```
#### Biost 518: Applied Biostatistics II
#### Emerson, Winter 2006

#### Homework #5 Key
#### Annotated Stata Log File
#### March 10, 2006

#### The following output was used to generate the numbers that I wanted to present
#### in tables, as well as the plots I wanted to present as figures in the paper.
#### I note that Stata does not present its output in a form suitable for presentation.
#### Numbers need to be rounded to an interesting number of significant digits, and
#### the columns and rows need to facilitate comparison of relevant measures.
#### I used Excel to manipulate this output into the form I wanted, then copied the
#### resulting tables to the MS-Word document.

#### Comments edited into the log file produced by Stata are
#### on the lines that start with the four '#' signs and are
#### printed in italics.

#### The Stata commands are put in bold face.

#### Stata output is displayed in regular typeface in blue.

#### Read in data. I use a dataset previously created in "wide" format.
. use dfmowide

#### Creating variables to model threshold (tx), dummy variables (dose075, dose200,
#### dose400), and polynomials (dosesqr, dosecub)
. g tx= dose
. recode tx 0.01/0.5=1
(82 changes made)

. table tx dose
----------------------------------
          |          dose
       tx |    0   .075    .2    .4
----------+-----------------------
        0 |   32
        1 |          29    25    28
----------------------------------

. g dose075= dose
. recode dose075 .01/.08=1 .08/0.5=0
(82 changes made)

. g dose200= dose
. recode dose200 0/.08=0 .08/.25=1 .25/.5=0
(82 changes made)

. g dose400= dose
. recode dose400 0/0.3=0 0.3/0.5=1
(82 changes made)

. table dose075 dose
----------------------------------
          |          dose
  dose075 |    0   .075    .2    .4
----------+-----------------------
        0 |   32          25    28
```

```
        1 |         29
----------------------------------

. table dose200 dose
----------------------------------
          |          dose
  dose200 |    0  .075    .2    .4
----------+-----------------------
        0 |   32    29          28
        1 |               25
----------------------------------

. table dose400 dose
----------------------------------
          |          dose

  dose400 |    0  .075    .2    .4
----------+-----------------------
        0 |   32    29    25
        1 |                     28
----------------------------------

. g dosesqr= dose^2
. g dosecub= dose^3

#### Generate log transforms of spermidine data. Note the need to check for zero values
#### when missing values are produced for logspd12.
. g logspd0= log(spd0)
. g logspd12= log(spd12)
(20 missing values generated)

. summ spd12
    Variable |     Obs        Mean   Std. Dev.       Min        Max
-------------+--------------------------------------------------------
       spd12 |      95    2.768562   1.233778          0      6.454

. summ spd12 if spd12>0
    Variable |     Obs        Mean   Std. Dev.       Min        Max
-------------+--------------------------------------------------------
       spd12 |      94    2.798014    1.20635       .293      6.454

. replace logspd12= log(.15) if spd12==0
(1 real change made)
. replace spd12= .15 if spd12==0
(1 real change made)

#### Problem 1
#### Descriptive statistics
. tabstat spd12, stat(n mean sd min q max) co(stat) by(dose) format
Summary for variables: spd12
     by categories of: dose

 dose |  N      mean        sd       min       p25       p50       p75       max
------+--------------------------------------------------------------------------
    0 | 28     3.256     1.314     1.013     2.262     2.816     4.273     5.910
 .075 | 26     2.920     0.994     1.352     2.127     2.859     3.635     4.923
 .200 | 21     2.712     1.395     0.293     1.757     2.509     3.777     6.454
```

```
 .400 | 20     1.950     0.799     0.000     1.475     1.929     2.456     3.417
------+-----------------------------------------------------------------------
Total | 95     2.769     1.234     0.000     1.987     2.553     3.586     6.454
-------------------------------------------------------------------------------

. replace spd12= .15 if spd12==0
(1 real change made)
. bysort dose: means spd12
-------------------------------------------------------------------------------
-> dose = 0
    Variable |    Type        Obs        Mean       [95% Conf. Interval]
-------------+-----------------------------------------------------------------
       spd12 | Arithmetic      28    3.255935        2.746536   3.765334
             |  Geometric      28    3.008319        2.564826   3.528499
             |   Harmonic      28    2.765647        2.351809   3.356228
-------------------------------------------------------------------------------

-> dose = .075

    Variable |    Type        Obs        Mean       [95% Conf. Interval]
-------------+-----------------------------------------------------------------
       spd12 | Arithmetic      26    2.919683        2.518294   3.321073
             |  Geometric      26    2.749752        2.375747   3.182634
             |   Harmonic      26    2.576789         2.23116    3.04913
-------------------------------------------------------------------------------

-> dose = .2

    Variable |    Type        Obs        Mean       [95% Conf. Interval]
-------------+-----------------------------------------------------------------
       spd12 | Arithmetic      21     2.71162        2.076408   3.346832
             |  Geometric      21      2.3247          1.7371   3.111065
             |   Harmonic      21    1.766532        1.144841   3.865809
-------------------------------------------------------------------------------

-> dose = .4

    Variable |    Type        Obs        Mean       [95% Conf. Interval]
-------------+-----------------------------------------------------------------
       spd12 | Arithmetic      20    1.949568        1.575666   2.323471
             |  Geometric      20    1.713356        1.257762   2.333977
             |   Harmonic      20    1.178535        .6687964   4.955442
-------------------------------------------------------------------------------

