# STAT:5400

## Computing in Statistics
## One approach to handling missing data
## Arrays and looping in SAS

**Lecture 28 Nov. 8, 2017**

**Kate Cowles**
**374 SH**
**kate-cowles@uiowa.edu**



## Example dataset

ACTG 320 (Hammer, Squires, *et al.*, 1997) was a randomized, double-blind, placebo-controlled trial comparing a three-drug regimen (indinavir, lamivudine, and either zidovudine or stavudine) with a two-drug regimen (zidovudine and lamivudine) in HIV-infected adults with CD4 counts <= 200 and at least 3 months of prior zidovudine therapy. The 1156 randomized patients were stratified according to their CD4 count ($\leq$ 50 cells/mm$^3$ or 50-200 cells/mm$^3$) at study entry. The primary endpoint was occurrence of an AIDS-defining event (according to the CDC definition) or death. In addition, blood specimens were collected at baseline and at weeks 4, 8, 24, and 40 during follow-up for analysis of CD4 counts and viral load. The ACTG 320 dataset available for purchase from the National Technical Information Service includes clinical endpoints and CD4 data for all patients but viral load data on only 198 patients who were randomly selected for a virology substudy.



## Example dataset, continued

- includes the 198 patients who have RNA data
- variables are

```
trt     -- 1/0 treatment group indicator
strat   -- 1/0 stratification group indicator
rna1    -- week 0 RNA
rna2    -- week 4 RNA
rna3    -- week 8 RNA
rna4    -- week 24 RNA
rna5    -- week 40 RNA
cd41    -- similar 5 cd4 values
cd42
cd43
cd44
cd45
obst    -- time at which clinical endpoint occurred, or last
           at which patient was observed and no clinical end
fail    -- 1: cliical endpoint; 0: no clinical endpoint
```

- Note: no patient identifier



## What we would like to do

- impute values for missing RNA and CD4 data
- calculate patient-specific rates of change of RNA by week and of CD4 by week
  - how will data file have to be laid out to do this?

## Last-value-carried forward

- one (not terribly good) method of imputing missing values of longitudinal data
- may make sense if values are "missing at random"
  - that is, if the probability that a value is missing doesn't depend on the value that would have been observed
  - not likely to be the case for this type of data



```
options linesize = 72 ;

data actg320 ;
infile '/group/ftp/pub/kcowles/datasets/combo1.dat' firstobs = 2 ;
input trt strat rna1 rna2 rna3 rna4 rna5 cd41 cd42 cd43 cd44
      cd45 obst fail ;
pid = _N_ ;      * copy observation number into permanent variable ;
run ;

proc print data = actg320 (obs = 12) ;
title 'no arrays used' ;
run ;
```



no arrays used 1
18:26 Sunday, June 29, 2003

| Obs | trt | strat | rna1 | rna2 | rna3 | rna4 | rna5 |
|---|---|---|---|---|---|---|---|
| 1 | 0 | 1 | 4.24790 | 3.27323 | 4.05660 | 3.98290 | . |
| 2 | 0 | 0 | 5.56951 | 5.10036 | 4.96781 | 5.41695 | 5.01041 |
| 3 | 0 | 0 | 4.96314 | 5.48520 | 4.60326 | 5.29003 | 5.96755 |
| 4 | 0 | 0 | 3.91666 | 3.53046 | 3.96881 | 2.03342 | . |
| 5 | 0 | 1 | 3.36286 | 2.29667 | 2.44560 | 3.55835 | 4.20850 |
| 6 | 1 | 1 | 4.65297 | 3.03342 | 2.71517 | . | . |
| 7 | 1 | 1 | 5.09670 | 2.87157 | 3.89856 | 2.14613 | . |
| 8 | 0 | 0 | 5.10824 | 5.14251 | 5.04452 | 2.37291 | 2.26007 |
| 9 | 1 | 0 | 5.52962 | 3.28825 | 2.88081 | 5.36319 | . |
| 10 | 1 | 1 | 4.97237 | 2.48001 | 2.13672 | 2.23045 | 2.30320 |
| 11 | 1 | 1 | 5.71936 | 3.45894 | 2.58995 | 2.25285 | . |
| 12 | 1 | 0 | 5.85708 | 3.18949 | 2.90309 | 2.02531 | 2.69984 |

