SAS MACROS FOR BOOTSTRAP SAMPLES WITH STRATIFICATION AND MULTIPLE OBSERVATIONS PER SUBJECT

Terry Hyslop

Biostatistics Section
Division of Clinical Pharmacology
Thomas Jefferson University
125 South 9th Street
Sheridan Building, Suite 403
Philadelphia, PA 19107

ABSTRACT

The development of powerful computers for the desktop has allowed the bootstrap to become a popular inferential method in statistics. A bootstrap sample is a random sample selected with replacement. While SAS macros for selecting bootstrap samples have been presented, the issues of multiple, and varying numbers of records per subject have not been addressed. This data structure would arise, for example, in the analysis of mixed-effects models, such as for data collected from crossover clinical trials. When applying mixed models to clinical trials data there is also the concern of maintaining the structure, such as strata or sequence, hence the bootstrap samples should be stratified. We present two SAS macros: The first selects bootstrap samples from a data file with multiple (including unequal) observations per subject. The second macro is an extension of the first macro in that it selects stratified bootstrap samples with varying numbers of records per subject, varying number of records per strata and varying number of strata. We have attempted to increase the efficiency of the sampling with replacement process in SAS by isolating the sampling variable and rejoining the original data to the bootstrap samples with Proc SQL.

INTRODUCTION

Assume we have sample data, denoted \((X_1, \ldots, X_n)\), where \(X\) is either scalar or vector. An analysis of this data has been completed, with the outcome being some statistics of interest, denoted \(\Theta_i\), \(i = 1\) to \(s\), where \(\Theta = \Theta_i(X_1, \ldots, X_n)\). The statistics may be simple to very complex depending on the situation. What we are interested in providing, through bootstrapping, is some idea of the distributions of the \(\Theta_i\).

The distribution of \(\Theta_i\) can be determined by 1) selecting a random sample, \(B_j = (X_1^*, \ldots, X_n^*)\), a sample of size \(n\) selected with replacement from the original data, 2) analyzing the bootstrap sample in the same way as the original sample data to obtain \(\Theta_{ij} = \Theta_i(B_j)\), 3) repeat steps 1 and 2 many times to obtain the distribution of each \(\Theta_i\).

See Efron and Tibshirani\(^1\) for a complete discussion of bootstrap methods.

This paper presents two macros for selecting bootstrap samples from data with multiple (varying) observations per subject and for selecting stratified bootstrap samples from data of this same structure.

METHODS

Bootstrap methods require the selection of a large number (200 to 1000) of random samples with replacement from the original data collected for a study. For simple data structures, this is achieved via the random number function, \(\text{ranuni}\), which can be used to calculate a random pointer for a set statement, so that observations are randomly selected, eg:
do j=1 to 1000 ;
data random ;
do i=1 to n ;
   iobs = int(ranuni(0)*n) + 1 ;
   set mydata point=iobs nobs=n ;
   output ;
end ;
end ;

For more complex data structures, such as multiple (and varying) observations per subject, the problem is no longer to randomly select observations with replacement, as in the above example, but rather to randomly select subjects (clusters) with replacement. The %BOOTMULT macro was developed for this purpose and is presented in Appendix 1.

The random selection process is only the first step for bootstrapping. Once the samples are selected, each sample then must be analyzed in the same way as the original data, providing a distribution of estimates of the statistics of interest. In our case, we were analyzing data from crossover clinical trials with PROC MIXED. The trials were from multi-sequence, multi-period designs. It was necessary to maintain the sequence structure of the original data set in the bootstrap samples, so the %BOOTMULT macro was modified to select stratified samples with replacement. This modified macro, %BTSTRAT, is presented in Appendix 2 and described below.

%BOOTMULT macro

The %BOOTMULT macro was written to randomly select subjects (clusters) with replacement rather than observations. It should be noted that subjects is used here as a general term and that any variable can be used for the selection process; that is, any variable which identifies a cluster of data which must be maintained in the selected samples. This is an important distinction, because this macro can then be used as a starting point for bootstrapping clustered data of any form.

