**ASSIGNMENTS**

**1. Define the Bayesian interpretation of probability.**

**2. Define probability of a union of two events with equation.**

**3. What is joint probability? What is its formula?**

**4. What is chain rule of probability?**

**5. What is conditional probability means? What is the formula of it?**

**6. What are continuous random variables?**

**7. What are Bernoulli distributions? What is the formula of it?**

**8. What is binomial distribution? What is the formula?**

**9. What is Poisson distribution? What is the formula?**

**10. Define covariance.**

**11. Define correlation**

**12. Define sampling with replacement. Give example.**

**13. What is sampling without replacement? Give example.**

**14. What is hypothesis? Give example.**

**SOLUTIONS**

1. ***The Bayesian interpretation of probability is a way of assigning probabilities to uncertain events based on prior knowledge or experience. It views probability as a measure of the degree of belief or confidence in the occurrence of an event, and updates this belief or confidence as new evidence becomes available.***
2. ***The probability of a union of two events A and B is given by the equation:***

***P(A or B) = P(A) + P(B) - P(A and B)***

1. ***Joint probability is the probability of two or more events occurring together. It is calculated by multiplying the probabilities of the individual events. The formula for joint probability is:***

***P(A and B) = P(A) \* P(B|A)***

***where P(B|A) is the probability of event B given that event A has occurred.***

1. ***The chain rule of probability is a method of calculating the probability of a series of events. It states that the probability of a series of events can be calculated by multiplying the conditional probabilities of each event, given the events that have already occurred. Mathematically, the chain rule can be written as:***

***P(A and B and C) = P(A) \* P(B|A) \* P(C|A and B)***

1. ***Conditional probability is the probability of an event A given that another event B has occurred. It is denoted as P(A|B) and is calculated using the formula:***

***P(A|B) = P(A and B) / P(B)***

1. ***Continuous random variables are variables that can take on any value within a certain range or interval. They are described by probability density functions (PDFs) instead of probability mass functions (PMFs).***
2. ***Bernoulli distributions are a type of discrete probability distribution that models the outcome of a single binary event with two possible outcomes, often denoted as success or failure. The formula for the Bernoulli distribution is:***

***P(X=k) = p^k \* (1-p)^(1-k)***

***where X is the random variable representing the outcome of the event, k=0 or 1, and p is the probability of success.***

1. ***The binomial distribution is a discrete probability distribution that describes the number of successes in a fixed number of independent trials, where each trial has the same probability of success. The formula for the binomial distribution is:***

***P(X=k) = (n choose k) \* p^k \* (1-p)^(n-k)***

***where P(X=k) is the probability of k successes in n trials, p is the probability of success on a single trial, (n choose k) is the binomial coefficient, which gives the number of ways to choose k items from a set of n items, and (1-p)^(n-k) is the probability of failure.***

1. ***The Poisson distribution is a discrete probability distribution that describes the number of events that occur in a fixed interval of time or space, given the average rate of occurrence. The formula for the Poisson distribution is:***

***P(X=k) = (lambda^k / k!) \* e^(-lambda)***

***where P(X=k) is the probability of k events occurring in the interval, lambda is the average rate of occurrence, k is the number of events, k! is the factorial of k, and e is the mathematical constant approximately equal to 2.718.***

1. ***Covariance is a measure of the degree to which two variables vary together. It measures the joint variability of two random variables, and is defined as:***

***cov(X,Y) = E[(X-E[X])(Y-E[Y])]***

***where cov(X,Y) is the covariance between X and Y, E[X] is the expected value of X, E[Y] is the expected value of Y, and the brackets denote expected value.***

1. ***Correlation refers to a statistical measure that measures the relationship between two variables. It is used to determine the degree to which two variables are related and how they move together. Correlation ranges from -1 to +1, where -1 indicates a perfect negative correlation (as one variable increases, the other decreases) and +1 indicates a perfect positive correlation (as one variable increases, the other also increases).***
2. ***Sampling with replacement is a statistical method in which a sample is taken from a population, and after each observation, the selected unit is returned to the population, and the next unit is selected randomly. This means that the same unit can be selected multiple times, and each time it is selected, it is considered an independent observation. For example, if a jar contains 100 marbles, and you randomly select 10 marbles with replacement, you could potentially select the same marble multiple times.***
3. ***Sampling without replacement is a statistical method in which a sample is taken from a population, and once an observation is selected, it is removed from the population, so it cannot be selected again. This means that each selected unit is considered a dependent observation, as the probability of selecting the next unit changes based on the previous selections. For example, if a deck of cards contains 52 cards, and you randomly select 5 cards without replacement, once you select a card, it is removed from the deck, and the probability of selecting the next card changes.***
4. ***A hypothesis is an educated guess or prediction about a phenomenon or relationship between variables that can be tested through research. It is a statement that explains the relationship between two or more variables, and it is formulated based on prior knowledge, observation, or theoretical understanding. For example, a hypothesis could be "Increased physical exercise is associated with a decrease in blood pressure." This hypothesis can be tested through research by collecting data on the amount of physical exercise and blood pressure levels in a sample population and analyzing the relationship between the two variables.***