| Obs | cd41 | cd42 | cd43 | cd44 | cd45 | obst | fail | pid |
|---|---|---|---|---|---|---|---|---|
| 1 | 188.5 | 152 | 178 | 148 | . | 32.8571 | 0 | 1 |
| 2 | 19.0 | 43 | 35 | 7 | . | 43.8571 | 0 | 2 |
| 3 | 15.0 | 15 | 19 | 15 | . | 38.8571 | 0 | 3 |
| 4 | 30.0 | 30 | 40 | 70 | . | 49.7143 | 0 | 4 |
| 5 | 190.5 | 205 | 255 | 301 | 243 | 46.0000 | 0 | 5 |
| 6 | 128.5 | 166 | 867 | . | . | 24.8571 | 0 | 6 |
| 7 | 33.0 | 96 | 100 | 159 | . | 43.8571 | 0 | 7 |
| 8 | 10.0 | 10 | 11 | 72 | 67 | 13.0000 | 1 | 8 |
| 9 | 8.0 | 33 | 48 | 7 | . | 13.0000 | 1 | 9 |
| 10 | 139.0 | 178 | 119 | 305 | 305 | 48.7143 | 0 | 10 |
| 11 | 90.0 | 243 | 154 | 266 | . | 42.5714 | 0 | 11 |
| 12 | 20.0 | 155 | 172 | 160 | 142 | 49.0000 | 0 | 12 |



## Arrays in SAS datasteps

- enable referencing a group of SAS variables by a single name and subscripts
- defined in `array` statements

```
array  arrayname[number of items] I
names or (list of
                values)
```

- exist during execution of data step

```
data actg320 ;
set actg320 ;
array arna[5] rna1 rna2 rna3 rna4 rna5 ;  * define array and make it a copy
                                            of existing variables in dataset
array acd4[5] cd41 cd42 cd43 cd44 cd45 ;
run ;

proc print data = actg320 (obs = 12) ;
title 'first way of doing arrays' ;
run ;
```



| Obs | trt | strat | rna1 | rna2 | rna3 | rna4 | rna5 |
|---|---|---|---|---|---|---|---|
| 1 | 0 | 1 | 4.24790 | 3.27323 | 4.05660 | 3.98290 | . |
| 2 | 0 | 0 | 5.56951 | 5.10036 | 4.96781 | 5.41695 | 5.01041 |
| 3 | 0 | 0 | 4.96314 | 5.48520 | 4.60326 | 5.29003 | 5.96755 |
| 4 | 0 | 0 | 3.91666 | 3.53046 | 3.96881 | 2.03342 | . |
| 5 | 0 | 1 | 3.36286 | 2.29667 | 2.44560 | 3.55835 | 4.20850 |
| 6 | 1 | 1 | 4.65297 | 3.03342 | 2.71517 | . | . |
| 7 | 1 | 1 | 5.09670 | 2.87157 | 3.89856 | 2.14613 | . |
| 8 | 0 | 0 | 5.10824 | 5.14251 | 5.04452 | 2.37291 | 2.26007 |
| 9 | 1 | 0 | 5.52962 | 3.28825 | 2.88081 | 5.36319 | . |
| 10 | 1 | 1 | 4.97237 | 2.48001 | 2.13672 | 2.23045 | 2.30320 |
| 11 | 1 | 1 | 5.71936 | 3.45894 | 2.58995 | 2.25285 | . |
| 12 | 1 | 0 | 5.85708 | 3.18949 | 2.90309 | 2.02531 | 2.69984 |

| Obs | cd41 | cd42 | cd43 | cd44 | cd45 | obst | fail | pid |
|---|---|---|---|---|---|---|---|---|
| 1 | 188.5 | 152 | 178 | 148 | . | 32.8571 | 0 | 1 |
| 2 | 19.0 | 43 | 35 | 7 | . | 43.8571 | 0 | 2 |



| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 3 | 15.0 | 15 | 19 | 15 | . | 38.8571 | 0 | 3 |
| 4 | 30.0 | 30 | 40 | 70 | . | 49.7143 | 0 | 4 |
| 5 | 190.5 | 205 | 255 | 301 | 243 | 46.0000 | 0 | 5 |
| 6 | 128.5 | 166 | 867 | . | . | 24.8571 | 0 | 6 |
| 7 | 33.0 | 96 | 100 | 159 | . | 43.8571 | 0 | 7 |
| 8 | 10.0 | 10 | 11 | 72 | 67 | 13.0000 | 1 | 8 |
| 9 | 8.0 | 33 | 48 | 7 | . | 13.0000 | 1 | 9 |
| 10 | 139.0 | 178 | 119 | 305 | 305 | 48.7143 | 0 | 10 |
| 11 | 90.0 | 243 | 154 | 266 | . | 42.5714 | 0 | 11 |
| 12 | 20.0 | 155 | 172 | 160 | 142 | 49.0000 | 0 | 12 |



```
data actg320 ;
set actg320 ;
array arna[5] rna1 - rna5 ;
array acd4[5] cd41 - cd45 ;
run ;

proc print data = actg320 (obs = 12) ;
title 'second way of doing arrays' ;
run ;
```



