Epidemiology/Biostats VHM 812/802 Course
Winter 2014, Atlantic Veterinary College, PEI
Javier Sanchez

Lecture 3a: Model building II

Index **Page**

- Datasets
  - daisy2red.dta
  - daisy2.dta (# obs. 9383)
  - coleman.dta

Model building strategies

- parsimony vs fit
- goals
  - prediction
  - estimates of effects
- **steps**
  - specify maximum model
  - address issues of missing values
  - functional form (linearity) of continuous predictors
  - criterion for selection
  - selection strategy
  - analysis
  - evaluate reliability
  - present results

Specifying the maximum model

- outcome of interest
- key predictors
- important confounders
- other variables of interest
  - lots / few
- causal model **DO NOT FORGET**
  - identify confounders
  - identify intervening variables
  - identify exposure-independent variables

Reducing the number of predictors

- 10 obs. per predictor
- screening – descriptive statistics
  - few missing values
  - substantial variability
  - small categories
- correlation
  - pairwise
- unconditional associations
  - liberal P-value
- multivariate analysis (eg princ. comp.)

Missing values

- complete case analysis
  - any missing value – entire observation ignored
- patterns of "missingness"
  - MCAR – missing completely at random
    - probability of being missing is truly random
    - no bias, just reduced power
  - MAR – missing at random
    - probability of being missing depends only on observed data (ie prob. can be fully explained by observed values)
  - MNAR – missing not at random (also NMAR)
    - probability of being missing depends on unobserved data (ie cannot be fully explained by observed values)
- imputation
  - predict what the missing values would be and insert these predicted values
  - eliminates "potential" bias from MAR and may reduce it from NMAR
- methods for analysis if incomplete data
  - not discussed

Functional form of continuous predictors (linearity)

- detecting non-linearity – in final model
  - plot residuals vs fitted values
    - simultaneous evaluation of all predictors
  - plot of residuals vs predictor
- detecting non-linearity – before / during model building
  - scatterplot of outcome vs predictor
    - smoothing functions
- detecting and correcting non-linearity
  - transformation of X
  - categorization of predictor
    - indicator dummy variable
    - hierarchical dummy variable
    - compare categorical and linear variables
  - explore polynomial functions of X
    - quadratic, cubic, etc
    - fractional polynomials
    - orthogonal polynomials

- Smoothed scatter-plots
  - best fitting line through a mass of data (not confined to any specific shape)
  - local-influence property
    - position of line only affected by "neighbours"
    - # of neighbours determined by bandwidth (width of the neighbourhood)
    - adjust bandwidth to control degree of smoothing
  - different types of smoother
    - lowess – commonly used
  - cautions
    - potential to mask important local effects
    - behave poorly at ends

Detecting and correcting non-linearity

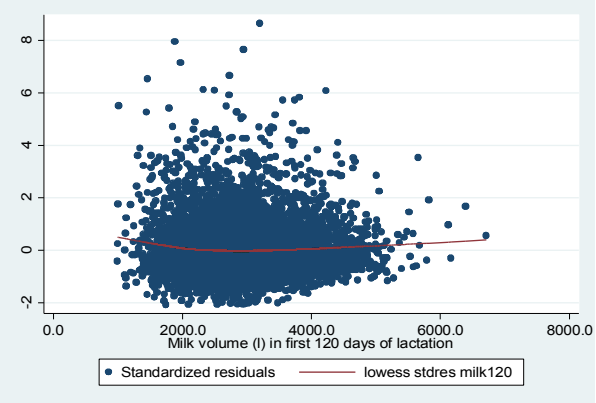
- transformation
  - of X, eg. log(x)

