Prerequisite Homework (Written Answers)

Thursday, January 19, 2023 1:49 PM

Landyn Francis

Problem 4: Probability and Statistics

1. Write the definitions of Random Variables, Expectation, and Variance.

   A random variable is a variable whose values are outcomes of a random event. It maps the outcomes of the random event to numerical values. There are two types, continuous and discrete. Continuous random variables take on any value in an interval. For example, the height of women in centimeters is continuous, there are infinitely many values in the range. Discrete random variables take on countable values in the interval. For example, a die being tossed, and counting the number of times a 1 is rolled. These values can only be integer values, 0, 1, 2, etc.

   The expectation is the theoretical mean of a random variable. It is calculated by summing each outcome weighted by it's probability of occurring. Sometimes called the long-term average. $E(x) = \sum_{\text{all } x} x \cdot p(x)$

   Variance is the average square distance from the mean, or the expected value of the square difference of the random variable to it's mean $V(x) = E[(X-\mu)^2]$ where μ is the mean

2. What is the difference between Expectation and a sample mean? When do they converge?

   The expectation or expected value is the mean of a probability distribution of a random variable and represents a theoretical mean, while the sample mean is the average value of already collected data and represents a real mean. These values converge when the sample size is infinitely large.

3. Write the three axioms of a probability distribution

   Axiom 1: The probability of any event is between 0 and 1.
   Axiom 2: For any sample space, the probability of the entire sample space is 1.
   Axiom 3: If two events A and B are mutually exclusive, then the probability of either A or B occurring is equal to the sum of each of their individual probabilities.

4. What is the difference between a probability density function, a probability mass function, and a cumulative distribution function?

   A probability mass function is a function that gives the probability of each outcome for a discrete random variable. A probability density function gives the probability of any outcome for a continuous random variable. Cumulative distribution functions are used to evaluate the accumulated probability at a given outcome, basically the probability that the outcome will be less than or equal to a given outcome. The probability density function is the derivative of the cumulative distribution function.

5. Prove that for any Random Variable X, the Variance V(X) = E[X^2] - E[X]^2 $\underbrace{V(x)}_{\text{LHS}} = \underbrace{E(x^2) - E[x]^2}_{\text{RHS}}$

   Variance can be defined as the following.
   $Var(X) = E\left[(X-\mu)^2\right]$
   This is the expected value of the square difference between the random variable and the mean.
   Expanding this we get the following:
   $Var(X) = E\left[X^2 - 2\mu X + \mu^2\right]$
   The mean can be written as the expected value of the random variable as sample size grows to infinity. And because the expectation is linear, we can apply it to each term.
   $= E\left[X^2\right] + E\left[-2XE[X]\right] + E\left[E[X]^2\right]$
   Because the mean is constant.
   $= E\left[X^2\right] - 2E[X]E[X] + E[X]^2E[1]$
   $= E\left[X^2\right] - 2E[X]^2 + E[X]^2$
   $= E\left[X^2\right] - E[X]^2$
   = Right Hand Side :)

6. Prove that E[X^2] > 0

   This quantity is known as the second moment, and can also be represented as the Variance plus the expected value squared. Assume the sample size is infinitely large, so the expected value can be treated as the mean. We get the following equation.
   $E\left[X^2\right] = Var(X) + \sigma^2$
   As seen in Question 5, Variance is defined as the square difference between a random variable and it's mean. It can be written:
   $Var(X) = E\left[(X-\mu)^2\right]$
   Because the inside term is squared, we know this will always be a positive term. As well, the mean is squared and will always be a positive term. Therefore, $E\left[X^2\right] > 0$.

   The only way this could be equal to zero is when the random variable is exactly equal to the mean, and the mean is also 0. This is a trivial case.

Problem 5:

1. Find the minimum point in the function f(x) = x^2 - 6x +33 using calculus. What property of derivatives did you use to reach the answer?