| Obs | trt | strat | rna1 | rna2 | rna3 | rna4 | rna5 |
|---|---|---|---|---|---|---|---|
| 1 | 0 | 1 | 4.24790 | 3.27323 | 4.05660 | 3.98290 | . |
| 2 | 0 | 0 | 5.56951 | 5.10036 | 4.96781 | 5.41695 | 5.01041 |
| 3 | 0 | 0 | 4.96314 | 5.48520 | 4.60326 | 5.29003 | 5.96755 |
| 4 | 0 | 0 | 3.91666 | 3.53046 | 3.96881 | 2.03342 | . |
| 5 | 0 | 1 | 3.36286 | 2.29667 | 2.44560 | 3.55835 | 4.20850 |
| 6 | 1 | 1 | 4.65297 | 3.03342 | 2.71517 | . | . |
| 7 | 1 | 1 | 5.09670 | 2.87157 | 3.89856 | 2.14613 | . |
| 8 | 0 | 0 | 5.10824 | 5.14251 | 5.04452 | 2.37291 | 2.26007 |
| 9 | 1 | 0 | 5.52962 | 3.28825 | 2.88081 | 5.36319 | . |
| 10 | 1 | 1 | 4.97237 | 2.48001 | 2.13672 | 2.23045 | 2.30320 |
| 11 | 1 | 1 | 5.71936 | 3.45894 | 2.58995 | 2.25285 | . |
| 12 | 1 | 0 | 5.85708 | 3.18949 | 2.90309 | 2.02531 | 2.69984 |

| Obs | cd41 | cd42 | cd43 | cd44 | cd45 | obst | fail | pid |
|---|---|---|---|---|---|---|---|---|
| 1 | 188.5 | 152 | 178 | 148 | . | 32.8571 | 0 | 1 |
| 2 | 19.0 | 43 | 35 | 7 | . | 43.8571 | 0 | 2 |
| 3 | 15.0 | 15 | 19 | 15 | . | 38.8571 | 0 | 3 |
| 4 | 30.0 | 30 | 40 | 70 | . | 49.7143 | 0 | 4 |
| 5 | 190.5 | 205 | 255 | 301 | 243 | 46.0000 | 0 | 5 |
| 6 | 128.5 | 166 | 867 | . | . | 24.8571 | 0 | 6 |
| 7 | 33.0 | 96 | 100 | 159 | . | 43.8571 | 0 | 7 |
| 8 | 10.0 | 10 | 11 | 72 | 67 | 13.0000 | 1 | 8 |
| 9 | 8.0 | 33 | 48 | 7 | . | 13.0000 | 1 | 9 |
| 10 | 139.0 | 178 | 119 | 305 | 305 | 48.7143 | 0 | 10 |
| 11 | 90.0 | 243 | 154 | 266 | . | 42.5714 | 0 | 11 |
| 12 | 20.0 | 155 | 172 | 160 | 142 | 49.0000 | 0 | 12 |

## Do loops in SAS data steps

- enable coding a task once and having SAS execute it repeatedly
- framework

```
do <...> ;
     .
     .
     .
end ;
```

- in simplest form, do statement includes a "loop-counter" such as

```
do i = 1 to 5 ;
```



```
data actg320lvcf ;
set actg320 ;
array arna[5] rna1 - rna5 ;
array acd4[5] cd41 - cd45 ;

* set up last-value-carried-forward ;
do i = 2 to 5 ;
if arna[i] eq . then arna[i] = arna[i-1] ;
if acd4[i] eq . then acd4[i] = acd4[i-1] ;
end ;
run ;

proc print data = actg320lvcf (obs = 12) ;
title 'after last value carried forward' ;
run ;
```



```
                after last value carried forward                4
                                      18:26 Sunday, June 29, 2003

Obs   trt   strat     rna1      rna2      rna3      rna4      rna5

  1    0      1     4.24790   3.27323   4.05660   3.98290   3.98290
  2    0      0     5.56951   5.10036   4.96781   5.41695   5.01041
  3    0      0     4.96314   5.48520   4.60326   5.29003   5.96755
  4    0      0     3.91666   3.53046   3.96881   2.03342   2.03342
  5    0      1     3.36286   2.29667   2.44560   3.55835   4.20850
  6    1      1     4.65297   3.03342   2.71517   2.71517   2.71517
  7    1      1     5.09670   2.87157   3.89856   2.14613   2.14613
  8    0      0     5.10824   5.14251   5.04452   2.37291   2.26007
  9    1      0     5.52962   3.28825   2.88081   5.36319   5.36319
 10    1      1     4.97237   2.48001   2.13672   2.23045   2.30320
 11    1      1     5.71936   3.45894   2.58995   2.25285   2.25285
 12    1      0     5.85708   3.18949   2.90309   2.02531   2.69984

Obs    cd41   cd42   cd43   cd44   cd45     obst    fail   pid   i

  1   188.5    152    178    148    148   32.8571     0      1   6
  2    19.0     43     35      7      7   43.8571     0      2   6
  3    15.0     15     19     15     15   38.8571     0      3   6
  4    30.0     30     40     70     70   49.7143     0      4   6
  5   190.5    205    255    301    243   46.0000     0      5   6
  6   128.5    166    867    867    867   24.8571     0      6   6
  7    33.0     96    100    159    159   43.8571     0      7   6
  8    10.0     10     11     72     67   13.0000     1      8   6
  9     8.0     33     48      7      7   13.0000     1      9   6
 10   139.0    178    119    305    305   48.7143     0     10   6
 11    90.0    243    154    266    266   42.5714     0     11   6
 12    20.0    155    172    160    142   49.0000     0     12   6
```



```
data actg320lvcf ;
set actg320 ;
array arna[5] rna1 - rna5 ;
array acd4[5] cd41 - cd45 ;

* set up last-value-carried-forward ;
do i = 2 to 5 ;
if arna[i] eq . then arna[i] = arna[i-1] ;
if acd4[i] eq . then acd4[i] = acd4[i-1] ;
end ;
drop i ;    * drop loop counter ;
run ;
```

```
data actg320lvcf ;
set actg320 ;
array aweek[5] (0, 4, 8, 12, 24) ; * define array and assign numeric values
                                    to its elements ;
array arna[5] rna1 - rna5 ;
array acd4[5] cd41 - cd45 ;

* set up better last-value-carried-forward ;
* do not carry values forward to times later than obst ;
do i = 2 to 5 ;
if (arna[i] eq . and obst > aweek[i]) then arna[i] = arna[i-1] ;
if (acd4[i] eq . and obst > aweek[i]) then acd4[i] = acd4[i-1] ;
end ;
drop i ;    * drop loop counter ;
run ;
```



