

```

* Stata Analysis Examples Replication for ASDA 2nd Edition
* Berglund April 2017
* Chapter 7

* Section 7.5: Application of Linear Regression
* Bivariate relationships

use "P:\ASDA 2\Data sets\nhanes 2011_2012\nhanes1112_sub_8aug2016.dta", clear
rename *, lower

gen bpxdil_1 = bpxdil
replace bpxdil_1 = . if bpxdil == 0

svyset sdmvpsu [pweight = wtmecl2yr], strata(sdmvstra) vce(linearized) singleunit(missing)
svy, subpop(age18p): mean age
gen agec = age-46.36

* bivariate tests
svy, subpop(age18p): regress bpxdil_1 i.ridreth1
test 2.ridreth1 3.ridreth1 4.ridreth1 5.ridreth1

svy, subpop(age18p): regress bpxdil_1 i.marcat
test 2.marcat 3.marcat

svy, subpop(age18p): regress bpxdil_1 i.riagendr
test 2.riagendr

svy, subpop(age18p): regress bpxdil_1 agec
test agec

* Naive analysis Table 7.2
regress bpxdil_1 i.ridreth1 i.riagendr agec if age18p == 1

* Weighted regression analysis Table 7.3
regress bpxdil_1 i.ridreth1 i.riagendr agec if age18p == 1 [pweight = wtmecl2yr]

* Correct analysis with weights and complex sample features Table 7.4
svyset sdmvpsu [pweight = wtmecl2yr], strata(sdmvstra) vce(linearized) singleunit(missing)
svy, subpop(age18p): regress bpxdil_1 i.ridreth1 i.riagendr agec
estat effects, deff

* post estimation predictions
predict resids, resid
scatter resids agec

* add age squared
gen agecsq = agec * agec

svy, subpop(age18p): regress bpxdil_1 i.ridreth1 i.riagendr agec agecsq
estat effects

predict resids2, resid
scatter resids2 agec

* test interactions of interest
svy, subpop(age18p): regress bpxdil_1 i.ridreth1 i.riagendr agec agecsq ///
i.ridreth1#c.agec i.ridreth1#c.agecsq

test 2.ridreth1#c.agec 3.ridreth1#c.agec 4.ridreth1#c.agec 5.ridreth1#c.agec 2.ridreth1#c.agecsq
///
3.ridreth1#c.agecsq 4.ridreth1#c.agecsq 5.ridreth1#c.agecsq

* add gender interaction
svy, subpop(age18p): regress bpxdil_1 i.ridreth1 i.riagendr agec agecsq ///
i.ridreth1#c.agec i.ridreth1#c.agecsq i.riagendr#c.agec i.riagendr#c.agecsq

test 2.riagendr#c.agec 2.riagendr#c.agecsq

```

```

* marginal predicted values plot
svy, subpop(age18p): regress bpxdil_1 i.ridreth1 i.riagendr agec c.agec#c.agec ///
i.ridreth1#c.agec i.ridreth1#c.agec#c.agec i.riagendr#c.agec i.riagendr#c.agec#c.agec

* plots from above model
margins ridreth1, at(agec=(-30(5)30))
marginsplot

margins riagendr, at(agec=(-30(5)30))
marginsplot

* model diagnostics
svy, subpop(age18p): regress bpxdil_1 i.ridreth1 i.riagendr agec c.agec#c.agec ///
i.ridreth1#c.agec i.ridreth1#c.agec#c.agec i.riagendr#c.agec i.riagendr#c.agec#c.agec
estat effects

predict ehat1, resid
symplot ehat1, name(sym_ehat1_1, replace) title(Symplot of Residuals)
histogram ehat1, normal name(h_ehat1, replace) title(Histogram of Residuals)
qnorm ehat1, name(qnorm_ehat1, replace) title(Normal Q-Q Plot of Residuals)
predict yhat1, xb
scatter ehat1 yhat1, name(ehat1xyhat1, replace) title(Residuals vs. Predicted Y)
graph combine sym_ehat1_1 h_ehat1 qnorm_ehat1 ehat1xyhat1, rows(2)

* Use Pfeiffermann method (Q weighted)
regress wtmecl2yr i.ridreth1 i.riagendr agec
predict w_hat, xb
gen q_wtmecl2yr = wtmecl2yr / w_hat
svyset sdmvpsu [pweight = q_wtmecl2yr], strata(sdmvstra) vce(linearized) singleunit(missing)

svy, subpop(age18p): regress bpxdil_1 i.ridreth1 i.riagendr agec c.agec#c.agec ///
i.ridreth1#c.agec i.ridreth1#c.agec#c.agec i.riagendr#c.agec i.riagendr#c.agec#c.agec

```