#### Six regressions with fitted values estimated after each one.
#### Note that in every case I can get the test of the spermidine- dose association
#### from the overall F test.
. regress logspd12 dose075 dose200 dose400, robust
Regression with robust standard errors                Number of obs =      95
                                                      F(  3,    91) =    4.20
                                                      Prob > F      =  0.0079
                                                      R-squared     =  0.1454
                                                      Root MSE      =  .51704
------------------------------------------------------------------------------
             |               Robust
    logspd12 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
```

```
------------------------------------------------------------------------------
     dose075 |  -.0898709   .1055491    -0.85   0.397    -.2995313    .1197894
     dose200 |  -.2577906   .1596293    -1.61   0.110    -.5748745    .0592934
     dose400 |  -.5629276   .1664787    -3.38   0.001     -.893617   -.2322382
       _cons |   1.101382   .0779899    14.12   0.000     .9464642    1.256299
------------------------------------------------------------------------------

. predict fitB
(option xb assumed; fitted values)

. replace fitB= exp(fitB)
(114 real changes made)

. regress logspd12 dose, robust
Regression with robust standard errors                Number of obs =      95
                                                      F(  1,    93) =   12.67
                                                      Prob > F      =  0.0006
                                                      R-squared     =  0.1451
                                                      Root MSE      =  .51155
------------------------------------------------------------------------------
             |               Robust
    logspd12 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        dose |   -1.40961   .3959712    -3.56   0.001    -2.195931   -.6232899
       _cons |   1.111247   .0616695    18.02   0.000     .9887833    1.23371
------------------------------------------------------------------------------

. predict fitC
(option xb assumed; fitted values)

. replace fitC= exp(fitC)
(114 real changes made)

. regress logspd12 dose dosesqr, robust
Regression with robust standard errors                Number of obs =      95
                                                      F(  2,    92) =    6.36
                                                      Prob > F      =  0.0026
                                                      R-squared     =  0.1454
                                                      Root MSE      =  .51422
------------------------------------------------------------------------------
             |               Robust
    logspd12 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        dose |  -1.167386   1.403428    -0.83   0.408    -3.954715    1.619942
     dosesqr |  -.6026757   3.518415    -0.17   0.864    -7.590551      6.3852
       _cons |   1.101725   .0734313    15.00   0.000     .9558845    1.247566
------------------------------------------------------------------------------

. predict fitD
(option xb assumed; fitted values)

. replace fitD= exp(fitD)
(114 real changes made)

. regress logspd12 tx, robust
Regression with robust standard errors                Number of obs =      95
                                                      F(  1,    93) =    7.30
                                                      Prob > F      =  0.0082
```

```
                                                R-squared     =  0.0558
                                                Root MSE      =  .53759
------------------------------------------------------------------------------
             |               Robust
    logspd12 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          tx |  -.2837134   .1050154    -2.70   0.008    -.4922532   -.0751736
       _cons |   1.101382   .0771467    14.28   0.000     .9481834    1.25458
------------------------------------------------------------------------------

. predict fitE
(option xb assumed; fitted values)

. replace fitE= exp(fitE)
(114 real changes made)

. regress logspd12 tx dose, robust
Regression with robust standard errors                Number of obs =      95
                                                      F(  2,    92) =    6.34
                                                      Prob > F      =  0.0026
                                                      R-squared     =  0.1453
                                                      Root MSE      =  .51425
------------------------------------------------------------------------------
             |               Robust
    logspd12 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          tx |   .0242732   .1282516     0.19   0.850    -.2304454    .2789919
        dose |  -1.458311   .5070687    -2.88   0.005    -2.465393   -.4512292
       _cons |   1.101382   .0775649    14.20   0.000      .947331    1.255432
------------------------------------------------------------------------------

. predict fitF
(option xb assumed; fitted values)

. replace fitF= exp(fitF)
(114 real changes made)

. regress logspd12 dose dosesqr dosecub, robust
Regression with robust standard errors                Number of obs =      95
                                                      F(  3,    91) =    4.20
                                                      Prob > F      =  0.0079
                                                      R-squared     =  0.1454
                                                      Root MSE      =  .51704
------------------------------------------------------------------------------
             |               Robust
    logspd12 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        dose |  -1.137711   2.613929    -0.44   0.664    -6.329959    4.054537
     dosesqr |  -.8383989   21.22799    -0.04   0.969    -43.00519     41.3284
     dosecub |   .4109465   39.01003     0.01   0.992    -77.07768    77.89957
       _cons |   1.101382   .0779899    14.12   0.000     .9464642    1.256299
------------------------------------------------------------------------------