The first section of the macro sorts the input data set by the subject identifier, reads through the input data set and keeps only unique subject (cluster) identifiers in the idvars temporary data set. The user controls the subject identifier by setting the idvar in the macro call.

%macro bootmult(data= , sample=500, idvar=) ;
proc sort data=&data ;
   by &idvar ;
run;
data idvars(keep=&idvar)
   set &data ;
   by &idvar ;
   if first.&idvar then do;
      output idvars ;
   end ;
run ;

The process is then completed by merging the randomly selected subjects with their original study data via PROC SQL.

proc sql ;
create table mylib.bootdata as
   select b.j, b.newid, d.*
   from bootsubj as b inner join
      &data as d
   on b.&idvar=d.&idvar
   order by 1,2 ;
quit;
Random samples with replacement are then taken from each strata. (Note the use of && in the do loop - this allows the macro processor to resolve to the variable name and then to the value within the variable so that the number of times the loop is performed can be varied for each strata.)

```
data bootsubj;
do j=1 to &sample;
   newid=0;
   %do k=1 %to &numstrat;
      %do i=1 %to &n&k;
         iobs = int(ranuni(0)*&n&k) + 1;
         newid+1;
         set idvars&k point=iobs;
         keep j &idvar &stratvar newid;
      %end;
   %end;
%end;
end;
stop;
run;
```

As described for the %BOOTMULT macro, the randomly selected subjects are then merged with their multiple records via Proc SQL.

```
proc sql;
   create table mylib.bootdata as
   select b.j, b.newid, d.*
   from bootsubj as b inner join &data
   as d
   on b.&idvar=d.&idvar
   order by 1,2;
quit;
```

Example 2.

Using the data set from Example 1, the %BTSTRAT macro call to obtain 500 samples from the 2 sequence study would be:

```
%btstrat(data=mydata, sample=500, idvar=subj, stratvar=seq,numstrat=2);
```

This call would create the temporary data sets idvars, idvars1, and idvars2, bootsubj as well as the permanent data set mylib.bootdata; an example listing of one sample is given in Appendix 4. Each bootstrap sample is then ready for analysis, with the end result being a distribution of the statistics of interest.

**DISCUSSION**

This paper has presented a method for bootstrap sample selection from clustered data. The macros given here can be used as a basis for inferential analysis of clustered data, including error estimation, hypothesis testing, and bootstrap confidence intervals.

Each of the random samples selected maintain all the variables from the original data, with added subject (cluster) and sample identifiers. Additional macros to provide the required analysis for each sample would further simplify the bootstrap process. Usage of utility procedures, such as PROC APPEND and the recently developed Output Delivery System (ODS), greatly facilitate the process of storing output statistics from each sample run.

**ACKNOWLEDGEMENT**

This work was supported in part by Contract #223-94-3001 between the Food and Drug Administration and Jefferson Medical College.

**REFERENCES**


Example 1.

The sample data set is from a 2-sequence, 4-period, 8 subject crossover clinical trial of two formulations. A portion of the SAS data set is listed below:

<table>
<thead>
<tr>
<th>OBS</th>
<th>SUBJ</th>
<th>SEQ</th>
<th>PER</th>
<th>TRT</th>
<th>ln_AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>T</td>
<td>5.10534</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>R</td>
<td>4.92217</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>T</td>
<td>5.09006</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>R</td>
<td>4.70863</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>T</td>
<td>5.59434</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>R</td>
<td>5.11620</td>
</tr>
<tr>
<td>7</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>T</td>
<td>5.45916</td>
</tr>
<tr>
<td>8</td>
<td>2</td>
<td>1</td>
<td>4</td>
<td>R</td>
<td>5.34425</td>
</tr>
<tr>
<td>9</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>R</td>
<td>5.12656</td>
</tr>
<tr>
<td>10</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>T</td>
<td>4.99179</td>
</tr>
<tr>
<td>11</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>R</td>
<td>4.51305</td>
</tr>
<tr>
<td>12</td>
<td>3</td>
<td>2</td>
<td>4</td>
<td>T</td>
<td>4.69318</td>
</tr>
</tbody>
</table>