# Using arrays to create a dataset with one record per week per patient

```
data actg320oneper ;
set actg320 ;   * might do this with actg320lvcf instead ;
array aweek[5] (0, 4, 8, 12, 24) ; * define array and assign numeric values
                                    to its elements ;
array arna[5] rna1 - rna5 ;
array acd4[5] cd41 - cd45 ;

do i = 1 to 5 ;
week = aweek[i] ;
rna = arna[i] ;
cd4 = acd4[i] ;
output ;
end ;
run ;

proc print data = actg320oneper (obs = 25) ;
title 'actg320oneper' ;
run ;
```



actg320oneper 5

18:26 Sunday, June 29, 2003

| 0 | s t | r | r | r | r | r | c | c |
|---|---|---|---|---|---|---|---|---|
| b | r | a | n | n | n | n | n | d | d |
| s | t | t | a | a | a | a | a | 4 | 4 |
| | | | 1 | 2 | 3 | 4 | 5 | 1 | 2 |
| 1 | 0 | 1 | 4.24790 | 3.27323 | 4.05660 | 3.98290 | . | 188.5 | 152 |
| 2 | 0 | 1 | 4.24790 | 3.27323 | 4.05660 | 3.98290 | . | 188.5 | 152 |
| 3 | 0 | 1 | 4.24790 | 3.27323 | 4.05660 | 3.98290 | . | 188.5 | 152 |
| 4 | 0 | 1 | 4.24790 | 3.27323 | 4.05660 | 3.98290 | . | 188.5 | 152 |
| 5 | 0 | 1 | 4.24790 | 3.27323 | 4.05660 | 3.98290 | . | 188.5 | 152 |
| 6 | 0 | 0 | 5.56951 | 5.10036 | 4.96781 | 5.41695 | 5.01041 | 19.0 | 43 |
| 7 | 0 | 0 | 5.56951 | 5.10036 | 4.96781 | 5.41695 | 5.01041 | 19.0 | 43 |
| 8 | 0 | 0 | 5.56951 | 5.10036 | 4.96781 | 5.41695 | 5.01041 | 19.0 | 43 |
| 9 | 0 | 0 | 5.56951 | 5.10036 | 4.96781 | 5.41695 | 5.01041 | 19.0 | 43 |
| 10 | 0 | 0 | 5.56951 | 5.10036 | 4.96781 | 5.41695 | 5.01041 | 19.0 | 43 |
| 11 | 0 | 0 | 4.96314 | 5.48520 | 4.60326 | 5.29003 | 5.96755 | 15.0 | 15 |
| 12 | 0 | 0 | 4.96314 | 5.48520 | 4.60326 | 5.29003 | 5.96755 | 15.0 | 15 |
| 13 | 0 | 0 | 4.96314 | 5.48520 | 4.60326 | 5.29003 | 5.96755 | 15.0 | 15 |
| 14 | 0 | 0 | 4.96314 | 5.48520 | 4.60326 | 5.29003 | 5.96755 | 15.0 | 15 |
| 15 | 0 | 0 | 4.96314 | 5.48520 | 4.60326 | 5.29003 | 5.96755 | 15.0 | 15 |
| 16 | 0 | 0 | 3.91666 | 3.53046 | 3.96881 | 2.03342 | . | 30.0 | 30 |
| 17 | 0 | 0 | 3.91666 | 3.53046 | 3.96881 | 2.03342 | . | 30.0 | 30 |
| 18 | 0 | 0 | 3.91666 | 3.53046 | 3.96881 | 2.03342 | . | 30.0 | 30 |
| 19 | 0 | 0 | 3.91666 | 3.53046 | 3.96881 | 2.03342 | . | 30.0 | 30 |
| 20 | 0 | 0 | 3.91666 | 3.53046 | 3.96881 | 2.03342 | . | 30.0 | 30 |
| 21 | 0 | 1 | 3.36286 | 2.29667 | 2.44560 | 3.55835 | 4.20850 | 190.5 | 205 |
| | | | a w | a w | a w | w | w | | |



