Problem set 5

S520 Fall 2021

Answer questions 1–6. Upload your answers through the Assignments tab on Canvas by 11:59 pm, Thursday 30th September.

Trosset question numbers refer to the published textbook. Show working (answers only will not get full credit.) You may work with others, but you must write up your homework independently — you should not have whole sentences in common with other students or other sources. You may (and sometimes have to) use R; include your R code where relevant.

1. (Adapted from the Spring 2017 takehome.) The tempo (speed) of a piece of music is usually measured in beats per minute (BPM.) A study[1] found that the BPM in disco songs was approximately normal with mean 120 and standard deviation 20. (Treat BPM as a continuous random variable for the purpose of this question.) The best tempo for dancing is considered to be 115 to 135 BPM, while anything above 160 BPM is exhausting.

   Suppose a DJ gets lazy and puts her MP3 player containing a large, representative collection of disco songs on shuffle.

   (a) Suppose the MP3 player randomly selects a disco song. What is the probability that the BPM is between 115 and 135?

   (b) Suppose the MP3 player randomly selects ten disco songs. What is the probability at least one of the ten songs has a BPM over 160?

   (c) Suppose the MP3 player randomly selects two disco songs. What is the probability the *average* BPM of these two songs is over 160? (Hint: What's the distribution of the sum of two independent normal random variables? How big does the sum have to be for the average to be over 160?)

2. Trosset chapter 5.6 exercise 8. Reminder: Normal$(1, 9)$ means the *variance* is 9, not the standard deviation.

3. Let X be a random variable representing the length of time (in years) that an external hard drive lasts. From past experience, I find X is well-modeled by an Exponential distribution with "rate" $\lambda = 0.1$.[2] That means the CDF of X is:

$$F(y) = \begin{cases} 0 & y < 0 \\ 1 - e^{-0.1y} & y \geq 0 \end{cases}$$

---

[1] https://github.com/nikhilunni/BPMFinder/

[2] This is not quite true—a decent hard drive has a very low annual failure rate for about the first three years and a higher failure rate after that—but let's pretend it is.

and the PDF is

$$f(x) = \begin{cases} 0 & x < 0 \\ 0.1\,e^{-0.1x} & x \geq 0 \end{cases}$$

(a) What is the probability an external hard drive lasts less than five years?

(b) What is the probability an external hard drive lasts between four years and eight years?

(c) Find the expected value of X. (Hint: Using a formula is much easier than doing the calculus.)

(d) Suppose I am paranoid and buy *two* external hard drives. What is the probability that *both* hard drives last less than five years?

4. I arrive at my bus stop at a random time. Let X be a random variable representing the time (in minutes) I have to wait until the next bus. Suppose the CDF of X is

$$F(y) = \begin{cases} 0 & y < 0 \\ \frac{y}{30} & 0 \leq y < 20 \\ \frac{1}{3} + \frac{y}{60} & 20 \leq y < 40 \\ 1 & y \geq 40. \end{cases}$$

(a) What is the median time I have to wait for a bus?

(b) What is the interquartile range of X? (Note: This should be one number.)

(c) Is the distribution of X symmetric, skewed to the left, or skewed to the right? How do you know?

5. (From the Fall 2015 midterm.) Let X be a continuous random variable with cumulative distribution function

$$F(y) = \begin{cases} 0 & y < 1 \\ 1 - \frac{1}{y^2} & y \geq 1. \end{cases}$$

and probability density function

$$f(x) = \begin{cases} \frac{2}{x^3} & x \geq 1 \\ 0 & \text{otherwise.} \end{cases}$$

(a) Find the median of X.

(b) Find the interquartile range of X.

(c) Is the expected value of X less than, equal to, or greater than its median? (You are not required to calculate the expected value, but you can if you want to.)

6. Trosset chapter 6.4 exercise 1. (Note: The PDF is on Trosset page 137.)