Homework 1

Tyler Poelking

1/16/2018

Question 1

"a. 5 out of the 26 letters in the alphabet are vowels. So there is a 5/26 = 0.1923 = 19.23% chance the letter is a vowel. "  
"b. 21 of the 26 letters in the alphabet are consonants. So there is a 21/26 = 0.8077 = 80.77% chance the letter is a consonant."  
"c. 12/26 = 0.4615 = 46.15% chance the letter occurs in the last 12 positions of the alphabet"  
"d. b,c,d,f, g, and h are the 6 consonants in between ‘a’ and ‘i’ in the conventional ordering, so the probability of picking any of these is 6/26 = 0.2307 = 23.07%."  
  
  
  
Question 2  
  
The size of the sample space of the experiment with replacement is n^r, where n=the number of things to choose from = 26 and r = the number of things we choose = 2. So n^r = 26^2 = 676. This includes combinations of two of the same letter such as (A,A) as well as combinations that are reverses of one another such as (A,B), (B,A).  
  
To find the probability that the first letter drawn precedes the second letter in the conventional ordering of the alphabet, we can consider a smaller example of a 4 letter alphabet [A, B, C, D]. In this scenario, say we pick A. There is a ¼ chance in picking A. From there, there is a ¾ chance we pick a letter that follows A (either B,C, or D). We can multiply these two quantities to get the chances we pick A and then get a letter such that A precedes it, obtaining (1/4)\*(3/4). From there, we can do the same for picking B first and get the quantities (1/4)\*(2/4). We then add these two numbers together because both are equate to the probabilities of outcomes of interest. This pattern continues for each letter, getting an end quantity of (1/4)\*(3/4)+ (1/4)\*(2/4)+ (1/4)\*(1/4)+ (1/4)\*(0/4). This equation can be simplified to (1/4^2)\*(sum of digits from 1-3) = (1/n^2)\*(sum of digits from 1 to (n-1)), where n is the number of letters in the alphabet. The sum of integers from 1 to (n-1) can be calculated by ((n-1)\*(n))/2. Extrapolating this to n=26, we get (1/676) \* ((25\*26)/2) = .4808 = 48.08%.  
  
The size of the sample space of the experiment without replacement is equal to 26 \* 25, since there are 26 letters to choose from on the first pick and 25 on the second. This equates to 650 possible outcomes. They are all the outcomes in the first sample space minus the 26 outcomes where the same letter was picked twice. This makes since because 676-26=650.   
  
To find the probability that the first letter drawn precedes the second letter in the conventional ordering of the alphabet, we can again consider our smaller example of [A, B, C, D]. We follow a very similar process, except now the second quantity in the multiplication has a denominator of 3 instead of 2. So For the instance where we pick A first, we get (1/4)\*(3/3) instead of (1/4)\*(3/4). The end equations looks like (1/(n)\*(n-1))\*(sum of digits from 1 to (n-1)). Using this, we get the probability equal to (1/650) \* ((25\*26)/2) = .5000 = 50.00%.

Question 3

1. There are 10 integers (0-9). The probability of choosing any one number (including an 8) one time is 1/10 = 0.1. The probability of drawing an 8 and then an 8 again is (1/10) \* (1/10) = 0.01. Finally, the probability of drawing an 8 twice then an 8 again is (1/10) \* (1/10) \* (1/10) = 0.001.

set.seed(42)  
  
#sample 3 million times  
values <- sample(0:9,3000000,TRUE)  
  
#create matrix that simulated drawing 3 numbers 1 million times  
samples <- matrix(values, nrow = 1000000, byrow = TRUE)  
  
#get total count each 3 number draw occured in simulation  
x1 <- data.frame(table(apply(samples, 1, paste, collapse = ", ")))  
  
#transform count to % frequency  
x1[2] <- x1[2]/1000000  
  
#display Note: 8,8,8 found on page 89  
head(x1, n=20)

## Var1 Freq  
## 1 0, 0, 0 0.000986  
## 2 0, 0, 1 0.001000  
## 3 0, 0, 2 0.000975  
## 4 0, 0, 3 0.001069  
## 5 0, 0, 4 0.000993  
## 6 0, 0, 5 0.001025  
## 7 0, 0, 6 0.001063  
## 8 0, 0, 7 0.001007  
## 9 0, 0, 8 0.001020  
## 10 0, 0, 9 0.000981  
## 11 0, 1, 0 0.000988  
## 12 0, 1, 1 0.000989  
## 13 0, 1, 2 0.000985  
## 14 0, 1, 3 0.001028  
## 15 0, 1, 4 0.001007  
## 16 0, 1, 5 0.000939  
## 17 0, 1, 6 0.000960  
## 18 0, 1, 7 0.000958  
## 19 0, 1, 8 0.000970  
## 20 0, 1, 9 0.001038

x1[889,]

## Var1 Freq  
## 889 8, 8, 8 0.001006

First we pick 3 million integers randomly sampled between 0-9, which a uniform distribution. We then reshape this vector into a 1000000, 3 matrix, simulating our scenario 1 million times, where each row corresponds to and instance of drawing three integers. In this table, the row corresponding to choosing (8,8,8) has a frequency of 0.001006. All other frequencies were around this value, demonstrating that getting three 8s is as likely as getting any other set of thee numbers.

Question 4

One of the largest factors of winning games is the players that make up the team. I would start by asking him which players he plans on having during the 2019 season. I would also ask if any of these players are injured or recently have been so we can take their potential absence from games into account. From there, we could compare our players to the players of other teams, using data such as that on fantasy league and ESPN websites. Another question I would ask is what our play schedule looks like so in our statistical analysis we can take into account which games are away and which games are played soon after other games (since we may not play as well if we are tired).