

Stata Program Notes

Biostatistics: A Guide to Design, Analysis, and Discovery

Chapter 6: Study Design

Note 6.1 – Drawing random samples

Stata can be used to draw random samples. Using the command **uniform()** to generate random numbers from a uniform distribution, a random sample with replacement of twenty observations from values ranging from 1-100 can be taken using the following commands:

Stata commands:

```
set obs 20
set seed 2
gen id = _n
gen i = round(uniform()*100)
list
```

Stata output:

	id	i
1.	1	85
2.	2	5
3.	3	63
4.	4	70
5.	5	52
6.	6	34
7.	7	17
8.	8	91
9.	9	39
10.	10	41
11.	11	96
12.	12	15
13.	13	65
14.	14	43
15.	15	42
16.	16	73
17.	17	43
18.	18	90
19.	19	9
20.	20	21

The Stata command **sample** can be used to draw random samples without replacement from data that is already in memory. For example, assume we have 50 observations numbered 1 to 50 and would like to draw a random sample of 10 observations without replacement.

Stata commands:

```
set obs 50
gen id = _n
set seed 2
sample 10, count
sort id
list
```

Stata output:

```
      +-----+
      | id |
      +-----+
1.   |  2 |
2.   |  7 |
3.   | 12 |
4.   | 19 |
5.   | 20 |
      +-----+
6.   | 32 |
7.   | 36 |
8.   | 40 |
9.   | 45 |
10.  | 48 |
      +-----+
```

The Stata command **bsample** can be used to draw random samples with replacement from data that is already in memory. For example, assume we have 50 observations numbered 1 to 50 and would like to draw a random sample of 10 observations with replacement.

Stata commands:

```
set obs 50
gen id = _n
set seed 2
bsample 10
sort id
list
```

Stata output:

```
      +-----+
      | id |
      +-----+
1.   |  3 |
2.   |  5 |
3.   |  5 |
4.   |  7 |
5.   | 10 |
      +-----+
6.   | 13 |
7.   | 13 |
8.   | 15 |
9.   | 15 |
10.  | 24 |
      +-----+
```