Probability Theory Lecture 03

Padmini Mukkamala

Budapest University of Technology and Economics

September 27, 2020

Overview

Random Variable

Definition

Any numerical function $X = X(\omega)$, defined on a (finite) sample space Ω is called a (simple) Random variable. So, $X : \Omega \to \mathbb{R}$.

Examples

Example 1: $X =$ number of heads in three coin tosses
Example 2: $Y =$ sum of two dice rolls.

Three important Random Variables

Indicator Random variable

Characteristic function of an event A. Let the probability of the event A be p, then this random variable is denoted by $1(p)$ or 1_A.

$$1_A(\omega) = \begin{cases} 1 & \omega \in A \\ 0 & \omega \notin A \end{cases}$$

Three important Random Variables cont.

Bernoulli Random variable

When there are only two outcomes in the sample space and we term one as success and the other as failure, giving success value 1 and failure value 0, then the random variable is called **Bernoulli** random variable. Examples: Coin toss, Pass/Fail on an exam.

$$X = \begin{cases} 1 & SUCCESS \\ 0 & FAILURE \end{cases}$$

Three important Random Variables cont.

Binomial Random variable

Denoted by $Bin(n, p)$, this denotes the number of successes in n independent, identical Bernoulli trials, each with success probability p. $X \in \{0, 1, 2, .., n\}$.
Examples: number of Heads in n coin tosses, number of students passing this course, number of people at an airport having Covid, number of road accidents etc.

Probability Mass Function

Definition of PMF for Simple Random Variables

Let X be a **simple** random variable that takes values $\{x_1, ..., x_m\}$. Let $A_i = \{\omega | X(\omega) = x_i\}$. We define the probability mass function (pmf) of X as follows:

$$P_X(x_i) = P(A_i)$$

This is also denoted as $P(X = x_i)$, $p(x_i)$ or $p_X(x_i)$. All notations are fairly standard and good to be familiar with.

Probability Mass Function

Random Variables that are not Simple

The above definition for pmf also holds for random variables that are not simple. But one must be more careful in defining the event space (sigma algebra) and probability measure (sigma additivity) in case of infinite Sample spaces. These details are not discussed in this class but we will use pmfs for such variables.

Probability Mass Function

Example 1: Number of heads in three coin tosses

Probability Mass Function

Example 1: Number of heads in three coin tosses

Value of X	$p(x_i)$
0	$\frac{1}{8}$
1	$\frac{3}{8}$
2	$\frac{3}{8}$
3	$\frac{1}{8}$

Table: Pmf of number of heads in three coin tosses

Probability Mass Function

Example 2: Sum of two dice rolls.

Value of Y	$p(y_i)$	Value of Y	$p(y_i)$
2	$\frac{1}{36}$	7	$\frac{6}{36}$
3	$\frac{2}{36}$	8	$\frac{5}{36}$
4	$\frac{3}{36}$	9	$\frac{4}{36}$
5	$\frac{4}{36}$	10	$\frac{3}{36}$
6	$\frac{5}{36}$	11	$\frac{2}{36}$
		12	$\frac{1}{36}$

Table: Pmf of sum of two dice

Probability Mass Function

Example 3: Indicator and Bernoulli random variables .

Value of 1_A	$p(x)$	Value of $1_{SUCCESS}$	$p(x)$
0	$1 - P(A)$	0	$1 - p$
1	$P(A)$	1	p

Note that the pmfs of Indicator and Bernoulli random variables are identical.

Probability Mass Function

Example 4: Binomial random variable.

$$P(X = i) = \binom{n}{i} p^i (1 - p)^{n-i}$$

Cumulative Distribution Function

Definition

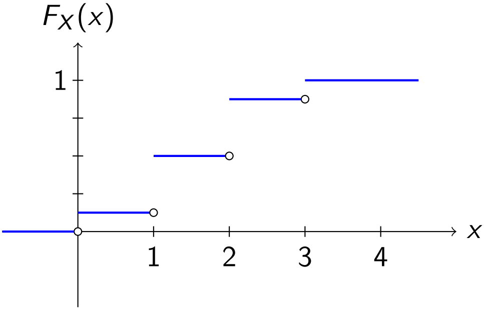
The Cumulative Distribution Function of a random variable X, denoted by $F_X(a)$, is defined as:

$$F_X(a) = P(X \leq a) = \sum_{x \leq a} p(x)$$

This is also called a Step Function because of its shape for discrete random variables.

Note: It is just as often defined with $x < a$ but in this course we will use $x \leq a$.

Cumulative Distribution Function



CDF for the number of Heads in three dice rolls.