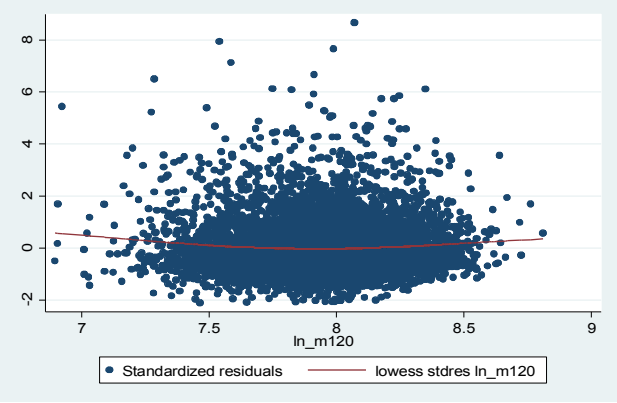
```
. reg cf milk120 -  R-squared     =  0.0001

------------------------------------------------------------------------------
          cf |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     milk120 |  -.0004435   .0004546    -0.98   0.329    -.0013345    .0004476
       _cons |   79.60298   1.379894    57.69   0.000     76.89801    82.30795
------------------------------------------------------------------------------

. reg cf ln_m120 -  R-squared     =  0.0006

------------------------------------------------------------------------------
          cf |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     ln_m120 |  -2.749959   1.305678    -2.11   0.035    -5.309441   -.1904759
       _cons |   100.1849   10.39878     9.63   0.000     79.80051    120.5694
------------------------------------------------------------------------------
```





  - transformation of Y (impact on normality, homoscedasticity)
- categorize X (more details L2a: Model build. I)
  - indicator variables

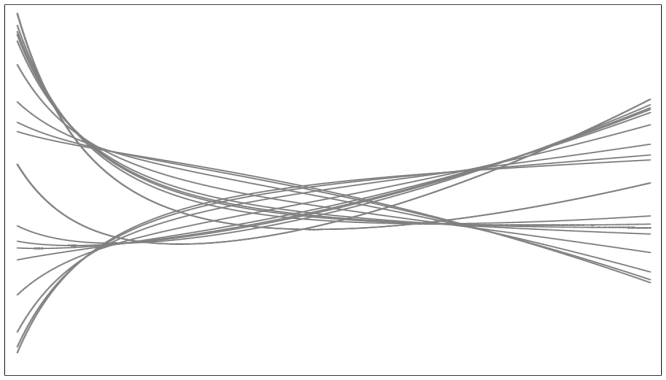
Polynomial functions of X

- quadratic (details L1b and L2b)

- fractional polynomials
  - extension polynomial regression
    - allow log, non-integer powers, repeated powers
  - select terms (usually one or two) of the form x^p
  - where "p" is form the set -2, -1, -0.5, 0, 0.5, 1, 2, 3
    - p=0 is taken to be ln(X)
    - eg $\beta_1 X^{-1}+\beta_2 X^2=\beta_1(1/X)+\beta_2(X^2)$
  - combination selected based on best fit (smallest log likelihood)

- usually 2 power terms (2 degree) can fit most shapes
  - 2-degree FP: x(-2,2) … $x^{(-2)}$ + $x^{(2)}$

- this graph shows some of the possibilities from a 2-degree FP



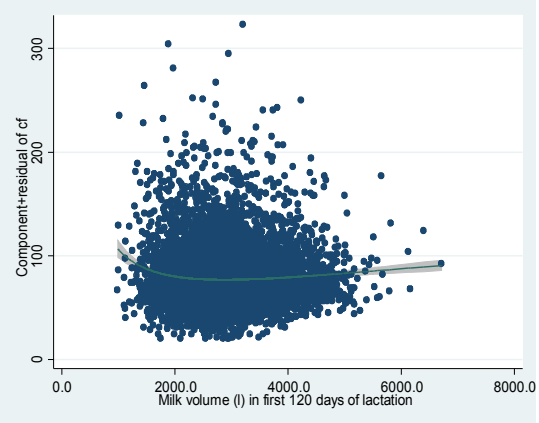
- Example – calving to first service and milk120

```
. fp <milk120>, scale center replace: reg cf <milk120>
(fitting 44 models)
(....10%....20%....30%....40%....50%....60%....70%....80%....90%....100%)

Fractional polynomial comparisons:
------------------------------------------------------------------------------
     milk120 |   df    Deviance  Res. s.d.   Dev. dif.   P(*)   Powers
-------------+----------------------------------------------------------------
     omitted |    0   73543.99     28.342     45.690    0.000
      linear |    1   73543.04     28.342     44.738    0.000   1
       m = 1 |    2   73525.42     28.310     27.118    0.000   -2
       m = 2 |    4   73498.30     28.262      0.000       --   -.5 0
------------------------------------------------------------------------------
(*) P = sig. level of model with m = 2 based on F with 7715 denominator dof.

      Source |       SS       df       MS              Number of obs =    7720
-------------+------------------------------           F(  2,  7717) =   22.90
       Model |  36587.3328     2  18293.6664           Prob > F      =  0.0000
    Residual |  6163761.19  7717  798.725048           R-squared     =  0.0059
-------------+------------------------------           Adj R-squared =  0.0056
       Total |  6200348.53  7719  803.258003           Root MSE      =  28.262

------------------------------------------------------------------------------
          cf |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   milk120_1 |   295.8967   46.01638     6.43   0.000     205.6922    386.1013
   milk120_2 |   87.36896   14.07522     6.21   0.000     59.77771    114.9602
       _cons |   76.95351    .378534   203.29   0.000     76.21148    77.69554
------------------------------------------------------------------------------
```



