```
#### Biost 517: Applied Biostatistics I
#### Emerson, Fall 2007

#### Annotated Stata Log File: Homework #3
#### October 29, 2007

#### In this file I give the Stata commands I used to produce
#### the key to Homework #2. In order to properly format
#### a table useful to casual readers, I cut and pasted some
#### of the output into Excel.

#### 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: The infile command was typed all on one line
. infile ptid time female age dose put spd spm using dfmolong.txt
'ptid' cannot be read as a number for ptid[1]
'time' cannot be read as a number for time[1]
'female' cannot be read as a number for female[1]
'age' cannot be read as a number for age[1]
'dose' cannot be read as a number for dose[1]
'put' cannot be read as a number for put[1]
'spd' cannot be read as a number for spd[1]
'spm' cannot be read as a number for spm[1]
'NA' cannot be read as a number for age[384]
'NA' cannot be read as a number for age[385]
'NA' cannot be read as a number for age[386]
'NA' cannot be read as a number for age[387]
(408 observations read)

#### Drop case corresponding to variable labels
. drop in 1
(1 observation deleted)

#### Examine range of data to see how to format output
. summ

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
```

```
        ptid |       407     8649.02    494.5656       8001       9089
        time |       407      7.7543     5.81391          0         15
      female |       407    .1547912    .3621507          0          1
         age |       403    63.61194    8.158043   45.42916   80.97467
        dose |       407    .1537469    .1501429          0         .4
-------------+---------------------------------------------------------
         put |       407    .8532547    .9618222          0      9.135
         spd |       407    2.981733    1.343019          0      7.844
         spm |       407    7.408888    4.102511          0   41.68433

. format age %9.1f

. format female put spd spm %9.2f

. format dose %9.3f

. format time %9.0f

#### Problem 1
. tabstat put spd spm, stat(n mean sd min q max) col(stat) format by(dose)

Summary for variables: put spd spm
     by categories of: dose

    dose |         N      mean        sd       min       p25       p50       p75       max
---------+--------------------------------------------------------------------------------
       0 |    117.00      1.09      1.29      0.00      0.48      0.72      1.13      9.14
         |    117.00      3.16      1.35      1.01      2.17      2.84      3.95      7.05
         |    117.00      7.17      3.86      1.46      4.88      6.47      8.69     35.55
---------+--------------------------------------------------------------------------------
.0750000 |    109.00      0.84      0.78      0.00      0.38      0.60      0.96      4.28
         |    109.00      3.00      1.17      0.00      2.24      2.82      3.63      7.02
         |    109.00      7.76      3.96      0.00      5.59      7.43      9.19     37.67
---------+--------------------------------------------------------------------------------
.2000000 |     90.00      0.71      0.65      0.00      0.27      0.51      0.93      3.21
         |     90.00      2.92      1.36      0.29      1.92      2.62      3.82      7.84
         |     90.00      7.43      4.60      1.93      4.82      7.13      8.89     41.68
---------+--------------------------------------------------------------------------------
.4000000 |     91.00      0.72      0.89      0.00      0.12      0.53      0.87      5.48
         |     91.00      2.80      1.49      0.00      1.86      2.39      3.28      7.60
         |     91.00      7.27      4.09      0.00      4.85      6.56      8.73     34.04
---------+--------------------------------------------------------------------------------
```

```
   Total |    407.00       0.85       0.96       0.00       0.34       0.60       0.96       9.14
         |    407.00       2.98       1.34       0.00       2.04       2.68       3.64       7.84
         |    407.00       7.41       4.10       0.00       5.01       6.86       8.89      41.68
-------------------------------------------------------------------------------------------------
```

```
#### Problem 2: creating variables
. egen nput= count(put) if put!=., by(ptid)

. egen nspd= count(spd) if spd!=., by(ptid)

. egen nspm= count(spm) if spm!=., by(ptid)

#### Tabulating counts for each variable
. table dose nput, row col

----------------------------------------
          |                nput
     dose |      1       2       3       4   Total
----------+-----------------------------------
    0.000 |      1       4      12     100     117
    0.075 |      1       4              104     109
    0.200 |      1       4       9      76      90
    0.400 |      3       8      12      68      91
          |
    Total |      6      20      33     348     407
----------------------------------------

. table dose nspd, row col

----------------------------------------
          |                nspd
     dose |      1       2       3       4   Total
----------+-----------------------------------
    0.000 |      1       4      12     100     117
    0.075 |      1       4              104     109
    0.200 |      1       4       9      76      90
    0.400 |      3       8      12      68      91
          |
    Total |      6      20      33     348     407
----------------------------------------

. table dose nspm, row col
```

```
-------------------------------------------
           |                nspm
      dose |      1       2       3       4   Total
-----------+-------------------------------------
     0.000 |      1       4      12     100     117
     0.075 |      1       4              104     109
     0.200 |      1       4       9      76      90
     0.400 |      3       8      12      68      91
           |
     Total |      6      20      33     348     407
-------------------------------------------
```

Verifying that cases have nonmissing values for put, spd, and spm

```
. table dose if nput==4 & nspd==4 & nspm==4

----------------------
     dose |      Freq.
