Name: Drashti Mehta

Lab Assignment 3

(40 points)

DIRECTIONS

- ***An uploaded pdf is due at the start of next Friday at 1PM.***

A. Solve the following exercises. Find the following Binomial probabilities by <u>hand and then with R</u>. For full credit, be sure to show all of your work, commands, and output. (30 pts total)

1. In the toss of 5 dice, what is the probability of obtaining 2 sixes? (5 pts)

By Hand: _____

P(X=k) = (^nCk) * p^k * q^(n-k) => [n! / k!*(n-k)!] * p^k * q^(n-k)
= [5! / 2!*(3)!] * (1/6)^2 * (5/6)^(3) = 0.16. Here, k = 2, n = 5, q = 1 – p
= 5/6 or 0.8333; p = probability of success = 1/6 or 0.16667

R Command: _____

> dbinom(x=2, size = 5, prob = 0.16667) = 0.1607555

2. In couples where each person is heterozygous for the sickle-cell gene, there is a probability of 0.25 that any child will have the disease, and a probability of 0.75 that any child will not have the disease. For a population of families of 4 children in which both parents are heterozygotes, what is the probability that in a randomly selected family, two or more children will have the disease? What are the R commands to create a barplot in R for P(X=k) where k is 0 to 4? Provide the R commands to color the bars and provide axis labels. (10 pts)

By Hand: _____

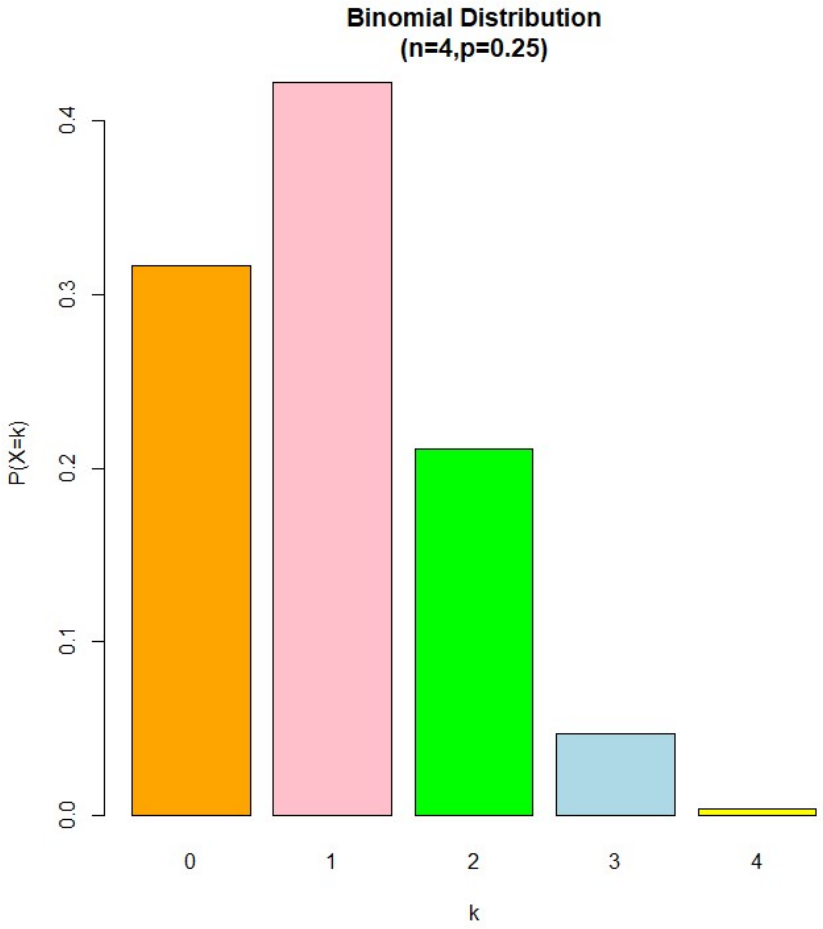
n = 4, p = 0.25, q = 0.75, k = 2

P(X<2) = ((4 choose 2)x (0.25)^2 x(0.75)^2) + ((4 choose 3)x (0.25)^2 x(0.75)^2) + ((4 choose 4)x (0.25)^2 x(0.75)^2).
Ans: 1 - Probability (x < 2)

R Command: _____

```
> n=4
```

```
> p=0.25
> colors = c("orange","pink","green","light blue","yellow")
> barplot(dbinom(0:n,size=n,prob=p),names.arg=0:n,ylab="P(X=k)",xlab="k",
  col=colors, main = "Binomial Distribution\n(n=4,p=0.25)")
```



3. Suppose 10% of the population has asthma. If a random sample of 20 people are drawn, find the probability that

a) **none will have asthma (5 pts)**

b) **one or more will have asthma (5pts)**

By Hand: _____

A) Here considering k = 0 in the first part and the probability of event k is 10% or 0.1. Also the total number of events are 20 Putting into formula

> P(X=k) = (^nCk) * p^k * q^(n-k)

> [n! / k!*(n-k)!] * p^k * q^(n-k)

> [20! / 0!*(20)!] * (0.1)^0 * (0.9)^(20) = 0.121

B) Here we subtract the probability of Px or P(X=k) from 1 to get atleast one patient with asthma. Hence 1 - Px = 0.878

R Command: _____

Ans a): Px <- dbinom(x=0, size = 20, prob = 0.1) => Ans: 0.1215767
Ans b): 1 - Px = 0.87843

4) We know the proportion in the population that has a disease is 3.8%. Let's randomly select 12 people who are independent of one another. We observe whether they have the disease. What is the probability that 3 will have the disease? (5 points).

By Hand: _____

Let us take X=3, n=12, p=0.038, q= 0.962. Putting into formula
> P(X=k) = (^nCk) * p^k * q^(n-k)
> [n! / k!*(n-k)!] * p^k * q^(n-k)
= [12! / 3!*(9)!] * (0.038)^3 * (0.962)^(9) = 0.00851

R Command: _____

Px = dbinom(x=3, size = 12, prob = 0.038) = 0.008518225

B. Solve this problem using R commands ONLY. (5 pts total)
1) DNA is made of these 4 nucleotides: A,G,C,T. A triplet contains three of these nucleotides. How many different triplets are possible if you can use any of the 4 nucleotides in each of the positions in the triplet (2.5 points)? How many if you can only use one of the 4 nucleotides once in a triplet (2.5 points)?

If repeats are allowed, method is n^n = 4^4 = 256. R code for this is base<-4; expo<-4; base^expo > Ans = 256.
If repeats are not allowed, calculation goes like n! = 4! = 24. R code for this is > factorial(4) >Ans = 24

C. The university policy department must write parking tickets to keep department revenues at budgeted levels. Suppose the mean number of tickets written per day was found to be 20.2. (5 pts)

What distribution can describe this random variable (2 pts)? Explain why using the properties of the distribution. Show the formula for the distribution (3 pts).

The distribution that describes this random variable is the poisson's distribution P(X=k) = (lambda^k) * (e^-lambda) / k! where k = 0,1,2.. λ = mean number of event occurrences in a

specified time interval or region in space. The events are independent. The events can not occur at the same time or the same place. The mean and variance are = lambda.