Cumulative Distribution Function

Properties

- $F_X : \mathbb{R} \to [0, 1]$

Cumulative Distribution Function

Properties

- $F_X : \mathbb{R} \to [0, 1]$
- It is a monotone increasing function, i.e,
  $\forall a \leq b, F_X(a) \leq F_X(b)$.

Cumulative Distribution Function

Properties

- $F_X : \mathbb{R} \to [0, 1]$
- It is a monotone increasing function, i.e, $\forall a \leq b, F_X(a) \leq F_X(b)$.
- It is right continuous, i.e, $\lim_{x \to a^+} F_X(x) \to F_X(a)$

Expected value of a random variable

Definition

The expected value of a discrete random variable X, denoted by $E(X)$ or μ_X is,

$$E(X) = \mu_X = \sum_x xP(X = x) = \sum_x xp(x)$$

Expected value of a random variable

Definition

The expected value of a discrete random variable X, denoted by $E(X)$ or μ_X is,

$$E(X) = \mu_X = \sum_x xP(X = x) = \sum_x xp(x)$$

Note: relative frequency tends to $p(x)$ for a large number of experiments, so expected value can be thought of as the average value of the random variable X for a large number of experiments.

Expected value of a random variable cont.

Examples

Example 1: number of heads in three coin tosses.

Expected value of a random variable cont.

Examples

Example 1: number of heads in three coin tosses.

$$E(X) = 0 \times \frac{1}{8} + 1 \times \frac{3}{8} + 2 \times \frac{3}{8} + 3 \times \frac{1}{8} = 1.5$$

Expected value of a random variable cont.

Examples

Example 2: result of a single dice roll.

Expected value of a random variable cont.

Examples

Example 2: result of a single dice roll.

$$E(X) = (1 + 2 + 3 + 4 + 5 + 6) \times \frac{1}{6} = 3.5$$

Example 3: sum of two dice rolls.

$$E(X) = 2 \times \frac{1}{36} + 3 \times \frac{2}{36} + 4 \times \frac{3}{36} + 5 \times \frac{4}{36} + 6 \times \frac{5}{36} + 7 \times \frac{6}{36} +$$

$$8 \times \frac{5}{36} + 9 \times \frac{4}{36} + 10 \times \frac{3}{36} + 11 \times \frac{2}{36} + 12 \times \frac{1}{36} = 7$$

Expected value of a random variable cont.

Examples

Example 4: of the Indicator random variable 1_A or $1(p)$.
Important

Expected value of a random variable cont.

Examples

Example 4: of the Indicator random variable 1_A or $1(p)$.
Important

$$E(1_A) = 0 \times (1 - P(A)) + 1 \times P(A) = P(A)$$

Example 5: of Bernoulli random variable $1(p)$.

Expected value of a random variable cont.

Examples

Example 4: of the Indicator random variable 1_A or $1(p)$.

Important

$$E(1_A) = 0 \times (1 - P(A)) + 1 \times P(A) = P(A)$$

Example 5: of Bernoulli random variable $1(p)$.

$$E(1(p)) = p$$

Properties of Expected Value

- Translation: $E(X + b) = E(X) + b$.

Properties of Expected Value

- Translation: $E(X+b) = E(X) + b$.
- Scaling: $E(aX) = aE(X)$.

Properties of Expected Value

- Translation: $E(X+b) = E(X) + b$.
- Scaling: $E(aX) = aE(X)$.
- Function: $E(g(X)) = \sum_x g(x)p(x)$ (also called the Law of the Unconscious Statistician).
  Proof for simple RVs: Use $Y = g(X)$, then,
  $E(g(X)) = E(Y) = \sum_y yp(y)$.

Example: Let me bet that I will pay 10$ for every heads that turns up in three coin tosses. Then $g(X) = 10X$ and so the expected amount of money I will pay you is 15$.

Properties of Expected Value cont.

Variance

When $g(X) = (X - \mu_X)^2$, then $E(g(X)) = E((X - \mu_X)^2)$ is called the Variance, $Var(X)$, of the random variable X.

Variance denotes how much the random variable deviates from its mean. We will discuss more in Lecture 7!

Properties of Expected Value cont.

Moments

When $g(X) = X^i$, then $E(g(X)) = E(X^i)$ is called the i^{th} moment of the random variable X.

Example

For the random variable counting the number of Heads in three coin tosses, the third moment is,

Properties of Expected Value cont.

Moments

When $g(X) = X^i$, then $E(g(X)) = E(X^i)$ is called the i^{th} moment of the random variable X.

Example

For the random variable counting the number of Heads in three coin tosses, the third moment is,

$$E(X^3) = 0 \times \frac{1}{8} + 1 \times \frac{3}{8} + 8 \times \frac{3}{8} + 27 \times \frac{1}{8} = \frac{54}{8}$$

Properties of Expected Value cont.