| 0 | c d 4 | c d 4 | c d 4 | o b s | f a i l | p i d | e e k 1 | e e k 2 | e e k 3 | e e k 4 | e e k 5 | w e i k | r a | c d 4 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 178 | 148 | . | 32.8571 | 0 | 1 | 0 | 4 | 8 | 12 | 24 | 1 | 0 | 4.24790 | 188.5 |
| 2 | 178 | 148 | . | 32.8571 | 0 | 1 | 0 | 4 | 8 | 12 | 24 | 2 | 4 | 3.27323 | 152.0 |
| 3 | 178 | 148 | . | 32.8571 | 0 | 1 | 0 | 4 | 8 | 12 | 24 | 3 | 8 | 4.05660 | 178.0 |
| 4 | 178 | 148 | . | 32.8571 | 0 | 1 | 0 | 4 | 8 | 12 | 24 | 4 | 12 | 3.98290 | 148.0 |
| 5 | 178 | 148 | . | 32.8571 | 0 | 1 | 0 | 4 | 8 | 12 | 24 | 5 | 24 | . | . |
| 6 | 35 | 7 | . | 43.8571 | 0 | 2 | 0 | 4 | 8 | 12 | 24 | 1 | 0 | 5.56951 | 19.0 |
| 7 | 35 | 7 | . | 43.8571 | 0 | 2 | 0 | 4 | 8 | 12 | 24 | 2 | 4 | 5.10036 | 43.0 |
| 8 | 35 | 7 | . | 43.8571 | 0 | 2 | 0 | 4 | 8 | 12 | 24 | 3 | 8 | 4.96781 | 35.0 |
| 9 | 35 | 7 | . | 43.8571 | 0 | 2 | 0 | 4 | 8 | 12 | 24 | 4 | 12 | 5.41695 | 7.0 |
| 10 | 35 | 7 | . | 43.8571 | 0 | 2 | 0 | 4 | 8 | 12 | 24 | 5 | 24 | 5.01041 | . |
| 11 | 19 | 15 | . | 38.8571 | 0 | 3 | 0 | 4 | 8 | 12 | 24 | 1 | 0 | 4.96314 | 15.0 |
| 12 | 19 | 15 | . | 38.8571 | 0 | 3 | 0 | 4 | 8 | 12 | 24 | 2 | 4 | 5.48520 | 15.0 |
| 13 | 19 | 15 | . | 38.8571 | 0 | 3 | 0 | 4 | 8 | 12 | 24 | 3 | 8 | 4.60326 | 19.0 |
| 14 | 19 | 15 | . | 38.8571 | 0 | 3 | 0 | 4 | 8 | 12 | 24 | 4 | 12 | 5.29003 | 15.0 |
| 15 | 19 | 15 | . | 38.8571 | 0 | 3 | 0 | 4 | 8 | 12 | 24 | 5 | 24 | 5.96755 | . |
| 16 | 40 | 70 | . | 49.7143 | 0 | 4 | 0 | 4 | 8 | 12 | 24 | 1 | 0 | 3.91666 | 30.0 |
| 17 | 40 | 70 | . | 49.7143 | 0 | 4 | 0 | 4 | 8 | 12 | 24 | 2 | 4 | 3.53046 | 30.0 |
| 18 | 40 | 70 | . | 49.7143 | 0 | 4 | 0 | 4 | 8 | 12 | 24 | 3 | 8 | 3.96881 | 40.0 |
| 19 | 40 | 70 | . | 49.7143 | 0 | 4 | 0 | 4 | 8 | 12 | 24 | 4 | 12 | 2.03342 | 70.0 |
| 20 | 40 | 70 | . | 49.7143 | 0 | 4 | 0 | 4 | 8 | 12 | 24 | 5 | 24 | . | . |
| 21 | 255 | 301 | 243 | 46.0000 | 0 | 5 | 0 | 4 | 8 | 12 | 24 | 1 | 0 | 3.36286 | 190.5 |



