardized residuals: predict rst, rstandard]
la var res "Residuals: predict command"
```

```
sort resid resc res
l resid resc res
su resid resc res
for varlist resid resc res: hist X, norm \ more
```

*Plot residuals vs. age:*

```
twoway scatter resid age || lfit resid age
```

*Easier way to plot residuals vs. age:*

```
reg weight age
rvpplot age, yline(0) [help regression postestimation]
```

**Ex 2.120**

d

l count

la var major "Major"

la var gender "Gender"

tab1 major gender [fw=count] [fw: frequency weight]

su [fw=count], d

help tabulate twoway [tab: abbreviated]

tab major gender [fw=count]

tab major gender [fw=count], col

tab major gender [fw=count], col nofreq

tab major gender [fw=count], row

tab major gender [fw=count], row nof

tab major gender [fw=count], cell

tab major gender [fw=count], cell nof

tab major gender [fw=count], cell row col

encode gender, gen(gend)

la var gend "Gender"

tab gend

tab gend, nolabel

recode gend 1=0 2=1

tab gend, nol

la def gen 0 "Female" 1 "Male"

la val gend gen

tab gend

tab gend, nol [abbreviated: nolabel]

tab gend major, row

tab gend major, col

tab gend major, cell

*Note:* tabi [‘immediate’ form of tabulate]

help tabulate

tabi 30 18 \ 18 14 [press ‘enter’]

tabi 30 18 \ 18 14, row

tabi 30 18 \ 18 14, col

tabi 30 18 \ 18 14, cell

tabi 30 18 38 \ 13 7 22

help table

table gend major [fw=count] [see options: tab ycat xcat, content(mean income)]

**Ex 3.xx**

help sample

seed 123 [‘seed’ to permit replication of the sample]

sample 10 [10% random sample]

seed 123

sample 10, by(gender)

```
seed 123
sample 10 if gender==1
```

```
seed 123
sample 100, count [random sample 100 observations]
```

```
help bsample [sample with replacement]
```

```
seed 123
bsample 10 [random sample 10 observations with replacement]
```

#### **Ex 4.xx**

*How to display a probability bar graph:*

```
input cardnum prob
```

```
1. 1 .3
2. 2 .1
3. 4 .2
4. 7 .2
5. 8 .1
6. 10 .1
7. end [press 'enter']
```

```
la var cardnum "Card's number"
la var prob "Probability"
```

```
d
su
hist prob, norm
gr bar prob, over(cardnum)
```

#### **Ex 5.xx**

```
help bitest [binominal test]
bitest bivar=.5, detail
bitest bivar, d [default: .95 significance]
bitest bivar, d level(90)
bitesti N #successes p, detail
bitesti 1000 47 .23, d
```

```
help prtest [proportion test]
prtest postbivar=prebivar
prtest postbivar=prebivar, l(99)
prtesti N p1 p2
prtesti 1000 .23 .29
```

#### **Ex 6.xx**

```
help ci [confidence interval]
help cii ['immediate' form]

cii #obs mean sd [default: .95 significance]
cii #obs mean sd, level(99)

ci x [default: .95 significance]
ci x, l(90)
```

*Power & Inference:*

findit sampsi

[then click blue-link to download sampsi]

help sampsi

*Note:* subcommands for multiple-hypothesis adjustments such as Bonferroni and Sidak are associated with specific *Stata* commands such as `pwcorr`. Otherwise, do a Bonferroni adjustment by using the Command window calculator to divide a selected P-value by the number of hypothesis tests conducted, e.g.: `di .05/10`

### **Ex 7.xx & Ex 8.xx**

cii #obs mean sd

cii #obs mean sd, l(99)

ci x

ci x, l(90)

*Binomial confidence interval:*

ci binomial-var, wilson

[or: `agresti, jeffries`]

ci binomial-var, l(99) wilson

help ttest [see 'unequal' option; see 'robvar' below to test equality of variances]

help ttesti

findit robvar [tests equality of variances: click blue-link to download robvar]

help prtest

help prtesti

*One sample test:*

ttesti obs mean sd val

['immediate' form]

ttesti 1000 63 17.4 51

ttesti 1000 63 17.4 51, l(99)

ttest x=60

ttest x=60, l(90)

help signtest

[nonparametric]

signtest x=60

prtesti obs

prtesti #obs1 #p1 #p2

[also: 'count' option]

prtesti 274 .43 .51

prtesti 274 .43 .51, l(99)

*Two sample tests:*

ttesti obs1 mean1 sd1 obs2 mean2 sd2

ttesti 1000 20 5 943 15 11

ttesti 1000 20 5 943 15 11, l(99)

prtesti N1 p1 N2 p2

prtesti 1000 29 777 31

prtesti 1000 29 777 31, l(90)

ttest income, by(gender) [see 'robvar']  
ttest income, by(gender) l(90)  
ttest income, by(gender) unequal l(99)

ranksum income, by(gender) [nonparametric]

prtest survive, by(gender)  
prtest survive, by(gender) l(90)

*Matched pairs tests:*

ttest postvar=prevar  
ttest postvar=prevar, l(99)

ttesti obs1 mean1 sd1 obs2 mean2 sd2  
ttesti 1000 20 5 943 15 11  
ttesti 1000 20 5 943 15 11, l(99)

signtest postvar=prevar [nonparametric]

prtest postbivar=prebivar  
prtest postbivar=prebivar, l(99)

prtesti obs p1 obs p2  
prtesti 1000 .23 987 .29  
prtesti 1000 .23 .987 29, l(90)

### **Ex 9.xx**

help tabulate twoway

tab fail gender  
tab fail gender, expected [chi2 contribution per cell: see tabchi, cont]  
tab fail gender, cchi2 ['nofreq' option]  
tab gender fail, cell chi2  
tab fail gender, col chi2  
tab gender fail, row chi2  
tab gender fail, all

bys ses: tab fail gender, col chi2

tabi 63 47 \ 51 92  
tabi 63 47 \ 57 92, col chi2

help tabchi  
tabchi fail gender, pearson  
tabchi fail gender, cont

tabchii 1 28 7 \ 9 53 68 \ 28 3 21

### **Ex 10.23**

u ta10\_004.dta", clear  
d  
l year