- orthogonal polynomials
  - similar to fractional polynomial
  - different parametrization
    - generates polynomials that have zero correlation with powers 1, 2, 3, etc.

Interactions

- 2 way
  - all possible
  - significant main effects
  - significant unconditional assoc.
  - biologically meaningful
  - with key predictor of interest
- 3 way
  - rarely

Selection criteria

- Non-statistical
  - predictor of interest
  - known confounder
  - evidence of being a confounder
  - component of an interaction term
- statistical - nested models
  - F-test for the predictor
  - Wald test or Likelihood ratio test (LRT)
  - always use these tests if appropriate

- other procedures – statistical – non-nested models
  - adjusted R^2 = $1-\frac{MSE}{MST}$
    - R^2 adjusted for the # predictors
    - linear regression models only
  - Mallow's Cp
    - linear regression only
    - $C_{pc}=\frac{RSS}{\sigma^2}+2k-n$
    - usually a positive value – might be negative if many predictors
    - lowest Cp = best
  - information criteria
    - AIC (Akaike's information criterion)
    - BIC (Bayesian information criterion)

Selection strategies

- all possible / best subset
  - look at all possible combinations of predictors
  - select best model based on some criterion (such as adjusted R2 or Mallow's CP)
  - best subset – computer finds "best" model with 1,2,3, etc predictors

- Example: coleman.dta – 20 schools in USA
  - outcome: test score 6th graders
  - predictors: staff salary, ses, educ mothers, etc (see 1bl)

```
. vselect y_test_scr x1_staff_sal x2_father_job x3_ses x4_test_teach x5_edu_mother,
best

Response :              y_test_scr
Fixed Predictors :
Selected Predictors:    x3_ses x4_test_teach x1_staff_sal
x5_edu_mother                                x2_father_job

Actual Regressions   5
Possible Regressions 32

 Optimal Models Highlighted:
   # Preds     R2ADJ         C        AIC       AICC        BIC
         1   .8518292  4.974883   90.89429   149.1518   92.88576
         2   .8740953  2.832759   88.49432   147.9185   91.48152
         3   .8820943  2.836089   87.96903   149.0123   91.95196
         4   .8756399  4.670221   89.74417   152.9633   94.72284
         5   .8728444         6   90.80893   156.8998   96.78332
Selected Predictors

1  :  x3_ses
2  :  x3_ses x4_test_teach
3  :  x3_ses x4_test_teach x1_staff_sal
4  :  x3_ses x4_test_teach x1_staff_sal x5_edu_mother
5  :  x3_ses x4_test_teach x1_staff_sal x5_edu_mother x2_father_job
```