   The maximum and minimum point in a function can found by determining when the slope of the function is equal to zero. For this we can find the derivative, and set it equal to 0.

   $y(x) = x^2 - 6x + 3$

   2u

to reach the answer?

The maximum and minimum point in a function can found by determining when the slope of the function is equal to zero. For this we can find the derivative, and set it equal to 0.

$y(x) = x^2 - 6x + 3$

$y'(x) = 2x - 6$

$0 = 2x - 6$

$3 = x$

$y(3) = 24$

The minimum point is at $(3, 24)$

2. Find the minimum point in the function f(x,y) = x^2 - 6x + y^2 - 8x - xy + 33 using calculus. What property of derivatives did you use to reach the answer?

   The maximum and minimum can also be found in this case by setting derivatives to zero. With multiple variables however, partial derivatives must be used.

$f(x,y)\ x^2 - 14x + y^2 - xy + 33$

$\frac{\partial}{\partial x} = 2x - 14 - y \qquad \frac{\partial}{\partial y} = 2y - x$

$0 = 2x - y - 14$ - solve the system

$0 = 2y - x$

$\boxed{x = \frac{28}{3} \quad y = \frac{14}{3}}$

To make sure this is a minimum we must look at the second derivatives

$\frac{\partial^2}{\partial x^2} = 2 \qquad \frac{\partial^2}{\partial x \partial y} = -1 \qquad \frac{\partial^2}{\partial y^2} = 2$

$D = \frac{\partial^2}{\partial x^2} \frac{\partial^2}{\partial y^2} - \left(\frac{\partial^2}{\partial y \partial x}\right)^2 = 3$

Since D is greater than zero we know this is a minimum point

Problem 6:

1. Given two matrices **A** of size **m x n** and **B** of size **p x q** when is the matrix multiplication **AB** valid? When is the matrix multiplication **BA** valid? When is the addition **A+B** valid?

   Matrix multiplication is valid when the columns in the first matrix are equal to the rows in the second matrix. For **AB n must equal p**, similarly, for **BA**, **q must equal p.** Matrix addition is only valid when the dimensions of both matrices are equal. So for **A+B m must equal p and n must equal q.**

2. Define dot product for two vectors. How do you test if two given vectors are perpendicular? Assume you have two n-dimensional vectors $\vec{a} = [a_1, a_2, \ldots a_n]$ and $\vec{b} = [b_1, b_2, \ldots b_n]$,
   Denote dot product as, $\vec{a} * \vec{b}$

   The dot product for two vectors is the value calculated by multiplying corresponding elements from each vector together, then adding the products. For example, $\vec{a} * \vec{b}$ would be calculated as $a_1 * b_1 + a_2 * b_2 + a_3 * b_3 + \cdots + a_n b_n$ if this value is equal to 0, then the two vectors are perpendicular.

3. Define cross product for two vectors? How to test when two vectors are parallel to each other?

   The cross product for two vectors is a vector that is perpendicular to both of the original vectors. It can be found using the either the cartesian coordinates of the vector, or the magnitude and angle of the vectors. The following equation shows the coordinate method.

   $AxB = (A_y B_z - A_z B_y)i + (A_z B_x - A_x B_z)j + (A_x B_y - A_y B_x)k$

Where $i, j, and\ k$ are unit vectors in the x,y and z directions.
Using the magnitudes and coordinates the equation is as follows:
$AxB = |A| * |B| * \sin(\theta) * n$
Where |A| and |B| are the magnitudes of the vector, θ is the angle between them, and
n is the "normal vector" or a unit vector perpendicular to both A and B
Two vectors are parallel to each other if the cross product is equal to the zero vector.

4. What is a unit vector? How can you find the magnitude of a vector?
   A unit vector is a vector with a magnitude of 1. They are often used for representing direction but not magnitude. The magnitude of a simple 2 dimensional vector with endpoints at $(x_1, y_1)\ and\ (x_2, y_2)$ can be found with the following equation.
   $|V| = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$