```
su count, d
hist count, norm
gr box count, marker(1, mlab(year))
```

```
twoway scatter count year || lfit count year
twoway scatter count year || qfit count year
```

```
corr count year
```

```
reg count year
predict resid, resid
hist resid, norm
rvfplot, yline(0)
```

#### **Ex 10.24**

```
reg count year if year~=2004
predict resid2, resid
for varlist resid-resid2: hist X, norm \ more
rvfplot, yline(0)
```

#### **Ex 11.27**

```
u ranking, clear
d
gen id=_n
la var id "ID#"
move id rank2006
```

```
l rank2006-name
l country-recruiter
l intlfaculty-facultystu
l citationfaculty_20- overallscore
```

```
sort rank2005
l name rank2005 country
```

```
sort rank2006
l name rank2006 country
```

```
sort rank2005 rank2006
l name rank2005 rank2006
```

```
su peereview_40, d [or: univar peereview_40]
hist box peereview_40, norm
gr box peereview_40, marker(1, mlab(name))
```

```
su facultystudent_20, d
hist facultystudent_20, norm
gr box facultystudent_20, marker(1, mlab(name))
```

```
su citationfac, d
hist citationfac, norm
gr box citationfac, marker(1, mlab(name))
```

```
su overallscore, d
hist overallscore, norm
gr box overallscore, marker(1, mlab(name))
```

```
help graph matrix
gr matrix overallscore peerreview_40 facultystudent_20 citationfaculty_20, half
```

```
twoway scatter peerreview_40 facultystudent_20, ml(id) || lfit peerreview_40
facultystudent_20
```

```
twoway scatter peerreview_40 facultystudent_20, ml(id) || qfit peerreview_40
facultystudent_20
```

```
twoway scatter peerreview_40 citationfaculty_20, ml(id) || lfit peerreview_
40 citationfaculty_20
```

```
twoway scatter peerreview_40 citationfaculty_20, ml(id) || qfit peerreview_
40 citationfaculty_20
```

```
twoway scatter citationfaculty_20 facultystudent_20, ml(id) || lfit peerreview_40
facultystudent_20
```

```
twoway scatter citationfaculty_20 facultystudent_20, ml(id) || qfit peerreview_40
facultystudent_20
```

### **Ex 11.28**

```
twoway scatter overallscore peerreview_40, ml(id) || lfit overallscore peerreview_40
twoway scatter overallscore peerreview_40, ml(id) || qfit overallscore peerreview_40
```

```
twoway scatter overallscore facultystudent_20, ml(id) || lfit overallscore
facultystudent_20
twoway scatter overallscore facultystudent_20, ml(id) || qfit overallscore
facultystudent_20
```

```
twoway scatter overallscore citationfaculty_20, ml(id) || lfit overallscore
citationfaculty_20
twoway scatter overallscore citationfaculty_20, ml(id) || qfit overallscore
citationfaculty_20
```

```
corr overallscore peerreview_40 facultystudent_20 citationfaculty_20
```

```
pwcorr overallscore peerreview_40 facultystudent_20 citationfaculty_20 if
overallscore~= . & peerreview_40~= . & facultystudent_20~= . &
citationfaculty_20~= ., obs sig star(.05) bonf
```

### **Ex 11.29**

```
reg overallscore peerreview_40
reg overallscore facultystudent_20
reg overallscore citationfaculty_20
```

```
help esttab
```

```

reg overallscore peerreview_40
eststo m1
reg overallscore peerreview_40 facultystudent_20
eststo m2
reg overallscore peerreview_40 facultystudent_20 citationfaculty_20
eststo m3
estab, se r2

```

*Note:* To clear eststo in order to start the sequence over again: eststo clear

See the course web document 'Editing Stata Esttab and Outreg2 Word Documents.'

### **Ex 11.31**