- stepwise estimation
  - stepwise command in Stata
  - old syntax
- forward selection
  - start with a null model
  - adds terms based on statistical significance (one at a time, always choosing the most significant predictor not yet in the model)
  - stop when no more terms are significant when added

```
. stepwise, pe(0.1): reg y_test_scr x1_staff_sal x2_father_job x3_ses x4_test_teach
x5_edu_mother
                     begin with empty model
p = 0.0000 <  0.1000  adding  x3_ses
p = 0.0566 <  0.1000  adding  x4_test_teach

      Source |       SS       df       MS              Number of obs =      20
-------------+------------------------------           F(  2,    17) =   66.95
       Model |  570.497872     2  285.248936            Prob > F      =  0.0000
    Residual |  72.4264222    17  4.26037777            R-squared     =  0.8873
-------------+------------------------------           Adj R-squared =  0.8741
       Total |  642.924294    19  33.8381207            Root MSE      =  2.0641

------------------------------------------------------------------------------
   y_test_scr |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
--------------+---------------------------------------------------------------
       x3_ses |   .5415607   .0500441    10.82   0.000     .4359768    .6471445
x4_test_teach |   .7498915   .3666402     2.05   0.057    -.0236517    1.523435
        _cons |    14.5827   9.175409     1.59   0.130    -4.775723    33.94112
------------------------------------------------------------------------------
```

- backward elimination
  - starts with a full model
  - eliminates terms that are not significant (one at a time, starting with the "least significant")
  - stop once all terms remaining in the model are significant

```
. stepwise, pr(0.1): reg y_test_scr x1_staff_sal x2_father_job x3_ses x4_test_teach
x5_edu_mother
                      begin with full model
p = 0.4267 >= 0.1000  removing x2_father_job
p = 0.6863 >= 0.1000  removing x5_edu_mother
p = 0.1616 >= 0.1000  removing x1_staff_sal

      Source |       SS       df       MS              Number of obs =      20
-------------+------------------------------           F(  2,    17) =   66.95
       Model |  570.497872     2  285.248936           Prob > F      =  0.0000
    Residual |  72.4264222    17  4.26037777           R-squared     =  0.8873
-------------+------------------------------           Adj R-squared =  0.8741
       Total |  642.924294    19  33.8381207           Root MSE      =  2.0641

------------------------------------------------------------------------------
   y_test_scr |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
--------------+---------------------------------------------------------------
       x3_ses |   .5415607   .0500441    10.82   0.000     .4359768    .6471445
x4_test_teach |   .7498915   .3666402     2.05   0.057    -.0236517    1.523435
        _cons |    14.5827   9.175409     1.59   0.130    -4.775723    33.94112
------------------------------------------------------------------------------
```

- stepwise
  - combines forward and backward
  - generally preferred approach is stepwise backward
    - starts with a full model and works backward using a stepwise approach

```
. stepwise, pe(0.1) pr(0.11): reg y_test_scr x1_staff_sal x2_father_job x3_ses
x4_test_teach x5_edu_mother
                     begin with full model
p = 0.4267 >= 0.1100  removing x2_father_job
p = 0.6863 >= 0.1100  removing x5_edu_mother
p = 0.1616 >= 0.1100  removing x1_staff_sal

      Source |       SS       df       MS              Number of obs =      20
-------------+------------------------------           F(  2,    17) =   66.95
       Model |  570.497872     2  285.248936           Prob > F      =  0.0000
    Residual |  72.4264222    17  4.26037777           R-squared     =  0.8873
-------------+------------------------------           Adj R-squared =  0.8741
       Total |  642.924294    19  33.8381207           Root MSE      =  2.0641

------------------------------------------------------------------------------
   y_test_scr |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
--------------+---------------------------------------------------------------
       x3_ses |   .5415607   .0500441    10.82   0.000     .4359768    .6471445
x4_test_teach |   .7498915   .3666402     2.05   0.057    -.0236517    1.523435
        _cons |    14.5827   9.175409     1.59   0.130    -4.775723    33.94112
------------------------------------------------------------------------------
```