| 0 | s t | r | r | r | r | r | c | c |
|---|---|---|---|---|---|---|---|---|
| b | r | a | n | n | n | n | n | d | d |
| | | | a | a | a | a | a | 4 | 4 |

```
s  t  t     1        2        3        4        5       1     2

22 0  1  3.36286  2.29667  2.44560  3.55835  4.20850  190.5  205
23 0  1  3.36286  2.29667  2.44560  3.55835  4.20850  190.5  205
24 0  1  3.36286  2.29667  2.44560  3.55835  4.20850  190.5  205
25 0  1  3.36286  2.29667  2.44560  3.55835  4.20850  190.5  205

                          a a a  a  a
                          w w w  w  w
    c    c  c     o    f   e e e  e  e    w
0   d    d  d     b    a p e e e  e  e    e    r      c
b   4    4  4     s    i i k k k  k  k    e    n      d
s   3    4  5     t    l d 1 2 3  4  5 i  k    a      4

22 255 301 243 46.0000 0 5 0 4 8 12 24 2  4 2.29667 205.0
23 255 301 243 46.0000 0 5 0 4 8 12 24 3  8 2.44560 255.0
24 255 301 243 46.0000 0 5 0 4 8 12 24 4 12 3.55835 301.0
25 255 301 243 46.0000 0 5 0 4 8 12 24 5 24 4.20850 243.0
```



```
data actg320oneper ;
set actg320oneper (keep = pid trt strat rna cd4 week) ;
run ;

proc print data = actg320oneper (obs = 25) ;
title 'actg320oneper after keep' ;
run ;
```



```
                    actg320oneper after keep                      7
                                      18:26 Sunday, June 29, 2003

     Obs    trt    strat    pid    week      rna       cd4

       1      0        1      1       0     4.24790    188.5
       2      0        1      1       4     3.27323    152.0
       3      0        1      1       8     4.05660    178.0
       4      0        1      1      12     3.98290    148.0
       5      0        1      1      24      .           .
       6      0        0      2       0     5.56951     19.0
       7      0        0      2       4     5.10036     43.0
       8      0        0      2       8     4.96781     35.0
       9      0        0      2      12     5.41695      7.0
      10      0        0      2      24     5.01041      .
      11      0        0      3       0     4.96314     15.0
      12      0        0      3       4     5.48520     15.0
      13      0        0      3       8     4.60326     19.0
      14      0        0      3      12     5.29003     15.0
      15      0        0      3      24     5.96755      .
      16      0        0      4       0     3.91666     30.0
      17      0        0      4       4     3.53046     30.0
      18      0        0      4       8     3.96881     40.0
      19      0        0      4      12     2.03342     70.0
      20      0        0      4      24      .           .
      21      0        1      5       0     3.36286    190.5
      22      0        1      5       4     2.29667    205.0
      23      0        1      5       8     2.44560    255.0
      24      0        1      5      12     3.55835    301.0
      25      0        1      5      24     4.20850    243.0
```



```
proc reg data = actg320oneper ;
model rna = week ;
by pid ;
where pid < 5 ;
run ;
```

```
                    actg320oneper after keep                      8
                                      18:26 Sunday, June 29, 2003

-------------------------------- pid=1 --------------------------------

                        The REG Procedure
                          Model: MODEL1
                    Dependent Variable: rna

                      Analysis of Variance

                                 Sum of          Mean
Source                    DF     Squares        Square  F Value  Pr > F

Model                      1  0.00000676  0.00000676     0.00  0.9965
Error                      2     0.54488     0.27244
Corrected Total            3     0.54488


        Root MSE              0.52196    R-Square     0.0000
        Dependent Mean        3.89016    Adj R-Sq    -0.5000
        Coeff Var            13.41736


                      Parameter Estimates

                     Parameter       Standard
Variable     DF       Estimate          Error    t Value    Pr > |t|
```

| | | | | | |
|---|---|---|---|---|---|
| Intercept | 1 | 3.89190 | 0.43670 | 8.91 | 0.0124 |
| week | 1 | -0.00029075 | 0.05836 | -0.00 | 0.9965 |



---------- pid=2 ----------

The REG Procedure
Model: MODEL1
Dependent Variable: rna

Analysis of Variance

| Source | DF | Sum of Squares | Mean Square | F Value | Pr > F |
|---|---|---|---|---|---|
| Model | 1 | 0.06870 | 0.06870 | 0.96 | 0.3986 |
| Error | 3 | 0.21385 | 0.07128 | | |
| Corrected Total | 4 | 0.28254 | | | |

| | | | |
|---|---|---|---|
| Root MSE | 0.26699 | R-Square | 0.2431 |
| Dependent Mean | 5.21301 | Adj R-Sq | -0.0091 |
| Coeff Var | 5.12156 | | |

Parameter Estimates

| Variable | DF | Parameter Estimate | Standard Error | t Value | Pr > \|t\| |
|---|---|---|---|---|---|
| Intercept | 1 | 5.34963 | 0.18337 | 29.17 | <.0001 |
| week | 1 | -0.01423 | 0.01450 | -0.98 | 0.3986 |





