

R TUTORIAL, #8: SAMPLING WITH AND WITHOUT REPLACEMENT

The (>) symbol indicates something that you will type in.

A bullet (•) indicates what the R program should output (and other comments).

SAMPLING WITH REPLACEMENT

• I have a closet full of red, green, and blue shirts. I will pick seven shirts for the seven days of next week. How many ways can I pick colors for the seven days? The answer is $3^7 = 2187$ ways.

> Type: `colors = c("red", "green", "blue")`

> Type: `sample(colors, size=7, replace=T)`

- Seven colors are selected with replacement.
- Ordinarily, ‘replace=F’ is assumed (it is the “default”), so we must type ‘replace=T’ here.
- Note that color(s) are repeated.

> (Repeat the above command a few times. You should get a variety of results.)

- Note that color(s) are repeated.

SAMPLING WITHOUT REPLACEMENT

• Five students are in a classroom. The teacher must choose three of them to clean up. Since students can't be repeated here, the teacher must sample without replacement.

• The number of ways to do this is ${}_5C_3 = \binom{5}{3} = 10$ ways.

> Type: `students = c("Butters", "Cartman", "Kenny", "Kyle", "Stan")`

> Type: `sample(students, size=3)`

- Three students are selected without replacement.
- Note that none of the students are repeated.

> (Repeat the above command a few times. You should get a variety of results.)

- Note that none of the students are repeated at any time.

IMPORTING THE 'CARDS' DATA FILE FROM THE WEB; POKER

- We will import the text file “RCards.txt” from the web directory <http://www.kkuniyuk.com/RFiles>.

> Type: `cards = read.table(url("http://www.kkuniyuk.com/RFiles/RCards.txt"), header=T)`

- ‘cards’ will be a data frame containing the imported data.
- ‘header=T’ indicates that it is “True” that the file has a header (‘card’).

> Type: `cards`

- You will see the data frame ‘cards’.

> Type: `attach(cards)`

- We attach the data frame.

> Type: `sample(card, size=5)`

- Five cards are sampled without replacement, as in some poker games.
- Note that ‘card’ is the name of the variable, not ‘cards’.

> (Repeat the above command a few times. You should get a variety of results.)

- Note that none of the cards are repeated at any time.

> Type: `detach(cards)`

- We detach the data frame.