- daisy2red dataset
  - reg wpc_sqrt parity1 aut_clv herd_size hs_sq dyst twin twdy rp vag_disch rpvd

```
*stepwise backward
stepwise, pe(0.05) pr(0.051) lockterm1: reg sqrt_wpc  parity1 aut_clv (hs_ct hs_sq)
      (dyst twin twdy) (rp vag_disch rpvd)
estimates store sw_1

stepwise, pe(0.05) pr(0.051) lockterm1: reg sqrt_wpc  parity1 aut_clv (hs_ct hs_sq)
      dyst twin twdy rp vag_disch rpvd
estimates store sw_2
estimates table sw_1 sw_2
```

```
. estimates table sw_1 sw_2
----------------------------------------
    Variable |    sw_1         sw_2
-------------+--------------------------
     parity1 |  .05586593    .04388077
     aut_clv | -.51283075   -.52440137
       hs_ct | -.02346682   -.02161068
       hs_sq |   .0000713    .00006694
        dyst |  .62771805
        twin |  1.6982381    1.4787911
        twdy | -2.7727951
          rp |  .39042354
   vag_disch | -.04220258
        rpvd |  1.4760578    1.8317287
       _cons |  3.0230207    3.4048174
----------------------------------------
```

Cautions with automated selection procedures

- don't let your computer decide what variables go into your model
- R^2 too high
- F-tests too large
- severe problems if collinearity
- ignores non-statistical considerations
  - exposures, confounders and intervening var.
- you need to take care of
  - dummy variables
  - interaction terms
  - missing data
- useful when faced with large number of predictor variables
  - help to identify predictors that potentially are statistically significant associated with the outcome
- perform residual and influential analysis of selected models

Conducting the analysis

Evaluate reliability

- validity
  - regression diagnostics
- reliability
  - ability to predict future observations
  - split sample analysis
    - split data into two parts (eg. 50:50)
    - build model using one part (1st model), generate predictions and compare them with observed values for the other part (2nd model)
    - cross-validation correlation
      - corr. predicted and obs. values 2nd mod.
    - shrinkage on cross-validation
      - difference R^2 from the 1st model and the square of the cross-validation correlation
      - guideline: <0.1 is Ok
  - leave-one-out analysis
    - fit the model using all data except one observation
    - generate prediction for the left-out obs.
    - compare the sum of residuals from predicted points to prediction error from full model

- analogous to Cook's D and DFFITS

- split-sample analysis – daisy2

```
. gen rand=uniform()

. reg  wpc_sqrt hs_ct hs_sq parity1 aut_calv twin i.dyst##vag_disch if rand<0.6

      Source |       SS       df       MS              Number of obs =    3637
-------------+------------------------------           F(  8,  3628) =    9.94
       Model |  688.168438     8  86.0210547           Prob > F      =  0.0000
    Residual |  31393.8604  3628    8.653214           R-squared     =  0.0215
-------------+------------------------------           Adj R-squared =  0.0193
       Total |  32082.0288  3636  8.82344027           Root MSE      =  2.9416

------------------------------------------------------------------------------
      wpc_sqrt |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
---------------+--------------------------------------------------------------
         hs_ct |   -.030159   .0047019    -6.41   0.000    -.0393776   -.0209404
         hs_sq |   .0000767   .0000112     6.87   0.000     .0000548    .0000986
       parity1 |   .0106027   .0334649     0.32   0.751    -.0550091    .0762145
      aut_calv |  -.3514468   .0988828    -3.55   0.000    -.5453182   -.1575754
          twin |   .9893748   .3763624     2.63   0.009     .2514719    1.727278
        1.dyst |   .1741367    .270418     0.64   0.520    -.3560497     .704323
    1.vag_disch |   .8600507   .3648321     2.36   0.018     .1447543    1.575347
                |
dyst#vag_disch |
          1 1  |  -1.787119   .9615773    -1.86   0.063    -3.672405     .098167
                |
         _cons |   3.121718    .731371     4.27   0.000     1.687778    4.555657
------------------------------------------------------------------------------

. predict pv
(option xb assumed; fitted values)
(1604 missing values generated)

. corr pv wpc_sqrt if rand>=0.6
(obs=2482)

             |       pv wpc_sqrt
-------------+------------------
          pv |   1.0000
    wpc_sqrt |   0.1193   1.0000


. *shrinkage on cross-validation
R2 first model= .02344636  and CV2= .01423657

. di "shrinkage on cross-validation = " r2_1-cv_sq
shrinkage on cross-validation = .00920979
```