---------- pid=3 ----------

The REG Procedure
Model: MODEL1
Dependent Variable: rna

Analysis of Variance

| Source | DF | Sum of Squares | Mean Square | F Value | Pr > F |
|---|---|---|---|---|---|
| Model | 1 | 0.49062 | 0.49062 | 2.53 | 0.2097 |
| Error | 3 | 0.58104 | 0.19368 | | |
| Corrected Total | 4 | 1.07166 | | | |

| | | | |
|---|---|---|---|
| Root MSE | 0.44009 | R-Square | 0.4578 |
| Dependent Mean | 5.26184 | Adj R-Sq | 0.2771 |
| Coeff Var | 8.36384 | | |

Parameter Estimates

| Variable | DF | Parameter Estimate | Standard Error | t Value | Pr > \|t\| |
|---|---|---|---|---|---|
| Intercept | 1 | 4.89673 | 0.30226 | 16.20 | 0.0005 |
| week | 1 | 0.03803 | 0.02390 | 1.59 | 0.2097 |





---------- pid=4 ----------

The REG Procedure
Model: MODEL1
Dependent Variable: rna

Analysis of Variance

| Source | DF | Sum of Squares | Mean Square | F Value | Pr > F |
|---|---|---|---|---|---|
| Model | 1 | 1.35792 | 1.35792 | 2.44 | 0.2584 |
| Error | 2 | 1.11145 | 0.55573 | | |
| Corrected Total | 3 | 2.46937 | | | |

| | | | |
|---|---|---|---|
| Root MSE | 0.74547 | R-Square | 0.5499 |
| Dependent Mean | 3.36234 | Adj R-Sq | 0.3249 |
| Coeff Var | 22.17118 | | |

Parameter Estimates

| Variable | DF | Parameter Estimate | Standard Error | t Value | Pr > \|t\| |
|---|---|---|---|---|---|
| Intercept | 1 | 4.14404 | 0.62370 | 6.64 | 0.0219 |
| week | 1 | -0.13028 | 0.08335 | -1.56 | 0.2584 |



```
proc mixed data = actg320oneper ;
model rna = trt strat week  / s ;
random int week / subject = pid s g ;
where pid < 25 ;
run ;
```

The Mixed Procedure

Model Information

| | |
|---|---|
| Data Set | WORK.ACTG320ONEPER |
| Dependent Variable | rna |
| Covariance Structure | Variance Components |
| Subject Effect | pid |
| Estimation Method | REML |
| Residual Variance Method | Profile |
| Fixed Effects SE Method | Model-Based |
| Degrees of Freedom Method | Containment |

Dimensions

| | |
|---|---|
| Covariance Parameters | 3 |
| Columns in X | 4 |
| Columns in Z Per Subject | 2 |
| Subjects | 24 |
| Max Obs Per Subject | 5 |
| Observations Used | 103 |
| Observations Not Used | 17 |
| Total Observations | 120 |

Iteration History

| Iteration | Evaluations | -2 Res Log Like | Criterion |
|---|---|---|---|
| 0 | 1 | 318.66503481 | |
| 1 | 2 | 299.27229641 | 0.00024588 |
| 2 | 1 | 299.25720786 | 0.00000135 |
| 3 | 1 | 299.25712850 | 0.00000000 |



Convergence criteria met.

Estimated G Matrix

| Row | Effect | Subject | Col1 | Col2 |
|---|---|---|---|---|
| 1 | Intercept | 1 | 0.2391 | |
| 2 | week | 1 | | 0.003092 |

Covariance Parameter Estimates

| Cov Parm | Subject | Estimate |
|---|---|---|
| Intercept | pid | 0.2391 |
| week | pid | 0.003092 |
| Residual | | 0.6756 |



The Mixed Procedure

Fit Statistics

| | |
|---|---|
| -2 Res Log Likelihood | 299.3 |
| AIC (smaller is better) | 305.3 |
| AICC (smaller is better) | 305.5 |
| BIC (smaller is better) | 308.8 |



Solution for Fixed Effects

| Effect | Estimate | Standard Error | DF | t Value | Pr > \|t\| |
|---|---|---|---|---|---|
| Intercept | 5.1153 | 0.2444 | 21 | 20.93 | <.0001 |
| trt | -0.1875 | 0.2885 | 55 | -0.65 | 0.5186 |
| strat | -1.0555 | 0.2900 | 55 | -3.64 | 0.0006 |
| week | -0.06458 | 0.01740 | 23 | -3.71 | 0.0012 |

