

## Drawing a Random Sample with SAS<sup>1</sup>

Sometimes it is necessary or useful to select a random sample from your data. Sometimes a specific number of cases is required, and sometimes rough percent is needed. The following SAS programs will show how to select either type.

```
* Random_SAMPLE.SAS ;
TITLE1 'Random_SAMPLE.SAS - not exact Number of Cases' ;

** Create a data set with 100,000 cases for this example ;
data Total ;
* create 100000 cases ;
do i = 1 to 100000 ;
  id = i ;
output ;
end ;

** Use data set created above to demonstrate creating a random sample ;

data Sample1 ; set Total ;
* CREATE A RANDOM VARIABLE ;
* This variable will have values ranging from 0 to 100 ;
* If I do not need an exact N, I can select based on values of the variable named "rand" ;
* A positive number within parentheses sets the "seed"
* so that I will get the same sample each time ;
* If seed a negative integer, the time of day is used to initialize the seed stream. ;
* and the sample will vary each time you run the program ;
rand = ranuni(692) * 100 ;
* To get roughly a 20% sample I could issue the following command ;
if rand ge 0 and rand le 20 ;
proc means ; run ;

TITLE1 'Random_SAMPLE.SAS - EXACT Number of Cases' ;
data total2 ; set Total ;
* CREATE A RANDOM VARIABLE ;
* This variable will have values ranging from 0 to 100 ;
* If I do not need an exact N, I can select based on values of the variable named "rand" ;
* A positive number within parentheses sets the "seed"
* so that I will get the same sample each time ;
* If seed a negative integer, the time of day is used to initialize the seed stream. ;
* and the sample will vary each time you run the program ;
rand = ranuni(937) * 100 ;
* To get an exact N of 20,000 cases, for example, I would sort by ;
* the variable "rand", create a sequential variable with values
* from 1 to N, then select only the cases with values of 20,000 ;
* or less on the sequential variable ;
proc sort ; by rand ;
data sample2 ; set total2 ;
sequent + 1 ;
if sequent le 20000 ;
proc means ; run ;
```

---

<sup>1</sup>Prepared by Patty Glynn, University of Washington, March 15, 2006. C:\all\help\helpnew\RandomSampleSAS.wpd.