Presenting the results

- standardized coefficients
  - scales all coefficients so that they represent a change of 1 SD.

$$\beta^{*}=\beta(\sigma_{x}/\sigma_{y})$$

    - use with caution to compare coeff. from diff. studies

- interquartile ranges (IQR)
  - how big is the change in the outcome if the predictor changes through out the IQR
    - IQR = diff. between 75^{th} and 25^{th} percentile
- predictors eliminated from the model
  - don't ignore predictors just because they have been eliminated from the model
    - not statistically sig. ≠ no effect
    - unconditional associations?
    - force back in to the final model?
- scale of results
  - dealing with transformed data
  - compute some expected effects of key predictors on the original scale at varous levels of the other factors

A Structured Approach to Data Analysis (VER 30)

- "jump to the finish"
  - start over
- data collections sheets
- data coding
- data entry
- keeping track of files
- keeping track of variables
- analyses
  - data editing
  - data verification
  - data processing - outcome variable
  - data processing - predictor variables
  - data processing - multilevel
  - unconditional associations
- keeping track of analyses

Stata code

```
* VHM 812 - Winter 2014
* Model Building II - Lecture 3a1

* change working directory and open a log file
cd "C:\Users\JavierTablet\Dropbox\vhm812\2014\Data"
      * capture log close
      * log using ch15_lect_E.log, text replace
set more off

* open the DAISY Red dataset
use daisy2.dta, clear
gen month=month(calv_dt)
gen aut_calv=(month>=2 & month<=7)
gen hs_ct=herd_size-251
gen hs_sq=herd_size^2
gen parity1=parity-1
gen wpc_sqrt=sqrt(wpc)


* specifying maximum model
* no analyses

* causal model
* no analyses

* reducing the number of predictors
* descriptive statistics
codebook cf vag_disch
sum  milk120 parity herd_size dyst rp vag_disch, detail

* correlation
corr  milk120 parity herd_size
pwcorr milk120 parity herd_size, obs star(0.05)

* indices
* no analyses

* unconditional associations
reg cf parity
reg cf vag_disch

* principle components / factor analysis / correspondence analysis
* not covered

* EVALUATING/FIXING LINEARITY OF EFFECT
* residual plots
reg cf milk120
predict stdres, rstandar

* lowess smoother
twoway (scatter stdres milk120) (lowess stdres milk120)

* transformation of X
reg cf milk120
capture drop stdres
predict stdres, rstandar
twoway (scatter stdres milk120) (lowess stdres milk120)

gen ln_m120=log(milk120)
```

```
reg cf ln_m120
capture drop stdres
predict stdres, rstandar
twoway (scatter stdres ln_m120) (lowess stdres ln_m120)


* categorization of predictor
reg cf parity
egen parity_c6=cut(parity), at(0 1 2 3 4 5 6 15) icodes
tab parity parity_c6
reg cf i.parity_c6


* quadratic function of X
reg cf c.milk120##c.milk120
test c.milk120#c.milk120
estat vif
vce, corr

* redoing the analysis with milk120 centred
summ milk120
gen m120_ct=(milk120-r(mean)) /*r(mean) - memory variable created by summ command
that store the mean of the variable*/
reg cf c.m120_ct##c.m120_ct
test c.m120_ct c.m120_ct#c.m120_ct
estat vif
vce, corr


* fractional polynomials
fp <milk120>,  scale center replace: reg cf <milk120>
fp plot, r(residuals)