----------+-----------
    0.000 |         100
    0.075 |         104
    0.200 |          76
    0.400 |          68
----------------------

. table nput nspd if time==0

----------------------------------
           |           nspd
     nput |    1      2     3     4
----------+-----------------------
        1 |    6
        2 |          10
        3 |                11
        4 |                      87
----------------------------------

. table nput nspm if time==0

----------------------------------
           |           nspm
     nput |    1      2     3     4
```

```
----------+-----------------------
        1 |     6
        2 |          10
        3 |                11
        4 |                      87
----------------------------------
```

Verifying that each unique ptid has a month 0 measurement

```
. egen mintime=min(time), by(ptid)

. table dose mintime

----------------
          |minti
          |  me
     dose |    0
----------+-----
    0.000 |  117
    0.075 |  109
    0.200 |   90
    0.400 |   91
----------------
```

Now getting counts for each patient

```
. table dose nput if time==0, row col

-----------------------------------------------
          |                nput
     dose |      1       2       3       4  Total
----------+------------------------------------
    0.000 |      1       2       4      25     32
    0.075 |      1       2              26     29
    0.200 |      1       2       3      19     25
    0.400 |      3       4       4      17     28
          |
    Total |      6      10      11      87    114
-----------------------------------------------

. table dose nspd if time==0, row col

-----------------------------------------------
          |               nspd
```

```
     dose |      1       2       3       4   Total
----------+------------------------------------
    0.000 |      1       2       4      25      32
    0.075 |      1       2               26      29
    0.200 |      1       2       3      19      25
    0.400 |      3       4       4      17      28
          |
    Total |      6      10      11      87     114
-------------------------------------------------
```

```
. table dose nspm if time==0, row col
```

```
-------------------------------------------
          |                nspm
     dose |      1       2       3       4   Total
----------+------------------------------------
    0.000 |      1       2       4      25      32
    0.075 |      1       2               26      29
    0.200 |      1       2       3      19      25
    0.400 |      3       4       4      17      28
          |
    Total |      6      10      11      87     114
-------------------------------------------
```

```
#### Problem 3: Creating variables
. egen mnput=mean(put), by(ptid)

. egen mnspd=mean(spd), by(ptid)

. egen mnspm=mean(spm), by(ptid)

. format mnput mnspd mnspm %9.2f

#### Descriptive statistics for all cases
. tabstat mnput mnspd mnspm, stat(n mean sd min q max) col(stat) format by(dose)

Summary for variables: mnput mnspd mnspm
     by categories of: dose

    dose |         N      mean        sd       min       p25       p50       p75       max
---------+--------------------------------------------------------------------------------
       0 |    117.00      1.09      0.66      0.34      0.53      0.99      1.25      3.24
```

```
         |    117.00      3.16      0.88      1.87      2.62      2.93      3.75      5.26
         |    117.00      7.17      2.30      4.68      5.81      6.79      7.81     16.43
---------+-------------------------------------------------------------------------------
.0750000 |    109.00      0.84      0.39      0.11      0.54      0.73      1.23      1.75
         |    109.00      3.00      0.70      1.95      2.52      2.98      3.47      6.62
         |    109.00      7.76      2.04      5.42      6.28      7.63      8.56     15.81
---------+-------------------------------------------------------------------------------
.2000000 |     90.00      0.71      0.32      0.12      0.44      0.69      0.95      1.28
         |     90.00      2.92      0.91      1.59      2.43      2.73      3.18      6.86
         |     90.00      7.43      2.28      5.77      5.96      6.90      7.97     16.56
---------+-------------------------------------------------------------------------------
.4000000 |     91.00      0.72      0.52      0.04      0.36      0.53      0.94      1.98
         |     91.00      2.80      0.78      1.54      2.11      2.66      3.19      5.21
         |     91.00      7.27      1.91      4.17      5.90      7.06      8.44     13.70
