Day-2: Assignments

Probability Theory along with Linear Algebra and Optimization Theory are the branches of Mathematics that ML primarily use. The data generating processes under ML's radar are not deterministic: The data is corrupted due to several reasons:

1 Measurement errors may corrupt the data

2 Inputs and/or outputs might be partially observed

3 There is inherent randomness in data generating proces

Probability Theory is used to express the randomness, and carry out the necessary computations for estimation, hypothesis testing and prediction purposes.
There are many good books on Probability Theory. It is a very standard and established theory, and each book tells more or less the same stuff with possibly differing priorities and presentation orders. In this course, we shall reference the book "Introduction to Probability Models" by Sheldon Ross (Note: Another good book is by Grimmett&Stirzaker, but it is a bit more advanced. Keep it in your library though). The assignment for Day-2 is composed of two parts and is as follows:

1 Part-1

Read Chapter-1, Chapter-2, Chapter-3, and sections 11.1 through 11.5. Work out the examples presented by the author.

2 Part-2

Solve the following exercises:

1 Let E,F,G be three events. Find expressions for the events that of E,F,G

a) only F occurs,

b) both E and F but not G occur,

c) at least one event occurs,

d) at least two events occur,

e) all three events occur,

f) none occurs,

g) at most one occurs,

h) at most two occur.

2 If two fair dice are tossed, what is the probability that the sum is i, $i = 2, 3, \cdots, 12$?

3 Suppose that 5 percent of men and 0.25 percent of women are color-blind. A color-blind person is chosen at random. What is the probability of this person being male? Assume that there are an equal number of males and females?

4 If the occurrence of B makes A more likely, does the occurrence of A make B more likely?

5 What is the conditional probability that the first die is six given that the sum of the dice is seven?

6 In a class there are four freshman boys, six freshman girls, and six sophomore boys. How many sophomore girls must be present if sex and class are to be independent when a student is selected at random?

7 Suppose a die is rolled twice. What are the possible values that the following random variables can take on?

i The maximum value to appear in the two rolls.

ii The minimum value to appear in the two rolls.

iii The sum of the two rolls.

iv The value of the first roll minus the value of the second roll.

8 If the distribution function F of X is given by

$$F(b) = \begin{cases} 0 & b < 0 \\ \frac{1}{2} & 0 \leq b < 1 \\ \frac{3}{5} & 1 \leq b < 2 \\ \frac{4}{5} & 2 \leq b < 3 \\ \frac{9}{10} & 3 \leq b < 3.5 \\ 1 & \leq b \geq 3.5 \end{cases}$$

Calculate the probability mass function of X.

9 Suppose X has a binomial distribution with parameters 6 and 12. Show that $X = 3$ is the most likely outcome.

10 An airline knows that 5 percent of the people making reservations on a certain flight will not show up. Consequently, their policy is to sell 52 tickets for a flight that can hold only 50 passengers. What is the probability that there will be a seat available for every passenger who shows up?

11 A point is uniformly distributed within the disk of radius 1. That is, its density is

$$f(x,y) = C,\ 0 \le x^2 + y^2 \le 1$$

Find the probability that its distance from the origin is less than x, $0 \le x \le 1$

12 If the density function of X equals

$$f(x) = \begin{cases} ce^{-2x} & 0 < x < \infty \\ 0 & x < 0 \end{cases}$$

Find c. What is $P\{X > 2\}$?

13 If X is uniformly distributed over the unit interval, calculate $E[X^2]$

14 Let $X_1, X_2, ..., X_{10}$ be independent Poisson random variables with mean 1. Use the central limit theorem to approximate

$$P\{X_1 + X_2 + \cdots + X_{10} >= 15\}$$

15 An unbiased die is successively rolled. Let X and Y denote, respectively,the number of rolls necessary to obtain a six and a five. Find

a) $E[X]$,

b) $E[X|Y = 1]$

c) $E[X|Y = 5]$

16 Let X be uniform over $(0, 1)$. Find $E[X|X < 12]$

17 Let $X_1, X_2, \cdots, X_n$ be independent random variables having a common distribution function that is specified up to an unknown parameter θ. Let $T = T(X)$ be a function of the data $X = (X_1, X_2, \cdots, X_n)$. If the conditional distribution of $X_1, X_2, \cdots, X_n$ given $T(X)$ does not depend on θ then T (X) is said to be a sufficient statistic for θ . In the following cases, show that $T(X) = \sum_{i=1}^{n} X_n$ is a sufficient statistic for θ:

(a) The X_i are normal with mean θ and variance 1.

(b) The density of X_i is $f(x) = \theta e^{-\theta x}, x > 0$

(c) The mass function of X_i is $p(x) = \theta x(1 - \theta)^{1-x}, x = 0, 1, 0 < \theta < 1$

(d) The X_i are Poisson random variables with mean θ

18 A total of 11 people, including you, are invited to a party. The times at which people arrive at the party are independent uniform $(0, 1)$ random variables.

(a) Find the expected number of people who arrive before you

(b) Find the variance of the number of people who arrive before you

19 (a) Show that

$$Cov(X, Y) = Cov(X, E[Y|X])$$

(b) Suppose, that, for constants a and b,

$$E[Y|X] = a + bX$$

Show that

$$b = \frac{Cov(X, Y)}{Var(X)}$$

20 Suppose X is a Poisson random variable with mean λ. The parameter λ is itself a random variable whose distribution is exponential with mean 1. Show that

$$P\{X = n\} = (\frac{1}{2})^{n+1}$$

21 Law of Large Numbers states that sample mean of a random variable X converges to its population mean in the limit. Now, do the following computation

1 Calculate the sample mean for N $i.i.d$ samples

2 Repeat Step-1 for $N \in \{1000, 10000, 100000, 1000000\}$

3 Record these 4 numbers

Do the computations separately for $X \sim U(0, 1)$ and $X \sim N(0, 2)$. Are your findings compatible with the law?

22 Central Limit Theorem states that sample mean converges to a normal random variable in the limit. Now, let X be a random variable with population mean μ, and population variance σ^2. Consider the scaled sample mean

$$\bar{X}_s \equiv \frac{\bar{X} - \mu}{\sigma/N}$$

where $\bar{X} = \frac{\sum_{i=1}^{N} X_N}{N}$

1 Calculate $\bar{X}_s$ for $N = 1000$

2 Repeat Step-1 for 100 times

3 Plot these 100 numbers

Do the computations separately for $X \sim U(0, 1)$ and $Poisson(5)$. Are your findings compatible with the law?