#### Testing for departures from a linear relationship, just for fun.
. test dosesqr dosecub
 ( 1)  dosesqr = 0.0
 ( 2)  dosecub = 0.0
```

```
       F(  2,    91) =    0.02
            Prob > F =    0.9821

. predict fitG
(option xb assumed; fitted values)

. replace fitG= exp(fitG)
(114 real changes made)

#### Table of fitted geometric means.
. tabstat fitB fitC fitD fitE fitF fitG, stat(mean) by(dose)
Summary statistics: mean
  by categories of: dose
    dose |      fitA      fitB      fitC      fitD      fitE      fitF
---------+--------------------------------------------------------------
       0 |  3.008319  3.038144  3.009354  3.008319  3.008319  3.008319
.0750000 |  2.749752  2.733345  2.747747  2.265212  2.762902  2.749752
.2000000 |    2.3247   2.29177  2.325979  2.265212  2.302493    2.3247
.4000000 |  1.713356  1.728757  1.713166  2.265212   1.72001  1.713356
---------------------------------------------------------------------------
   Total |  2.474566  2.475325   2.47458  2.473803  2.474675  2.474566
---------------------------------------------------------------------------

#### Problem 2
#### Six regressions adjusted for baseline.
#### Note that in several cases I have to use "test" or "testparm" to test
#### the dose effect, because in several models dose is modeled using several
#### predictors.
. regress logspd12 dose075 dose200 dose400 logspd0, robust
Regression with robust standard errors                Number of obs =      95
                                                      F(  4,    90) =    4.73
                                                      Prob > F      =  0.0017
                                                      R-squared     =  0.1675
                                                      Root MSE      =  .51314
------------------------------------------------------------------------------
             |               Robust
    logspd12 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     dose075 |  -.0929842   .0988132    -0.94   0.349    -.2892939    .1033256
     dose200 |  -.2612635   .1568764    -1.67   0.099    -.5729258    .0503988
     dose400 |  -.5753144   .1688008    -3.41   0.001    -.9106666   -.2399622
     logspd0 |   .1820632   .0897642     2.03   0.045      .003731    .3603954
       _cons |   .9011687   .1234444     7.30   0.000     .6559248    1.146413
------------------------------------------------------------------------------

. testparm dose*
 ( 1)  dose075 = 0.0
 ( 2)  dose200 = 0.0
 ( 3)  dose400 = 0.0
       F(  3,    90) =    4.28
            Prob > F =    0.0071

. regress logspd12 dose logspd0, robust
Regression with robust standard errors                Number of obs =      95
                                                      F(  2,    92) =    9.75
                                                      Prob > F      =  0.0001
                                                      R-squared     =  0.1671
```

```
                                                      Root MSE      =  .50765
------------------------------------------------------------------------------
             |               Robust
    logspd12 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        dose |  -1.438404   .4021031    -3.58   0.001    -2.237016    -.639793
     logspd0 |   .1818201   .0887094     2.05   0.043     .0056355    .3580046
       _cons |   .9113681   .1208245     7.54   0.000     .6714002    1.151336
------------------------------------------------------------------------------

. display 2*ttail(92,3.58)
.00055105

. regress logspd12 dose dosesqr logspd0, robust
Regression with robust standard errors                Number of obs =      95
                                                      F(  3,    91) =    6.37
                                                      Prob > F      =  0.0006
                                                      R-squared     =  0.1675
                                                      Root MSE      =  .51031