The call to the macro to obtain 500 samples from this data set is (assuming the data set containing the example data is named mydata):

```sas
%bootmult (data=mydata, sample=500, idvar= subj) ;
```

This call to %BOOTMULT will create the temporary data sets idvars, bootsubj, and the permanent data set mylib.bootdata, which contains 500 bootstrap samples from mydata. An example listing of one sample from mylib.bootdata is provided in Appendix 3. The temporary data set idvars is listed below:

<table>
<thead>
<tr>
<th>OBS</th>
<th>SUBJ</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
</tr>
</tbody>
</table>

%BTSTRAT Macro

This macro was written as an extension of the %BOOTMULT macro in that it allows for multiple (and varying) number of records per subject, but it also maintains an underlying data structure in the bootstrap samples for analysis purposes. Its purpose is to select stratified random samples with replacement. If the input data set is from a 4-sequence 4-period clinical trial, then each bootstrap sample will represent a 4-sequence 4-period trial with the same number of subjects per sequence as in the original study. Currently, the macro requires that the strata variable be numeric and that it be sequentially numbered.

The first section of the macro sorts the input file, reads through the input data set, keeping only the unique subject and strata identifier variables.

```sas
%macro btstrat(data=, sample=500, idvar=, stratvar=, numstrat=4) ;
proc sort data=&data by &idvar ;
run;
```

The identifier variable data set is then split into separate data sets for each strata level. In the example of a 4-sequence clinical trial, 4 data sets with unique subject identifiers would be created. Each data set would contain the identifier of the subjects in a sequence from the trial. The utility macro %numobs is used to count the number of observations in each of the stratified data sets.

```sas
%do k=1 %to &numstrat ;
  data idvars&k ;
  set idvars(where=(&stratvar= &k)) ;
  %global nidvar ;
  %let nidvar=&num
```


Appendix 1. %BOOTMULT macro

%macro bootmult(data=&sysdsn, sample=100, idvar=) ;

/* Sort the input data set by the id variable, &idvar */
/* Create a temporary SAS data set containing the unique id variables */
data idvars(keep=&idvar) ;
set &data ;
by &idvar ;
if first.&idvar then do;
output idvars idcnt+1;
end;
run;

/* Randomly select n subjects with replacement &sample times */
data bootsubj ;
%do j=l to &sample ;
  newid=0 ;
  %do i=l to n ;
    iobs = int(ranuni(0)*n) + 1 ;
  %end ;
  set idvars point=iobs nobs=n ;
  keep j &idvar newid ;
  output ;
%end ;
run ;

/* Combine selected subjects with study data, sort by sample number then new id within sample */
proc sql ;
create table mylib.bootdata as
select b.j, b.idobs, b.newid, d.* from bootsubj as b inner join &data as d
on b.&idvar=d.&idvar order by 1,3 ;
quit;

Appendix 2. %BTSTRAT macro

%macro btstrat(data=, sample=500, idvar=, stratvar=, numstrat=);  
proc sort data=&data;  
  by &idvar;  
run;  
*****************************************************************************  
/* Create multiple data sets, which */  
/* contain one record per id */  
/* variable for the bootstrap */  
/* samples, one file per strata */  
/* for stratified sampling */  
*****************************************************************************  
data idvars(keep=&idvar &stratvar);  
set &data;  
  by &idvar;  
  if first.&idvar then do;  
  output idvars  
    idcnt.+1;  
  end;  
run;  
%do k=l %to &numstrat;  
data idvars&k;  
  set idvars(where=&stratvar=&k);  
run;  
%end;  
quit;  
*****************************************************************************  
/* Randomly select n&k subjects from */  
/* each strata with replacement */  
/* &sample times */  
*****************************************************************************  
data bootsubj;  
do j=1 to &sample;  
  newid=0;  
  %do k=1 %to &numstrat;  
    %do i=1 %to &n&k;  
      iobs = int(ranuni(0)*&n&k) + 1;  
      newid+1;  
    end;  
    set idvars&k point=iobs;  
    keep j &idvar &stratvar newid;  
    output;  
  %end;  
%end;  
end;  
stop;  
run;  
proc sql;  
  create table mylib.bootdata as  
  select b.j, b.newid, d.*  
  from bootsubj as b inner join  
    &data as d  
  on b.&idvar=d.&idvar  
  order by 1,2;
Appendix 3.

