

Homework 1 R Exercise

R Markdown

This is a sampling exercise from a known probability distribution. We will first sample a single draw from the M&M problem in the text (3.19), then sample a draw of 100. Graphs will help us interpret results.

RStudio

We start out by setting up a character vector of colors, labelled **MnM**, that represents our sample space, the set of all possible outcomes, as well as a numeric vector, **Prob_MnM** that represents the probability of each color in **MnM**. We sum **Prob_MnM** just to check that it adds to 1. Why?

```
MnM=c("Brown","Yellow","Red","Blue","Orange","Green")
Prob_MnM=c(0.13,0.14,0.13,0.24,0.20,0.16)
sum(Prob_MnM)
```

To sample a single outcome from a known probability distribution, we can use the **sample** command. We need to specify the sample space, the number of draws we will make, whether we are sampling with replacement (the default is that we *are not* sampling with replacement), and the probability of each outcome. Note that we do not need the last argument if each outcome is equally likely. There are two versions of the command here; one that specifies the arguments and one that uses the default ordering of the arguments. Note that they both do the same thing, though they may generate different outcomes in this case. Report your results—there is no “wrong” answer here—why?

```
sample(x=MnM,size=1,replace=TRUE,prob=Prob_MnM)
sample(MnM,1,TRUE,Prob_MnM)
```

We can save results by assigning the output to a variable; in this case, we are saving a single outcome to the variable **MnM_Outcome**. When we simply enter the name of the variable in R’s console, the console will print its contents.

```
MnM_Outcome=sample(MnM,1,TRUE,Prob_MnM)
MnM_Outcome
```

Using the last **sample** command from above, change the **size** argument from 1 to 100 and save it to **MnM_Outcome**. Rather than look at a long vector of color labels, we can construct a table of frequencies, as well as a bar chart. The first command below is just to make sure our vector has length 100; if it does not, we have made a mistake.

```
length(MnM_Outcome)
table(MnM_Outcome)
library(lattice)
barchart(MnM_Outcome)
```

Are the frequencies in the table roughly what you would expect based on the probability model (do not expect an *exact* match)? It would be nice if the bars in the barchart matched their respective colors, but we do not need that particular bell-and-whistle at this point.