Solution for Random Effects

| Effect | Subject | Estimate | Std Err Pred | DF | t Value | Pr > \|t\| |
|---|---|---|---|---|---|---|
| Intercept | 1 | 0.05587 | 0.3731 | 55 | 0.15 | 0.8815 |
| week | 1 | 0.02040 | 0.04434 | 55 | 0.46 | 0.6472 |
| Intercept | 2 | 0.2085 | 0.3672 | 55 | 0.57 | 0.5725 |
| week | 2 | 0.04076 | 0.03281 | 55 | 1.24 | 0.2194 |
| Intercept | 3 | 0.08236 | 0.3672 | 55 | 0.22 | 0.8234 |
| week | 3 | 0.06641 | 0.03281 | 55 | 2.02 | 0.0478 |
| Intercept | 4 | -0.6154 | 0.3682 | 55 | -1.67 | 0.1003 |
| week | 4 | -0.05256 | 0.04420 | 55 | -1.19 | 0.2395 |
| Intercept | 5 | -0.4863 | 0.3722 | 55 | -1.31 | 0.1968 |
| week | 5 | 0.05163 | 0.03295 | 55 | 1.57 | 0.1229 |
| Intercept | 6 | -0.02646 | 0.3748 | 55 | -0.07 | 0.9440 |
| week | 6 | -0.02388 | 0.05000 | 55 | -0.48 | 0.6348 |
| Intercept | 7 | 0.1118 | 0.3707 | 55 | 0.30 | 0.7640 |
| week | 7 | -0.02876 | 0.04426 | 55 | -0.65 | 0.5185 |
| Intercept | 8 | -0.03262 | 0.3672 | 55 | -0.09 | 0.9295 |
| week | 8 | -0.04778 | 0.03281 | 55 | -1.46 | 0.1510 |
| Intercept | 9 | -0.1672 | 0.3701 | 55 | -0.45 | 0.6532 |
| week | 9 | 0.001733 | 0.04424 | 55 | 0.04 | 0.9689 |
| Intercept | 10 | -0.1715 | 0.3698 | 55 | -0.46 | 0.6447 |
| week | 10 | -0.01662 | 0.03285 | 55 | -0.51 | 0.6150 |
| Intercept | 11 | 0.1805 | 0.3707 | 55 | 0.49 | 0.6283 |
| week | 11 | -0.04794 | 0.04426 | 55 | -1.08 | 0.2835 |
| Intercept | 12 | -0.3679 | 0.3693 | 55 | -1.00 | 0.3235 |
| week | 12 | -0.04137 | 0.03284 | 55 | -1.26 | 0.2130 |
| Intercept | 13 | -0.2126 | 0.3731 | 55 | -0.57 | 0.5712 |
| week | 13 | -0.00422 | 0.04434 | 55 | -0.10 | 0.9246 |
| Intercept | 14 | 0.6663 | 0.3922 | 55 | 1.70 | 0.0950 |
| week | 14 | 0.01846 | 0.05412 | 55 | 0.34 | 0.7344 |
| Intercept | 15 | 0.2220 | 0.3865 | 55 | 0.57 | 0.5681 |
| week | 15 | -0.00891 | 0.03298 | 55 | -0.27 | 0.7880 |
| Intercept | 16 | 0.1836 | 0.3672 | 55 | 0.50 | 0.6191 |





The Mixed Procedure

Solution for Random Effects

| Effect | Subject | Estimate | Std Err Pred | DF | t Value | Pr > \|t\| |
|---|---|---|---|---|---|---|
| week | 16 | -0.00960 | 0.03281 | 55 | -0.29 | 0.7710 |
| Intercept | 17 | 0.2460 | 0.3693 | 55 | 0.67 | 0.5081 |
| week | 17 | 0.06613 | 0.03284 | 55 | 2.01 | 0.0489 |
| Intercept | 18 | -0.1078 | 0.3693 | 55 | -0.29 | 0.7715 |
| week | 18 | 0.02780 | 0.03284 | 55 | 0.85 | 0.4008 |
| Intercept | 19 | 0.1005 | 0.3722 | 55 | 0.27 | 0.7881 |
| week | 19 | 0.01553 | 0.03295 | 55 | 0.47 | 0.6393 |
| Intercept | 20 | 0.02655 | 0.3725 | 55 | 0.07 | 0.9434 |
| week | 20 | 0.007548 | 0.04997 | 55 | 0.15 | 0.8805 |

| | | | | | | |
|---|---|---|---|---|---|---|
| Intercept | 21 | 0.7732 | 0.3701 | 55 | 2.09 | 0.0414 |
| week | 21 | 0.07491 | 0.04424 | 55 | 1.69 | 0.0961 |
| Intercept | 22 | -0.4401 | 0.3748 | 55 | -1.17 | 0.2453 |
| week | 22 | -0.02374 | 0.05000 | 55 | -0.47 | 0.6368 |
| Intercept | 23 | -0.2525 | 0.3693 | 55 | -0.68 | 0.4970 |
| week | 23 | -0.05515 | 0.03284 | 55 | -1.68 | 0.0987 |
| Intercept | 24 | 0.02324 | 0.3672 | 55 | 0.06 | 0.9498 |
| week | 24 | -0.03077 | 0.03281 | 55 | -0.94 | 0.3524 |

Type 3 Tests of Fixed Effects

| Effect | Num DF | Den DF | F Value | Pr > F |
|---|---|---|---|---|
| trt | 1 | 55 | 0.42 | 0.5186 |
| strat | 1 | 55 | 13.25 | 0.0006 |
| week | 1 | 23 | 13.77 | 0.0012 |