fp plot, r(none)

*old command
fracpoly reg cf milk120

*reproduce fp by hand
capture drop milk120_fp*
capture drop milk120k
gen milk120k=milk120/1000 if e(sample)
summ milk120k
gen milk120_fp1=(milk120k^-0.5) if e(sample)
replace milk120_fp1=milk120_fp1-(r(mean)^(-0.5)) if e(sample)
gen milk120_fp2=ln(milk120k) if e(sample)
replace milk120_fp2=milk120_fp2-ln(r(mean)) if e(sample)

reg cf milk120_fp1 milk120_fp2


* DEALING WITH NON-LINEARITY - advanced methods
* interactions
* 2-way
reg  wpc_sqrt hs_ct hs_sq aut_calv twin dyst##vag_disch

* 3-way retpla*vag_dysc*dyst
* only consider if absolutely necessary
reg  wpc_sqrt hs_ct hs_sq aut_calv twin dyst##vag_disch##rp


* selection criteria
```

```
* non-statistical
* no analyses

* statistical - nested models
* Wald test
reg  wpc_sqrt hs_ct hs_sq aut_calv twin i.dyst##vag_disch


* AIC and BIC
* statistical non-nested - this in logistic regression


*Selection strategy - Coleman dataset
* best subset - install command vselect
use coleman.dta, clear
rename x1 x1_staff_sal
rename x2 x2_father_job
rename x3 x3_ses
rename x4 x4_test_teach
rename x5 x5_edu_mother
rename y y_test_scr

vselect y_test_scr x1_staff_sal x2_father_job x3_ses ///
      x4_test_teach x5_edu_mother, best

* forward selection
stepwise, pe(0.1): reg y_test_scr x1_staff_sal x2_father_job x3_ses x4_test_teach
x5_edu_mother
* backward elimination
stepwise, pr(0.1): reg y_test_scr x1_staff_sal x2_father_job x3_ses x4_test_teach
x5_edu_mother
* stepwise selection
stepwise, pe(0.1) pr(0.11): reg y_test_scr x1_staff_sal x2_father_job x3_ses
x4_test_teach x5_edu_mother

*Selection strategy - daisy2red
use daisy2red, clear
gen hs_ct=herd_size-251
gen hs_sq=herd_size^2
gen parity1=parity-1
gen month=month(calv_dt)
gen aut_clv=(month>=2 & month<=7) if  !missing(calv_dt)
gen sqrt_wpc=sqrt(wpc)
gen twdy=twin*dyst
gen rpvd=rp*vag_disch

reg sqrt_wpc  parity1 aut_clv  hs_ct hs_sq twin dyst twdy  rp vag_disch rpvd

*stepwise backward
stepwise, pe(0.05) pr(0.051) lockterm1: reg sqrt_wpc  parity1 aut_clv (hs_ct hs_sq)
///
      (dyst twin twdy) (rp vag_disch rpvd)
estimates store sw_1

stepwise, pe(0.05) pr(0.051) lockterm1: reg sqrt_wpc  parity1 aut_clv (hs_ct hs_sq)
///
      dyst twin twdy rp vag_disch rpvd
estimates store sw_2
estimates table sw_1 sw_2
```

```
* reliability
* split sample analysis
use daisy2, clear
gen wpc_sqrt=sqrt(wpc)
gen hs_ct=herd_size-251
gen hs_sq=herd_size^2
gen parity1=parity-1
gen month=month(calv_dt)
gen aut_calv=(month>=2 & month<=7)

gen rand=uniform()
reg  wpc_sqrt hs_ct hs_sq parity1 aut_calv twin i.dyst##vag_disch if rand<0.6
scalar r2_1 = e(r2)
predict pv
*cross validation correlation
corr pv wpc_sqrt if rand>=0.6
*shrinkage on cross-validation
scalar cv_sq = r(rho)^2   /* this computes r2 - r(rho) is a "saved result" from
-corr-  */
di "R2 first model= " r2_1 "  and CV2= " cv_sq
di "shrinkage on cross-validation = " r2_1-cv_sq


* presenting results
* IQR
reg  wpc_sqrt hs_ct hs_sq parity1 aut_calv twin dyst##vag_disch
codebook parity1  /* note IQR = 3 - 0 = 3  */
                  /* coefficient for parity = 0.024 */
margins, over(dyst vag_disch) at(parity1=(0 3) hs_ct=0 hs_sq=0 aut_calv=0 twin=1)
expression(predict(xb)^2)
```