

# Math 151 - Probability Theory - Homework 3

your name here

Due: Friday, September 11, midnight PDT

## Important Note:

You should work to turn in assignments that are clear, communicative, and concise. Part of what you need to do is not print pages and pages of output. Additionally, you should remove these exact sentences and the information about HW scoring below.

Click on the *Knit to PDF* icon at the top of R Studio to run the R code and create a PDF document simultaneously. [PDF will only work if either (1) you are using R on the network, or (2) you have LaTeX installed on your computer. Lightweight LaTeX installation here: <https://yihui.name/tinytex/>]

Either use the college's RStudio server (<https://rstudio.pomona.edu/>) or install R and R Studio on to your personal computer. See: <https://research.pomona.edu/johardin/math151f20/> for resources.

## Assignment

### 1: PodQ

Describe one thing you learned from someone in your pod this week (it could be: content, logistical help, background material, R information, etc.) 1-3 sentences.

### 2: 1.7.8

An elevator building starts with five passengers and stops at seven floors. If every passenger is equally likely to get off at each floor and all the passengers leave independently of each other, what is the probability that no two passengers will get off at the same floor?

### 3: 1.7.10

A box contains 100 balls, of which  $r$  are red. Suppose that the balls are drawn from the box one at a time, at random, without replacement. Determine:

- the probability that the first ball drawn will be red;
- the probability that the 50th ball drawn will be red; and
- the probability that the last ball drawn will be red.

### 4: 1.8.4

A box contains 24 light bulbs, of which four are defective. If a person selects four bulbs from the box at random, without replacements, what is the probability that all four bulbs will be defective?

### 5: 1.8.6

Suppose that  $n$  people are seated in a random manner in a row of  $n$  theater seats. What is the probability that two particular people  $A$  and  $B$  will be seated next to each other?

### 6: 1.10.7

Suppose that a school band contains 10 students from the freshman class, 20 students from the sophomore class, 30 students from the junior class, and 40 students from the senior class. If 15 students are selected at random from the band, what is the probability that at least one student will be selected from each of the four classes? *Hint:* First determine the probability that at least one of the four classes will not be represented in the selection.

### 7: R - Baseball cards

Assume that every time you buy a box of Wheaties, you receive one of the pictures of the  $n$  players on the LA Dodgers. Over a period of time, you buy  $m \geq n$  boxes of Wheaties. [From Introduction to Probability by Grinstead and Snell]

- a. Use Theorem 1.10.2 to show that the probability that you get all  $n$  pictures is

$$1 - \binom{n}{1} \left(\frac{n-1}{n}\right)^m + \binom{n}{2} \left(\frac{n-2}{n}\right)^m - \dots + (-1)^{n-1} \binom{n}{n-1} \left(\frac{1}{n}\right)^m.$$

*Hint:* consider the event that you do not get the  $k^{\text{th}}$  player's picture.

- b. Write a computer program to ESTIMATE this probability. Use the program to estimate, for given  $n$ , the smallest value of  $m$  which will give probability greater than 0.5 of getting all  $n$  pictures. Consider  $n = 10, 50, 100$  and show that  $m = n \log(n) + n \log(2)$  is a good (if slightly conservative) estimate for the number of boxes needed. (For a derivation of this estimate, see Feller(1968).)

```
n.Dodgers=10
n.boxes = 20
n.reps=5
mycereal = sample(1:n.Dodgers, n.boxes,replace=TRUE)
mycereal          # remove this line after you know what it does

## [1] 8 6 6 1 9 3 10 6 3 5 9 2 6 10 2 8 1 6 2 9
unique(mycereal)  # remove this line after you know what it does

## [1] 8 6 1 9 3 10 5 2
length(unique(mycereal)) # remove this line after you know what it does

## [1] 8
mycerprop = c()
for (m in n.Dodgers:n.boxes){
  mycercount = c()
  for (j in 1:n.reps){
    mycereal = sample(1:n.Dodgers,m, replace=TRUE)
    mycercount = c(mycercount, length(unique(mycereal)))
  }
  mycerprop = c(mycerprop, sum(mycercount==n.Dodgers)/n.reps)
}

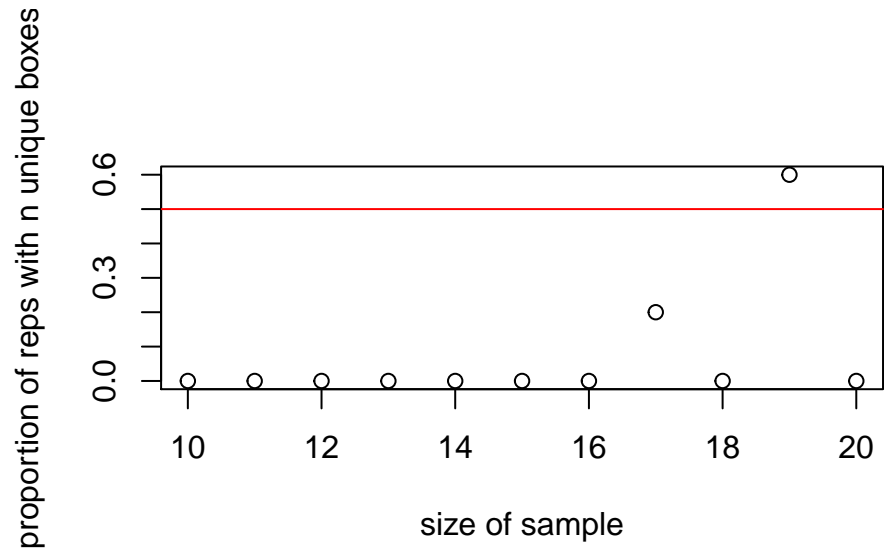
mycercount # remove this line later changes, note it gets written over every time m changes

## [1] 7 9 8 9 9
mycerprop  # remove this line later, note it should be as long as m

## [1] 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.2 0.0 0.6 0.0
```

After you have the simulation running, bump up the reps to get a good empirical estimate of the probability. A plot will give you an understanding of what is being estimated:

```
plot(n.Dodgers:n.boxes, mycerprop, xlab="size of sample", ylab="proportion of reps with n unique boxes",  
     abline(h=0.5, col="red")  
     abline(v=n.Dodgers*log(n.Dodgers) + n.Dodgers*log(2), col="blue"))
```



Write a few sentences reflecting on what you observe. Note that your sentences should come outside of the R chunk, but they should be within the R Markdown file and therefore typed (not pencil) on the pdf which you print and turn in. [n.b. This is the whole point of R Markdown: to create a way for reproducible code + analysis in one document.]