“In the Name of God”

Engineering probabilities and Statistics **CE 40181**
Sharif University of Technology

Homework 3
Wednesday, Nov 29 2017

Reminders:

- **Typing the solutions has extra marks.**
- **Collaboration is permitted, but you must write solutions *by yourself without* assistance.**
- **Getting solutions from outside sources such as the Web or students not enrolled in the class is strictly forbidden.**
- **Late submissions will be treated according to the course policy.**

1. Suppose that a program’s execution time is normally distributed with mean 39.8 and standard deviation 2.05 seconds, respectively. What is the probability that of 20 random selections of the program’s execution, five take at least 40 seconds to finish?

2. Starting at 5A.M., every half hour there is a flight from Mehrabad Airport in Tehran to Shahid Dastgheib Airport in Shiraz. Suppose that none of these planes is completely sold out and that they always have room for passengers. A person who wants to fly to Shiraz arrives at Mehrabad Airport at a random time between 8:45 A.M. and 9:45 A.M. Find the probability that she waits
   (a) At most 10 minutes.
   (b) At least 15 minutes.
   (c) And also compute the expected value of her waiting time.

3. The time (in hours) required to repair a machine is an exponential distributed random variable with parameter $\lambda = \frac{1}{2}$. What is
   (a) the probability that a repair time exceeds 2 hours
   (b) Calculate the conditional probability that an ongoing repair takes at least 10 hours, given that so far it has taken 9 hours.
   (c) Calculate the expected value and the variance of the time required to repair the machine.

4. Let the random variable N(t) be the number of packets arriving during time (0, t]. Suppose N(t) is Poisson with pmf

$$p_N(n) = \frac{(\lambda t)^n}{n!} e^{-\lambda t} \qquad \text{for n = 0, 1, 2, } \ldots$$

Let the random variable Y be the time to get the n-th packet. Find the pdf of Y.

5. A CRT[1] monitor in our laboratory is supposed to show each pixel in its exact position as a colorful bright dot but unfortunately it has some error and the result is a bit skew. Consider the random variables X and Y which show the horizontal and vertical displacement of bright dot from its correct position, respectively. Suppose that the joint pdf of X and Y is as follow:

$$f_{X,Y}(x,y) = \begin{cases} c & \text{if } |x| + |y| \leq \frac{1}{\sqrt{2}} \\ 0 & \text{otherwise} \end{cases}$$

where c is a constant.

   (a) Find c.
   (b) Find $f_X(x)$ and $f_{X|Y}(x|y)$.
   (c) Are X and Y independent random variables? Justify your answer.

6. Random variables X and Y correspond to temperature of two elements of a computer system. The joint probability density function of X and Y is given by

$$f(x,y) = \begin{cases} \lambda x y^2 & 0 \leq x \leq y \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

   (a) Determine the value of λ.
   (b) Find the marginal probability density functions of X and Y .
   (c) Calculate E(X) and E(Y).

7. For $n \geq 1$, let X_n be a continuous random variable with the probability density function

$$f_n(x) = \begin{cases} \frac{c_n}{x^{n+1}} & \text{if } x > c_n \\ 0 & \text{otherwise} \end{cases}$$

X_ns are called Pareto random variables and are used to study income distributions [2].

   (a) Calculate $c_n, n \geq 1$

[1] Cathode ray tube

[2] Pareto (Italian civil engineer, economist, and sociologist) originally used this distribution to describe the allocation of wealth among individuals since it seemed to show rather well the way that a larger portion of the wealth of any society is owned by a smaller percentage of the people in that society. He also used it to describe distribution of income. This idea is sometimes expressed more simply as the Pareto principle or the "80-20 rule" which says that 20% of the population controls 80% of the wealth. However, the 80-20 rule corresponds to a particular value of n, and in fact, Pareto's data on British income taxes in his Course "d'conomie politique" indicates that about 30% of the population had about 70% of the income. Which can be derived with a different value of n

(b) Find $E(X_n), n \geq 1$

(c) Determine the density function of $Z_n = \ln X_n, n \geq 1$

8. A hardware chip's task is to calculate $Y = 1 - 3X^2$ in which X and Y are indicating the input and the output, respectively. We have observed that the input of this chip is a continues random variable with following pdf:

$$f_X(x) = \begin{cases} 4x^3 & \text{if } 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

Considering the given information, find the pdf of Y, i.e. output of the chip.

9. A factory produces 5000 monitors, everyday. 2500 of the daily produced monitors are type I and 2500 are type II. If a sample of 40 monitors is selected at random to be examined for defects, what is the approximate probability that this sample contains at least 18 monitors of each type?

R Problem

Many measurements are *continuous*, e.g. time, length, scores, to name but a few. Many of these measurements are *random*, e.g. the time you must wait until the bus arrives, the score you get from a course, or the future height of your child brother or sister. Therefore, the concept of continuous random variables rises.

Poor Professor!

Suppose that you are a professor in MIT in your future life and you receive many emails from students all over the world. On average, you receive a qualified applicants email each 10 days. However, due to uncertainty in the length of this period, you are facing a random variable, which follows an exponential distribution. (simulations = 1000 use **rexp** function to simulate exponential distribution.)

1. Use **ggplot2** to visualize the distribution mentioned above.

2. Write a function in R which gets n as an input to simulate this process for n hypothetical years and returns the number of qualified applications received during a year.

3. Plot the outputs of the function mentioned in b for $n = 10$, $n = 100$, and $n = 1000$.

4. Write a function in R which gets n as an input and returns the time you must wait until you receive 10^{th} qualified application.

5. Use **ggplot2** to visualize the distribution in 4.