Example 1 sample output:
*BOOTMULT macro

One bootstrap sample from clinical trial data set described in Example 1.

<table>
<thead>
<tr>
<th>J</th>
<th>NEWID</th>
<th>SUBJ</th>
<th>SEQ</th>
<th>PER</th>
<th>TRT</th>
<th>LN_AUCT</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>T</td>
<td>5.45916</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>R</td>
<td>5.11520</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>T</td>
<td>5.59434</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>7</td>
<td>2</td>
<td>2</td>
<td>T</td>
<td>5.12871</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>7</td>
<td>2</td>
<td>3</td>
<td>R</td>
<td>5.24650</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>7</td>
<td>2</td>
<td>4</td>
<td>T</td>
<td>5.48852</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>R</td>
<td>5.01926</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>T</td>
<td>5.11520</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>T</td>
<td>5.45916</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>6</td>
<td>2</td>
<td>4</td>
<td>T</td>
<td>5.33320</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>6</td>
<td>2</td>
<td>2</td>
<td>T</td>
<td>5.08140</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>6</td>
<td>2</td>
<td>1</td>
<td>R</td>
<td>5.27146</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>6</td>
<td>2</td>
<td>3</td>
<td>R</td>
<td>5.03500</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>6</td>
<td>2</td>
<td>4</td>
<td>T</td>
<td>5.33320</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>6</td>
<td>2</td>
<td>2</td>
<td>T</td>
<td>5.08140</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>6</td>
<td>2</td>
<td>1</td>
<td>R</td>
<td>5.27146</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>6</td>
<td>2</td>
<td>3</td>
<td>R</td>
<td>5.03500</td>
</tr>
<tr>
<td>2</td>
<td>6</td>
<td>3</td>
<td>2</td>
<td>4</td>
<td>T</td>
<td>4.59318</td>
</tr>
<tr>
<td>2</td>
<td>6</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>R</td>
<td>4.51305</td>
</tr>
<tr>
<td>2</td>
<td>6</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>T</td>
<td>4.99179</td>
</tr>
<tr>
<td>2</td>
<td>6</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>R</td>
<td>5.21656</td>
</tr>
<tr>
<td>2</td>
<td>7</td>
<td>2</td>
<td>1</td>
<td>4</td>
<td>R</td>
<td>5.34425</td>
</tr>
<tr>
<td>2</td>
<td>7</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>T</td>
<td>5.45916</td>
</tr>
<tr>
<td>2</td>
<td>7</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>R</td>
<td>5.11520</td>
</tr>
<tr>
<td>2</td>
<td>7</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>T</td>
<td>5.45943</td>
</tr>
<tr>
<td>2</td>
<td>8</td>
<td>6</td>
<td>2</td>
<td>4</td>
<td>T</td>
<td>5.33320</td>
</tr>
<tr>
<td>2</td>
<td>8</td>
<td>6</td>
<td>2</td>
<td>3</td>
<td>R</td>
<td>5.03500</td>
</tr>
<tr>
<td>2</td>
<td>8</td>
<td>6</td>
<td>2</td>
<td>2</td>
<td>T</td>
<td>5.08140</td>
</tr>
<tr>
<td>2</td>
<td>8</td>
<td>6</td>
<td>2</td>
<td>1</td>
<td>R</td>
<td>5.27146</td>
</tr>
</tbody>
</table>

Note: This bootstrap sample contains 3 subjects from sequence 1 and 5 subjects from sequence 2. Other bootstrap samples would contain varying numbers of subjects from the two sequences.
Appendix 4.