```

u happiness, clear
gen id=_n
la var id "ID#"
move id country
la country "Country"
la var lsi "Life Satisfaction Index"
la var gini "Gini Coefficient"
la var corrupt "Government Corruption Index"
la var life "Average Life Expectancy"
la var democracy "Civil and Political Liberties Index"

d

l country

help tabstat
tabstat lsi gini corrupt democracy life, s(n mean sd min p25 p50 p75 max) f(%9.2f)

hist lsi, norm
gr box lsi, marker(1, mlab(country))

hist gini, norm
gr box gini, marker(1, mlab(country))

hist corrupt, norm
gr box corrupt, marker(1, mlab(country))

hist life, norm
gr box life, marker(1, mlab(country))

hist democracy, norm
gr box democracy, marker(1, mlab(country))

help graph matrix
gr matrix lsi gini corrupt democracy life, half [list dependent variable first]

twoway scatter lsi gini, mlab(country) || lfit lsi gini
twoway scatter lsi gini, mlab(country) || qfit lsi gini

```

twoway scatter lsi corrupt, mlab(country) || lfit lsi corrupt  
twoway scatter lsi corrupt, mlab(country) || qfit lsi corrupt

twoway scatter lsi life, mlab(country) || lfit lsi life  
twoway scatter lsi life, mlab(country) || qfit lsi life

twoway scatter lsi democracy, mlab(country) || lfit lsi democracy  
twoway scatter lsi democracy, mlab(country) || qfit lsi democracy

twoway scatter gini corrupt, mlab(country) || lfit lsi corrupt  
twoway scatter gini corrupt, mlab(country) || qfit lsi corrupt  
twoway scatter gini life, mlab(country) || lfit gini life  
twoway scatter gini life, mlab(country) || qfit gini life

twoway scatter gini democracy, mlab(country) || lfit lsi democracy  
twoway scatter gini democracy, mlab(country) || qfit lsi democracy

twoway scatter corrupt life, mlab(country) || lfit corrupt life  
twoway scatter corrupt life, mlab(country) || qfit corrupt life

twoway scatter corrupt democracy, mlab(country) || lfit corrupt democracy  
twoway scatter corrupt democracy, mlab(country) || qfit corrupt democracy

twoway scatter life democracy, mlab(country) || lfit democracy life  
twoway scatter life democracy, mlab(country) || qfit life democracy life

help ladder [possible transformations: indicated by insignificant chi2 test]  
ladder lsi  
ladder gini  
ladder corrupt  
ladder life  
ladder democracy

*Is there sufficient linearity to justify estimating correlations and a regression model?  
The relationship of lsi and gini is curvilinear enough to warrant further exploration:*

hist lsi, norm  
hist gini, norm  
twoway qfit lsi gini [some curvilinearity]

ladder gini  
gen lgini=ln(gini)  
la var lgini "Log Gini Coefficient"  
note lgini: TS gen lgini=ln(gini)

for varlist gini-lgini: hist X, norm \ more [improved distribution]  
twoway scatter qfit lsi gini  
twoway scatter qfit lsi lgini

*There's no practical improvement in linearity, and the relationship between lsi and gini is sufficiently linear to justify both correlation and regression.*

*Use pwcorr instead of corr only if there are no missing observations:*  
pwcorr lsi gini life democracy corrupt, obs sig star(.05) bonf

*Here's a way to identify the linearity/curvilinearity of each y/x relationship in a model, holding the other variables constant.*  
findit mrunning [download] [graphed curvilinear y/x relationships, holding the other x's constant]  
mrunning lsi gini corrupt life democracy, combine(saving(graph1))

**Ex 11.32**

findit esttab [download estadd, estout, \_eststo, estso, esttab]  
help esttab  
help eststo

```
reg lsi gini
eststo m1
reg lsi gini life
eststo m2
reg lsi gini life democracy
eststo m3
reg lsi gini life democracy corrupt
eststo m4
```

**Ex 11.33**

esstab, se r2 [see the course web document 'Editing Stata's Esttab...']

```
reg lsi gini life democracy
help linktest [test for functional form's adequacy: hatsq should test insignificant]
help ovtest [test for functional form's adequacy: F should test insignificant]
linktest
estat ovtest
```

```
help vif [test for multicollinearity: tolerance<.1; or
estat vif individual vif>10]
```

```
rvfplot, yline(0) mlab(country) [distribution of residuals: should be random]
predict resid, resid
hist resid, norm [distribution of residuals: should be normal]
```

```
help hettest [test for heteroscedasticity: should test insig]
hettest, rhs mt(bonf)
rvpplot gini, yline(0) mlab(country) [distribution of residuals should be random]
rvpplot life, yline(0) mlab(country)
rvpplot democracy, yline(0) mlab(country)
avplots [each plot holds the other variables constant]
```

```
findit lvr2plot [download] [influential observations display in the
help lvr2plot upper right-hand quadrant]
lvr2plot, mlab(country)
```

*Note:* for alternative diagnostic graphs, 'findit modeldiag' and download.

*Note:* How to create an interaction variable

```
gen giniXlife=gini*life
su gini life giniXlife
```