---------+-------------------------------------------------------------------------------
   Total |    407.00      0.85      0.52      0.04      0.47      0.75      1.12      3.24
         |    407.00      2.98      0.83      1.54      2.42      2.92      3.34      6.86
         |    407.00      7.41      2.15      4.17      5.91      7.01      8.34     16.56
-------------------------------------------------------------------------------------------
```

Descriptive statistics on a per patient basis

```
. tabstat mnput mnspd mnspm if time==0, stat(n mean sd min q max) col(stat) format by(dose)

Summary for variables: mnput mnspd mnspm
     by categories of: dose

    dose |         N      mean        sd       min       p25       p50       p75       max
---------+-------------------------------------------------------------------------------
       0 |     32.00      1.06      0.65      0.34      0.59      0.99      1.19      3.24
         |     32.00      3.14      0.89      1.87      2.57      2.91      3.61      5.26
         |     32.00      7.09      2.26      4.68      5.81      6.76      7.71     16.43
---------+-------------------------------------------------------------------------------
.0750000 |     29.00      0.80      0.41      0.11      0.50      0.72      1.13      1.75
         |     29.00      3.09      0.92      1.95      2.52      2.98      3.63      6.62
         |     29.00      7.79      2.02      5.42      6.54      7.63      8.56     15.81
---------+-------------------------------------------------------------------------------
.2000000 |     25.00      0.70      0.32      0.12      0.44      0.67      0.94      1.28
         |     25.00      3.01      1.05      1.59      2.49      2.84      3.18      6.86
         |     25.00      7.44      2.25      5.77      6.11      6.90      7.97     16.56
---------+-------------------------------------------------------------------------------
.4000000 |     28.00      0.71      0.54      0.04      0.34      0.52      0.96      1.98
         |     28.00      2.76      0.83      1.54      2.10      2.59      3.18      5.21
```

```
         |     28.00      7.20      1.89      4.17      5.80      7.19      8.39     13.70
---------+--------------------------------------------------------------------------------
   Total |    114.00      0.83      0.52      0.04      0.44      0.72      1.10      3.24
         |    114.00      3.01      0.92      1.54      2.41      2.90      3.34      6.86
         |    114.00      7.37      2.10      4.17      5.91      7.00      8.34     16.56
------------------------------------------------------------------------------------------

#### Problem 4: For each part, I create variables, then produce the statistics in a table
. egen grbg=mean(put) if time>0, by(ptid)
(114 missing values generated)

. egen mtrtput=mean(grbg), by(ptid)
(6 missing values generated)

. drop grbg

. egen grbg=mean(spd) if time>0, by(ptid)
(114 missing values generated)

. egen mtrtspd=mean(grbg), by(ptid)
(6 missing values generated)

. drop grbg

. egen grbg=mean(spm) if time>0, by(ptid)
(114 missing values generated)

. egen mtrtspm=mean(grbg), by(ptid)
(6 missing values generated)

. drop grbg

. format mtrtput mtrtspd mtrtspm %9.2f

. tabstat mtrtput mtrtspd mtrtspm if time==0, stat(n mean sd min q max) col(stat) format by(dose)

Summary for variables: mtrtput mtrtspd mtrtspm
     by categories of: dose

    dose |         N      mean        sd       min       p25       p50       p75       max
---------+--------------------------------------------------------------------------------
       0 |     31.00      1.23      0.87      0.14      0.59      1.06      1.34      4.26
```

```
         |     31.00      3.14      0.88      1.95      2.53      2.96      3.92      5.24
         |     31.00      6.76      1.85      3.98      5.47      6.55      8.08     11.75
---------+--------------------------------------------------------------------------------
.0750000 |     28.00      0.87      0.46      0.00      0.51      0.88      1.20      1.97
         |     28.00      2.80      0.54      1.79      2.48      2.80      3.05      3.93
         |     28.00      7.56      1.57      4.41      6.38      7.66      8.90     10.19
---------+--------------------------------------------------------------------------------
.2000000 |     24.00      0.72      0.38      0.16      0.49      0.60      1.06      1.64
         |     24.00      2.89      1.29      1.27      2.28      2.63      3.09      7.84
         |     24.00      6.95      1.79      4.74      5.56      6.28      8.71     11.52