------------------------------------------------------------------------------
             |               Robust
    logspd12 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        dose |  -1.177407   1.381165    -0.85   0.396    -3.920921    1.566108
     dosesqr |  -.6494769   3.502937    -0.19   0.853    -7.607632    6.308678
     logspd0 |   .1820503   .0892683     2.04   0.044     .0047297    .3593709
       _cons |   .9008543   .1207236     7.46   0.000     .6610518    1.140657
------------------------------------------------------------------------------

. testparm dose*
 ( 1)  dose = 0.0
 ( 2)  dosesqr = 0.0
       F(  2,    91) =    6.49
            Prob > F =    0.0023

. regress logspd12 tx logspd0, robust
Regression with robust standard errors                Number of obs =      95
                                                      F(  2,    92) =    6.75
                                                      Prob > F      =  0.0018
                                                      R-squared     =  0.0744
                                                      Root MSE      =  .53515
------------------------------------------------------------------------------
             |               Robust
    logspd12 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          tx |  -.2892058   .1015678    -2.85   0.005    -.4909282   -.0874834
     logspd0 |   .1668214   .0923632     1.81   0.074    -.0166199    .3502627
       _cons |   .9179299   .1255277     7.31   0.000     .6686211    1.167239
------------------------------------------------------------------------------

. regress logspd12 tx dose logspd0, robust
Regression with robust standard errors                Number of obs =      95
                                                      F(  3,    91) =    6.40
                                                      Prob > F      =  0.0006
                                                      R-squared     =  0.1674
                                                      Root MSE      =  .51035
------------------------------------------------------------------------------
```

```
             |               Robust
    logspd12 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          tx |    .024438   .1235557     0.20   0.844    -.2209902    .2698661
        dose |  -1.487438    .515701    -2.88   0.005    -2.511815   -.4630615
     logspd0 |   .1818333   .0892754     2.04   0.045     .0044987    .359168
       _cons |   .9014214   .1227733     7.34   0.000     .6575474    1.145295
------------------------------------------------------------------------------

. test tx dose
 ( 1)  tx = 0.0
 ( 2)  dose = 0.0
       F(  2,    91) =    6.47
            Prob > F =    0.0024


. regress logspd12 dose dosesqr dosecub logspd0, robust
Regression with robust standard errors                Number of obs =      95
                                                      F(  4,    90) =    4.73
                                                      Prob > F      =  0.0017
                                                      R-squared     =  0.1675
                                                      Root MSE      =  .51314
------------------------------------------------------------------------------
             |               Robust
    logspd12 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        dose |  -1.205761   2.460956    -0.49   0.625    -6.094879    3.683356
     dosesqr |  -.4242516   20.34999    -0.02   0.983    -40.85305    40.00455
     dosecub |  -.3926507   37.74516    -0.01   0.992    -75.37999    74.59469
     logspd0 |   .1820632   .0897642     2.03   0.045      .003731    .3603954
       _cons |   .9011687   .1234444     7.30   0.000     .6559248    1.146413
------------------------------------------------------------------------------

. testparm dose*
 ( 1)  dose = 0.0
 ( 2)  dosesqr = 0.0
 ( 3)  dosecub = 0.0
       F(  3,    90) =    4.28
            Prob > F =    0.0071