Example 2 sample output:
%BTSTRAT macro

One stratified bootstrap sample from clinical trial data set described in Example 2.

<table>
<thead>
<tr>
<th>J</th>
<th>NEWID</th>
<th>SUBJ</th>
<th>SEQ</th>
<th>PER</th>
<th>TRT</th>
<th>LN_AUCT</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>T</td>
<td>5.10534</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>R</td>
<td>4.70863</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>T</td>
<td>5.09006</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>R</td>
<td>4.92217</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>T</td>
<td>4.99179</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>R</td>
<td>5.21656</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>R</td>
<td>4.51305</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>4</td>
<td>T</td>
<td>4.69318</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>R</td>
<td>5.21656</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>R</td>
<td>4.51305</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>4</td>
<td>T</td>
<td>4.69318</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>T</td>
<td>4.99179</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>4</td>
<td>R</td>
<td>5.34425</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>T</td>
<td>5.59434</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>R</td>
<td>5.11620</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>T</td>
<td>5.45916</td>
</tr>
<tr>
<td>1</td>
<td>5</td>
<td>6</td>
<td>2</td>
<td>4</td>
<td>T</td>
<td>5.33320</td>
</tr>
<tr>
<td>1</td>
<td>5</td>
<td>6</td>
<td>2</td>
<td>3</td>
<td>R</td>
<td>5.03500</td>
</tr>
<tr>
<td>1</td>
<td>5</td>
<td>6</td>
<td>2</td>
<td>2</td>
<td>T</td>
<td>5.08140</td>
</tr>
<tr>
<td>1</td>
<td>5</td>
<td>6</td>
<td>2</td>
<td>1</td>
<td>R</td>
<td>5.27146</td>
</tr>
<tr>
<td>1</td>
<td>6</td>
<td>5</td>
<td>2</td>
<td>4</td>
<td>T</td>
<td>5.21330</td>
</tr>
<tr>
<td>1</td>
<td>6</td>
<td>5</td>
<td>2</td>
<td>1</td>
<td>R</td>
<td>5.15618</td>
</tr>
<tr>
<td>1</td>
<td>6</td>
<td>5</td>
<td>2</td>
<td>3</td>
<td>R</td>
<td>4.98703</td>
</tr>
<tr>
<td>1</td>
<td>6</td>
<td>5</td>
<td>2</td>
<td>2</td>
<td>T</td>
<td>5.16878</td>
</tr>
<tr>
<td>1</td>
<td>7</td>
<td>8</td>
<td>1</td>
<td>1</td>
<td>T</td>
<td>4.13196</td>
</tr>
<tr>
<td>1</td>
<td>7</td>
<td>8</td>
<td>1</td>
<td>4</td>
<td>R</td>
<td>5.24597</td>
</tr>
<tr>
<td>1</td>
<td>7</td>
<td>8</td>
<td>1</td>
<td>2</td>
<td>R</td>
<td>5.24913</td>
</tr>
<tr>
<td>1</td>
<td>7</td>
<td>8</td>
<td>1</td>
<td>3</td>
<td>T</td>
<td>4.84968</td>
</tr>
<tr>
<td>1</td>
<td>8</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>R</td>
<td>4.92217</td>
</tr>
<tr>
<td>1</td>
<td>8</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>R</td>
<td>4.70863</td>
</tr>
<tr>
<td>1</td>
<td>8</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>T</td>
<td>5.10534</td>
</tr>
<tr>
<td>1</td>
<td>8</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>T</td>
<td>5.09006</td>
</tr>
</tbody>
</table>

Note: This data set is stratified by sequence, and contains 4 subjects per sequence as in the original data set. All bootstrap samples for this data set using %BTSTRAT will contain 4 subjects per sequence.