

SAS Program Notes

Biostatistics: A Guide to Design, Analysis, and Discovery

Chapter 6: Study Design

Note 6.1 – Drawing random samples

In the SAS commands below, we demonstrate how the SAS procedure **PROC SURVEYSELECT** can be used to draw a simple random sample (SRS) without replacement of 25 from 50 sequentially numbered subjects as discussed in **Example 6.13** in the textbook. First we use a **DO** statement to create a data set named OLD containing the numbers 1 to 50. Next we use **PROC SURVEYSELECT** and the DATA option to select the input dataset which has been named OLD in this case. The OUT option tells SAS where to store the output dataset which has been named NEW in this case. The METHOD option allows you to specify the method of sample selection. Some options are **SRS**-which is selection with equal probability and without replacement, **SYS**-which is systematic random sampling, and **URS**-which is selection with equal probability and with replacement. The option **SAMPSIZE** allows you to specify the sample size and the option **SEED** indicates the random number seed.

SAS commands:

```
DATA OLD;
DO I = 1 TO 50;
OUTPUT;
END;
PROC SURVEYSELECT DATA=OLD OUT=NEW METHOD=SRS SAMPSIZE=25 SEED=1;
RUN;

PROC PRINT DATA=NEW;
RUN;
```

SAS output:

The SAS System

Obs	I
1	2
2	3
3	4
4	5
5	6
6	10
7	16
8	17
9	18
10	21

11	23
12	25
13	26
14	27
15	29
16	30
17	31
18	33
19	36
20	37
21	39
22	42
23	44
24	48
25	50

SAS can be used to generate random numbers. The command `RANUNI` is used to generate random numbers from a uniform distribution ranging from 0 to 1, and the command `RANNOR` is used to generate random numbers from a standard normal distribution. To use the command, we need to specify a *seed*. The *seed*, a numerical value, is placed in parentheses after the command. If a different numerical value or *seed* is specified, the commands will produce a different set of random numbers. For example, the value produced by `RANUNI(1)` is different from the value produced by `RANUNI(2)` as shown using the SAS commands below.

SAS commands:

```
DATA RANDOM_UNIFORM;
X = RANUNI (1) ;
Y = RANUNI (2) ;
PROC PRINT;
RUN;
```

SAS output:

The SAS System		
Obs	X	Y
1	0.18496	0.97009

Note that we could have used a negative number (ie. `RANUNI(-1)`) which uses SAS's internal clock as the seed thus generating a different random number each time you run the command. These commands are most beneficial when they allow you to generate a set of random values. For example, to generate ten random values from a uniform distribution ranging from 0 to 1, we can use the following SAS commands.

SAS commands:

```
DATA RANDOM_UNIFORM;  
DO I = 1 TO 10;  
X = RANUNI(1);  
OUTPUT;  
END;  
PROC PRINT;  
RUN;
```

SAS output:

The SAS System

Obs	I	X
1	1	0.18496
2	2	0.97009
3	3	0.39982
4	4	0.25940
5	5	0.92160
6	6	0.96928
7	7	0.54298
8	8	0.53169
9	9	0.04979
10	10	0.06657