#### Problem 3
#### Six regressions on a dichotomized variable.
#### Note that my commands are all the same, except I substitute "logistic" for
#### "regress" and "decspd" for "logspd12".
#### Note that in every case I can get the test of the spermidine- dose association
#### from the overall F test.

. g decspd = spd12 - spd0
(19 missing values generated)
. recode decspd min/0=1 0/max=0
(decspd: 95 changes made)

. drop fitB fitC fitD fitE fitF fitG

. tabulate dose decspd, row

           |        decspd
      dose |         0          1 |     Total
```

```
-----------+----------------------+----------
         0 |         15         13 |        28
           |      53.57      46.43 |    100.00
-----------+----------------------+----------
      .075 |         10         16 |        26
           |      38.46      61.54 |    100.00
-----------+----------------------+----------
        .2 |          8         13 |        21
           |      38.10      61.90 |    100.00
-----------+----------------------+----------
        .4 |          4         16 |        20
           |      20.00      80.00 |    100.00
-----------+----------------------+----------
     Total |         37         58 |        95
           |      38.95      61.05 |    100.00
```

```
. logistic decspd dose075 dose200 dose400, robust
Logistic regression                               Number of obs   =         95
                                                  Wald chi2(3)    =       5.18
                                                  Prob > chi2     =     0.1594
Log pseudolikelihood = -60.623017                 Pseudo R2       =     0.0454

------------------------------------------------------------------------------
             |               Robust
      decspd | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     dose075 |   1.846154    1.02681     1.10   0.270      .620639    5.491572
     dose200 |      1.875   1.107978     1.06   0.287     .5888542    5.970281
     dose400 |   4.615385   3.133498     2.25   0.024      1.21985    17.46262
------------------------------------------------------------------------------
. predict fitB
(option p assumed; Pr(decspd))

. logistic decspd dose, robust
Logistic regression                               Number of obs   =         95
                                                  Wald chi2(1)    =       4.71
                                                  Prob > chi2     =     0.0299
Log pseudolikelihood = -60.925995                 Pseudo R2       =     0.0407

------------------------------------------------------------------------------
             |               Robust
      decspd | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        dose |   30.85402   48.73495     2.17   0.030     1.395759     682.045
------------------------------------------------------------------------------
. predict fitC
(option p assumed; Pr(decspd))

. logistic decspd dose dosesqr, robust
Logistic regression                               Number of obs   =         95
                                                  Wald chi2(2)    =       4.75
                                                  Prob > chi2     =     0.0931
Log pseudolikelihood = -60.925026                 Pseudo R2       =     0.0407

------------------------------------------------------------------------------
             |               Robust
      decspd | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
```

```
-------------+----------------------------------------------------------------
        dose |   38.64213   210.7046     0.67   0.503     .0008825     1692087
     dosesqr |   .5588149    7.48834    -0.04   0.965     2.19e-12    1.42e+11
------------------------------------------------------------------------------
. predict fitD
(option p assumed; Pr(decspd))

. logistic decspd tx, robust
Logistic regression                               Number of obs   =         95
                                                  Wald chi2(1)    =       3.45
                                                  Prob > chi2     =     0.0631
Log pseudolikelihood = -61.748291                 Pseudo R2       =     0.0277

------------------------------------------------------------------------------
             |               Robust
     decspd | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          tx |    2.36014   1.090563     1.86   0.063     .9541532    5.837909
------------------------------------------------------------------------------

. predict fitE
(option p assumed; Pr(decspd))

. logistic decspd tx dose, robust
Logistic regression                               Number of obs   =         95
                                                  Wald chi2(2)    =       5.14
                                                  Prob > chi2     =     0.0765
Log pseudolikelihood = -60.804892                 Pseudo R2       =     0.0426

------------------------------------------------------------------------------
             |               Robust
     decspd | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          tx |   1.349357   .8225588     0.49   0.623      .408545    4.456702
        dose |   15.85019   32.04232     1.37   0.172     .3014867    833.2988
------------------------------------------------------------------------------

. predict fitF
(option p assumed; Pr(decspd))

. logistic decspd dose dosesqr dosecub, robust
Logistic regression                               Number of obs   =         95
                                                  Wald chi2(3)    =       5.18
                                                  Prob > chi2     =     0.1594
Log pseudolikelihood = -60.623017                 Pseudo R2       =     0.0454

------------------------------------------------------------------------------
             |               Robust
     decspd | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        dose |   545039.1    7386349     0.97   0.330     1.59e-06    1.87e+17
     dosesqr |   2.99e-34   3.00e-32    -0.77   0.442     1.0e-119    8.81e+51
     dosecub |   2.17e+58   3.78e+60     0.77   0.441     1.10e-90    4.3e+206
------------------------------------------------------------------------------

#### Testing for departures from a linear relationship, just for fun.
```

```
. test dosesqr dosecub
( 1)  dosesqr = 0
 ( 2)  dosecub = 0

           chi2(  2) =    0.60
         Prob > chi2 =    0.7424

. predict fitG
(option p assumed; Pr(decspd))

#### Table of fitted geometric means.
. tabstat fitB fitC fitD fitE fitF fitG, stat(mean) by(dose)
Summary statistics: mean
  by categories of: dose

    dose |      fitB      fitC      fitD      fitE      fitF      fitG
---------+------------------------------------------------------------
       0 |  .4642857  .4931918  .4910965  .4642857  .4642857  .4642857
.0750000 |  .6153846  .5572381  .5585266  .6716418  .5899517  .6153846
.2000000 |  .6190476    .65895  .6619449  .6716418  .6702163  .6190476
.4000000 |        .8  .7932245   .791338  .6716418  .7793357        .8
---------+------------------------------------------------------------
   Total |  .6191183   .619527  .6194601  .6134366  .6187943  .6191183
------------------------------------------------------------------------
```