**Random sample**



How do you take the sample?

There are a number of ways to take a sample. In fact, we have a whole course on it! We will focus primarily on one way to take a sample that we saw earlier named a random sample.

Random sample – A sample from the population where every item has an equal probability of being included within the sample.

The easiest way to take a random sample

Number all items from your population with 1, 2, …, N. Use the sample() function in R to take a random sample of size n.

Example: Random sample (random\_sample.R)

Suppose the population is very small with N = 20 and the sample size is n = 10. Items are numbered 1 to 20.

> N <- 20

> n <- 10

> set.seed(9812)

> sample(x = 1:N, size = n)

[1] 17 14 19 10 13 3 9 16 5 7

What if I used a different seed number?

Suppose the population has a size of N = 25,000 and the sample has a size of n = 100.

> N <- 25000

> n <- 100

> set.seed(7823)

> sample(x = 1:N, size = n)

 [1] 19628 13653 5074 19885 21474 20136 12240 17954 15127

 [10] 22689 18885 3330 1935 9221 12763 16808 20535 847

 [19] 20588 979 2521 24099 10006 8414 9066 19460 1499

 [28] 23788 13400 23016 1453 16761 14654 19271 19559 15535

 [37] 802 991 5570 1241 10906 3436 17237 3372 3681

 [46] 3860 1350 17968 931 6032 2478 23082 1614 1595

 [55] 14010 8183 13473 23012 1287 19611 16513 16467 6044

 [64] 15029 24735 18273 3517 19328 18802 7715 14776 15591

 [73] 18550 4553 12582 16415 15253 15408 13201 8137 13825

 [82] 1491 4882 15282 20068 95 13846 15991 7351 3617

 [91] 10551 9348 3830 629 15957 11921 23424 5792 5470

[100] 17549

Should you use the same seed number every time?

How does one choose a seed number?

One way is to use simulate as a sample from a uniform probability distribution.

> runif(n = 5, min = 0, max = 1)

[1] 0.8529989 0.1941843 0.5008234 0.6343356 0.5932707

Notice a seed number is not set here with runif(). The resulting seed number to use for the random sample is 8529989.