---------+--------------------------------------------------------------------------------
.4000000 |     25.00      0.73      0.59      0.12      0.32      0.54      1.04      2.40
         |     25.00      2.42      0.79      1.28      1.90      2.07      2.81      4.61
         |     25.00      7.09      2.04      2.83      5.97      7.05      8.37     12.25
---------+--------------------------------------------------------------------------------
   Total |    108.00      0.91      0.65      0.00      0.48      0.81      1.19      4.26
         |    108.00      2.83      0.93      1.27      2.17      2.70      3.14      7.84
         |    108.00      7.08      1.82      2.83      5.77      6.89      8.35     12.25
--------------------------------------------------------------------------------------------
```

```
. g dmtrtput=mtrtput-put
(6 missing values generated)

. g dmtrtspd=mtrtspd-spd
(6 missing values generated)

. g dmtrtspm=mtrtspm-spm
(6 missing values generated)

. format dmtrtput dmtrtspd dmtrtspm %9.2f

. tabstat dmtrtput dmtrtspd dmtrtspm if time==0, stat(n mean sd min q max) col(stat) format by(dose)

Summary for variables: dmtrtput dmtrtspd dmtrtspm
     by categories of: dose

    dose |         N      mean        sd       min       p25       p50       p75       max
---------+--------------------------------------------------------------------------------
       0 |     31.00      0.56      0.98     -0.99      0.16      0.40      0.75      4.08
         |     31.00     -0.14      1.31     -3.06     -0.84      0.01      0.85      1.96
         |     31.00     -1.51      5.32    -25.49     -2.01     -1.07      0.56      6.47
---------+--------------------------------------------------------------------------------
```

```
.0750000 |     28.00      0.21      0.57     -1.70     -0.13      0.23      0.65      0.91
         |     28.00     -0.56      1.41     -4.51     -1.48     -0.31      0.55      1.21
         |     28.00     -0.85      5.95    -29.14     -1.19     -0.23      1.45      4.53
---------+-------------------------------------------------------------------------------
.2000000 |     24.00      0.11      0.53     -1.37     -0.20      0.16      0.35      1.43
         |     24.00     -0.48      1.47     -3.88     -1.19     -0.37      0.44      1.98
         |     24.00     -2.20      7.23    -33.50     -2.83     -1.08      0.36      6.27
---------+-------------------------------------------------------------------------------
.4000000 |     25.00      0.07      0.63     -1.40     -0.18      0.04      0.40      1.66
         |     25.00     -1.28      2.15     -5.35     -2.92     -0.82      0.16      2.41
         |     25.00     -1.19      6.18    -27.13     -1.50     -0.81      1.42      7.62
---------+-------------------------------------------------------------------------------
   Total |    108.00      0.26      0.73     -1.70     -0.11      0.19      0.50      4.08
         |    108.00     -0.59      1.63     -5.35     -1.43     -0.22      0.53      2.41
         |    108.00     -1.42      6.07    -33.50     -1.94     -0.89      1.24      7.62
-------------------------------------------------------------------------------------------

. egen grbg=mean(put) if time>0 & time < 15, by(ptid)
(206 missing values generated)

. egen mdrgput=mean(grbg), by(ptid)
(8 missing values generated)

. drop grbg

. egen grbg=mean(spd) if time>0 & time < 15, by(ptid)
(206 missing values generated)

. egen mdrgspd=mean(grbg), by(ptid)
(8 missing values generated)

. drop grbg

. egen grbg=mean(spm) if time>0 & time < 15, by(ptid)
(206 missing values generated)

. egen mdrgspm=mean(grbg), by(ptid)
(8 missing values generated)

. drop grbg

. format mdrgput mdrgspd mdrgspm %9.2f
```

```
. tabstat mdrgput mdrgspd mdrgspm if time==0, stat(n mean sd min q max) col(stat) format by(dose)

Summary for variables: mdrgput mdrgspd mdrgspm
     by categories of: dose

    dose |         N      mean        sd       min       p25       p50       p75       max
---------+--------------------------------------------------------------------------------
       0 |     31.00      1.09      0.94      0.03      0.59      0.80      1.44      5.47
         |     31.00      3.39      1.18      1.78      2.53      3.20      4.08      6.91
         |     31.00      6.89      2.24      2.32      5.14      6.79      8.33     11.59
---------+--------------------------------------------------------------------------------
.0750000 |     28.00      0.75      0.54      0.00      0.38      0.58      0.95      2.39
         |     28.00      2.75      0.73      1.56      2.34      2.67      3.27      4.43
