
Tutorial Sheet 4: Sets and Probability

Course: CSEG 2036P | *School of Computing Sciences, UPES*

Faculty: Dr. Mrityunjay Guha Majumdar

1. Suppose a radioactive particle is contained in a unit square. We can define random variables X and Y to denote the x and y coordinates of the particle's location in the unit square, with the bottom left corner placed at the origin. Radioactive particles follow completely random behavior, meaning that the particle's location should be uniformly distributed over the unit square. This implies that the joint density function of X and Y should be constant over the unit square, which we can write as

$$f(x, y) = c, 0 \leq x \leq 1, 0 \leq y \leq 1 \text{ and } 0, \text{ elsewhere}$$

- (a) What is the value of c ? What can we say from this result regarding any random variables X and Y that are uniformly distributed on $[0, 1]$?
- (b) What is the value of $E(X)$ and $E(Y)$?
- (c) What are the marginal density functions $f_X(x)$ and $f_Y(y)$?
2. Let A , B , and C be sets. Prove or disprove the following identities using properties of sets:
- (a) Prove that $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.
- (b) Prove that $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.
- (c) Prove that $A \cup (B - C) = (A \cup B) - (A \cap C)$.
- (d) Prove that $A - (B \cap C) = (A - B) \cup (A - C)$.
3. Consider three sets A , B , and C defined as follows: Set A contains all prime numbers less than 10, Set B contains all even numbers less than 20, Set C contains all multiples of 3 less than 30. Now, define the set D as follows: Set D contains elements that are either in A or in the union of B and C , but not both. Express set D in terms of set operations. Find D and provide an explanation of how you arrived at your answer.
4. Let X be a random variable representing the time (in minutes) between consecutive arrivals at a bus stop, which follows an exponential distribution with a rate parameter $\lambda = 0.2$. Calculate the expectation of $Y = 2X + 3$. What does the expectation of Y represent in the context of bus arrival times?
5. If the i^{th} central moment around the point $x = 10$ is given by the expression

$$\mu'_i = i^{2/3} + 1$$

for a distribution, find the values of $E(X^2 + 5X + 7)$ and $\text{Var}(3X + 11)$.

6. If independent random variables X and Y have means $\mu_X = 5$ and $\mu_Y = 9$, variances $\sigma_X^2 = 24$ and $\sigma_Y^2 = 19$ respectively, what is the correlation coefficient between X and $X + Y$?

Student	Exam Score	Study Hours
1	85	10
2	92	12
3	78	8
4	95	14
5	88	11

- You have data on the exam scores (out of 100) and the number of hours students studied for an upcoming math exam. Calculate the correlation coefficient between exam scores and study hours for the following dataset: Interpret the correlation coefficient in terms of the relationship between study hours and exam scores.
- In a marketing study, you have data on monthly sales revenue (in lakhs of rupees) and monthly advertising spending (in lakhs of rupees) for a retail store over the past year. Calculate the correlation coefficient for the data given below. Interpret the correlation

Month	Sales Revenue	Advertising Spending
Jan	75	5
Feb	82	7
Mar	70	6
Apr	88	8
May	92	9
Jun	95	10
Jul	98	12
Aug	90	9
Sep	86	8
Oct	80	6
Nov	78	5
Dec	85	7

coefficient and discuss the relationship between advertising spending and sales revenue.

- You have data on daily maximum temperatures (in degrees Fahrenheit) and daily ice cream sales (in rupees) for a local ice cream shop over a summer month. Calculate the correlation coefficient for the data in Table 3. Interpret the correlation coefficient and explain the relationship between temperature and ice cream sales.
- Three coins are tossed. Let X denote the number of heads appearing and Y the tail runs (consecutive tails). Find the mean and standard deviation of X and Y .
- Consider a random variable X representing the number of minutes a customer spends browsing in a store, which follows a uniform distribution between 5 and 15 minutes. Calculate the variance of $Y = X^2 - 10X + 25$. What does the variance of Y tell you about the variability in customer browsing times?
- In the Indian Super League (ISL), there are two prominent goal scorers: Bartholomew Ogbeche of Hyderabad F.C. and Sunil Chhetri of Bengaluru F.C. Both players have

Day	Max Temperature	Ice Cream Sales
1	85	12000
2	88	13000
3	90	14000
4	92	15000
5	78	10000
6	80	11000
7	95	16000
8	97	17000
9	88	13000
10	86	12000

a reputation for being excellent scorers, but they have different playing styles. Player Ogbeche is known for being consistent; he scores in 80% of the matches he plays. Chhetri, on the other hand, is known for being explosive; he scores in only 30% of the matches he plays, but when he does score, he usually scores multiple goals. For Bartholomew Ogbeche: In matches where he scores, he scores exactly one goal 80% of the time, two goals 15% of the time, and three goals 5% of the time. For Sunil Chhetri: In matches where he scores, he scores one goal 10% of the time, two goals 60% of the time, and three goals 30% of the time. Now, a fan is attending a match involving one of these players, but they don't know which player it is. The fan watches the match and notices that no goals have been scored. What is the probability that Bartholomew Ogbeche is on the field, given this observation?