Statistics 251: Lab 3

Handout

First Name: _____Musa_____   Last Name: __Rasheed_____

Student Number: _____25618232_____

Students should write the name they are registered with. Do not put nicknames/short forms on the handout. Please write down your answers neatly and do show your work.

1. Finding probabilities:
   Let X = the time between questions. [Hint: An expression for a probability can look like one of the following: $P(a < X < b), P(a \le X \le b), P(a \le X < b),$
   $P(a < X), P(a \le X), P(X > b), P(X \ge b)$, etc.]

   a. The probability that the *time between questions* is **one minute or less**.
      i. Write an expression for the required probability. Draw a rough sketch of the probability density function of X by hand and shade the area of interest.
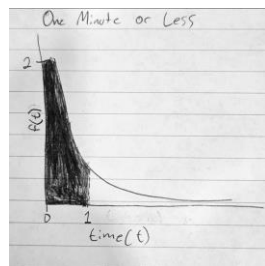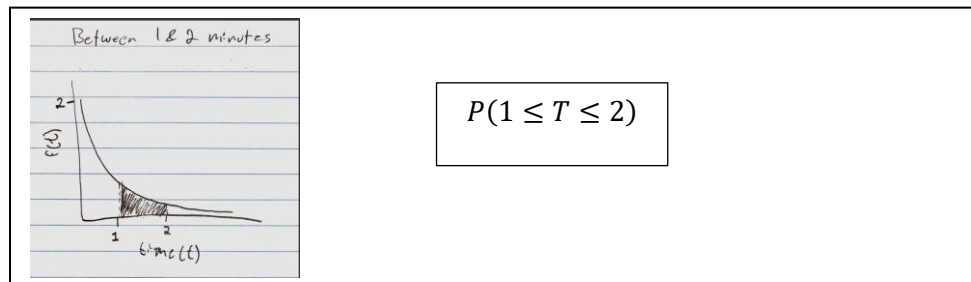

$$P(T \le 1)$$

where T is the random variable time.

      ii. Find the required probability using R. (Write the commands and result.)

```
pexp(1,2)

>> 0.864664716763387
```

b. The probability that the *time between questions* is **between one and two minutes**.

    i.   Write an expression for the required probability. Draw a rough sketch of the probability density function of X by hand and shade the area of interest.
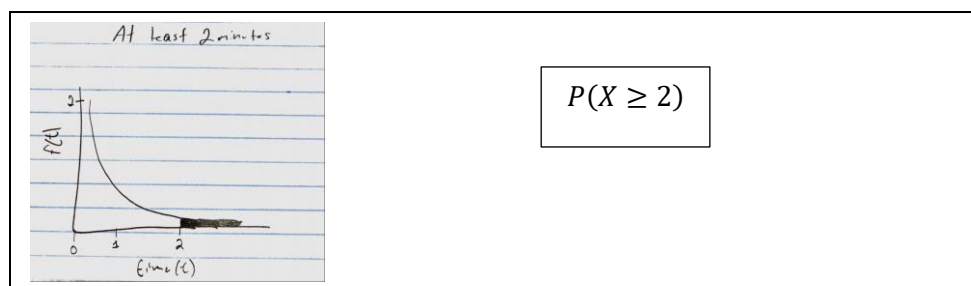

Between 1 & 2 minutes

$$P(1 \leq T \leq 2)$$

    ii.   Find the required probability using R. (Write the commands and result.)

```
pexp(2,2) - pexp(1,2)

>>0.117019644347878
```

c. The probability that the *time between questions* is **at least two minutes**.

    i.   Write an expression for the required probability. Draw a rough sketch of the probability density function of X by hand and shade the area of interest.

At least 2 minutes

$$P(X \geq 2)$$

    ii.   Find the required probability using R. (Write the commands and result.)

```
1 - pexp(2,2)

>>0.0183156388887342
```

2. Find the first quartile, median, and third quartile of X. Write your R commands and result.

```
qexp(0.25,2) # Q1

qexp(0.50,2) # Median

qexp(0.75,2) # Q3

>>0.14384103622589

>>0.346573590279973

>>0.693147180559945
```

3. Simulation:
    a. Generate a random sample of <u>six</u> *times between questions*. Draw a histogram representing the probability density of the sample. On top of the histogram, draw the pdf of an exponential distribution with rate = 2. Write your R commands.

```
x = rexp(6,2) #The Data

hist(P, prob=TRUE) #Creating the density histogram

curve(dexp(x,2),from=0,to=1,add=TRUE)#What the distribution would look like

>>
```


Probability Density Diagram
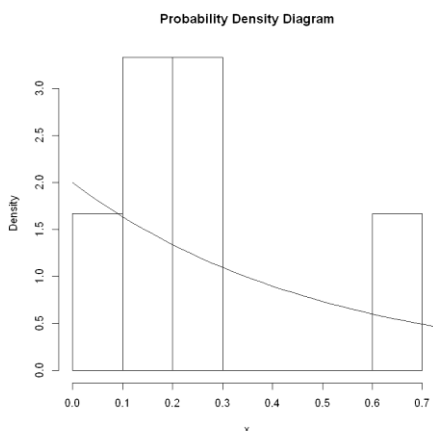
b. Generate a random sample of <u>6,000</u> *times between questions*. Draw a histogram representing the probability density of the sample. On top of the histogram, draw the pdf of an exponential distribution with rate = 2. Write your R commands.

```
x = rexp(6000,2)#The Data
hist(x, prob=TRUE, main = "Probability Density Diagram")#Creating the histogram
curve(dexp(x,2),from=0,to=5,add=TRUE)#What the distribution would look like
```
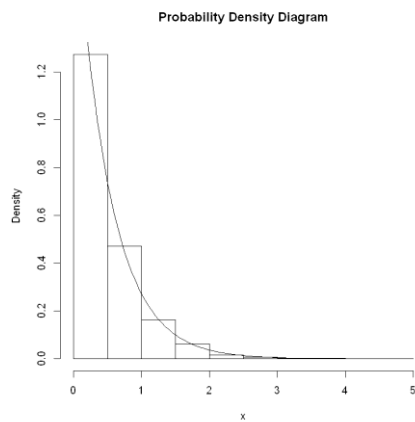

c. Explain which sample appears more representative of its true distribution.

The one with 6000 observations. Generally, the more observations you have, the more representative the distribution will be. We can see that it also follows the exponential curve shape much more closely than the sample with only 6 observations. Larger samples are also less susceptible to being skewed by outliers, thus making them more accurate.