         |     28.00      7.94      1.90      3.87      6.71      7.92      9.07     11.77
---------+--------------------------------------------------------------------------------
.2000000 |     23.00      0.63      0.47      0.12      0.30      0.47      1.05      1.72
         |     23.00      2.75      1.49      0.83      1.67      2.53      3.23      7.84
         |     23.00      7.22      2.50      3.41      4.94      7.02      8.09     12.04
---------+--------------------------------------------------------------------------------
.4000000 |     25.00      0.59      0.75      0.00      0.08      0.33      0.57      3.02
         |     25.00      2.34      0.95      1.07      1.85      2.07      2.59      4.70
         |     25.00      7.14      2.37      2.71      6.08      7.26      8.73     12.25
---------+--------------------------------------------------------------------------------
   Total |    107.00      0.78      0.74      0.00      0.33      0.54      1.20      5.47
         |    107.00      2.84      1.16      0.83      2.00      2.62      3.41      7.84
         |    107.00      7.29      2.25      2.32      5.95      7.29      8.73     12.25
--------------------------------------------------------------------------------------------

. g dmdrgput=mdrgput-put
(8 missing values generated)

. g dmdrgspd=mdrgspd-spd
(8 missing values generated)

. g dmdrgspm=mdrgspm-spm
(8 missing values generated)

. format dmdrgput dmdrgspd dmdrgspm %9.2f

. tabstat dmdrgput dmdrgspd dmdrgspm if time==0, stat(n mean sd min q max) col(stat) format by(dose)
```

```
Summary for variables: dmdrgput dmdrgspd dmdrgspm
     by categories of: dose
```

dose	N	mean	sd	min	p25	p50	p75	max
0	31.00	0.43	1.06	-1.11	0.08	0.26	0.59	5.28
	31.00	0.11	1.43	-2.89	-1.03	0.04	1.43	2.42
	31.00	-1.38	5.37	-25.49	-2.02	-0.40	0.64	6.47
.0750000	28.00	0.08	0.67	-1.97	-0.33	-0.00	0.55	1.59
	28.00	-0.61	1.33	-3.65	-1.52	-0.48	0.26	1.37
	28.00	-0.47	5.93	-28.59	-1.09	0.24	1.78	4.76
.2000000	23.00	-0.01	0.59	-1.62	-0.28	-0.14	0.28	1.51
	23.00	-0.60	1.41	-4.02	-1.25	-0.42	0.33	1.98
	23.00	-1.98	7.83	-34.30	-3.38	-0.94	1.26	9.50
.4000000	25.00	-0.06	0.78	-1.40	-0.57	-0.18	0.19	2.28
	25.00	-1.36	2.26	-5.61	-3.10	-1.31	0.15	2.50
	25.00	-1.14	6.50	-27.80	-1.22	-0.10	1.44	7.62
Total	107.00	0.13	0.83	-1.97	-0.28	0.06	0.43	5.28
	107.00	-0.58	1.70	-5.61	-1.59	-0.37	0.51	2.50
	107.00	-1.21	6.30	-34.30	-2.15	-0.10	1.26	9.50

Problem 5

```
. egen grbg=min(put) if time>0 & time < 15, by(ptid)
(206 missing values generated)

. egen mindrgput=mean(grbg), by(ptid)
(8 missing values generated)

. drop grbg

. format mindrgput  %9.2f

. tabstat mindrgput if time==0, stat(n mean sd min q max) col(stat) format by(dose)

Summary for variables: mindrgput
     by categories of: dose
```

```
    dose |         N      mean        sd       min       p25       p50       p75       max
---------+--------------------------------------------------------------------------------
       0 |     31.00      0.67      0.40      0.00      0.43      0.61      0.81      1.80
.0750000 |     28.00      0.39      0.23      0.00      0.26      0.38      0.56      0.82
.2000000 |     23.00      0.33      0.34      0.00      0.15      0.26      0.42      1.43
.4000000 |     25.00      0.28      0.44      0.00      0.00      0.02      0.38      1.73
---------+--------------------------------------------------------------------------------
   Total |    107.00      0.43      0.39      0.00      0.17      0.38      0.60      1.80
--------------------------------------------------------------------------------------------
```

```
. g dmindrgput= mindrgput - put
(8 missing values generated)

. format dmindrgput  %9.2f

. tabstat dmindrgput if time==0, stat(n mean sd min q max) col(stat) format by(dose)

Summary for variables: dmindrgput
     by categories of: dose

    dose |         N      mean        sd       min       p25       p50       p75       max
---------+--------------------------------------------------------------------------------
       0 |     31.00      0.01      0.60     -1.19     -0.21     -0.06      0.29      1.62
.0750000 |     28.00     -0.27      0.55     -2.31     -0.51     -0.24      0.07      0.71
.2000000 |     23.00     -0.30      0.47     -1.78     -0.57     -0.36      0.02      0.70
.4000000 |     25.00     -0.37      0.47     -1.40     -0.63     -0.38     -0.13      0.56
---------+--------------------------------------------------------------------------------
   Total |    107.00     -0.22      0.55     -2.31     -0.51     -0.21      0.08      1.62
--------------------------------------------------------------------------------------------
```