Linearity of Expectation

For any two random variables X, Y (**not necessarily independent**), the following always holds:

$$E(X+Y) = E(X) + E(Y)$$

Proof

Let $Z = X + Y$, and we are interested in $E(Z)$.
$E(Z) = \sum_z zP(X+Y=z) = \sum_z \sum_{x,y,x+y=z} (x+y)P(X = x, Y = z-x) = \sum_x \sum_y (x+y)P(X=x, Y=y) = \sum_x \sum_y xP(X=x, Y=y) + \sum_y \sum_x yP(X=x, Y=y) = E(X) + E(Y)$.

Properties of Expected Value cont.

Examples of Linearity of expectation

Example 1: Sum of two dice rolls. We know the expected value of a single dice roll is 3.5. Then the expected value of two dice rolls is $3.5 + 3.5 = 7$.

Properties of Expected Value cont.

Examples of Linearity of expectation cont.

Example 2: Expected value of Binomial Random variable.
First Method: without using Linearity of Expectation:

$$E(X) = \sum_i i \binom{n}{i} p^i (1-p)^{n-i}$$

Properties of Expected Value cont.

Examples of Linearity of expectation cont.

Example 2: Expected value of Binomial Random variable.
First Method: without using Linearity of Expectation:

$$E(X) = \sum_i i \binom{n}{i} p^i (1-p)^{n-i}$$

Second Method: using Linearity of Expectation: Let $X = Bin(n, p)$. Then we set $X = X_1 + X_2 + ... + X_n$ where each X_i is independent random variable with distribution $1(p)$. Then, by Linearity of Expectation,

$$E(X) = \sum_i E(X_i) = np$$

Expected value of a random variable cont.

Example of Use of 1_A

We notice that given any two events A and B, the following always holds:

$1_{A \cup B} = 1_A + 1_B - 1_{A \cap B}$

Expected value of a random variable cont.

Example of Use of 1_A

We notice that given any two events A and B, the following always holds:

$1_{A\cup B} = 1_A + 1_B - 1_{A\cap B}$

Taking the expectation and using Linearity of expectation, we get,

$E(1_{A\cup B}) = E(1_A) + E(1_B) - E(1_{A\cap B})$

or, $P(A\cup B) = P(A) + P(B) - P(A\cap B)$

This gives another proof of the formula of Inclusion-Exclusion.

Geometric Distribution

Definition

Let X be a random variable that denotes the number of times a coin is tossed until it turns up Heads. (Here Heads can be thought of as Success, that is number of times some experiment is repeated until it results in Success). Then, X is said to have **Geometric distribution** and the probability mass function of X is given by:

$$P(X = i) = (1 - p)^{i-1} p$$

Properties

Expected value: $E(X) = \frac{1}{p}$.

Expected value of an infinite random variable

Example where $E(X)$ does not exist

Suppose we toss a coin until it turns up Heads, and I will pay you 2^i dollars if it took i turns. What is the expected amount of money you will make?

$$2 \times \frac{1}{2} + 4 \times \frac{1}{4} + 8 \times \frac{1}{8} + \ldots$$

$E(X)$ for infinite random vairable

For an infinite R.V., if the sum $\sum_i x_i p(x_i)$ absolutely converges, i.e., $\sum_i |x_i| p(x_i) = L$ for some real number L, then $E(X) = \sum_i x_i p(x_i)$.

Coupon Collector's Problem

Coupon Collector's Problem

If you shop at Auchan for groceries, you receive a gift packet for every 5000 forints you spend. The gift packet contains a character from the Harry Potter story. Suppose every character is equally likely to turn up, how much money should you spend to be able to collect all the 10 figures they are giving?

Solution

Let X_i be the number of gift packets opened after receiving the $i-1$ character and until the i^{th} character.

Coupon Collector's Problem

Coupon Collector's Problem

If you shop at Auchan for groceries, you receive a gift packet for every 5000 forints you spend. The gift packet contains a character from the Harry Potter story. Suppose every character is equally likely to turn up, how much money should you spend to be able to collect all the 10 figures they are giving?

Solution

Let X_i be the number of gift packets opened after receiving the $i-1$ character and until the i^{th} character. Notice that X_i has geometric distribution with probability

Coupon Collector's Problem

Coupon Collector's Problem

If you shop at Auchan for groceries, you receive a gift packet for every 5000 forints you spend. The gift packet contains a character from the Harry Potter story. Suppose every character is equally likely to turn up, how much money should you spend to be able to collect all the 10 figures they are giving?

Solution

Let X_i be the number of gift packets opened after receiving the $i-1$ character and until the i^{th} character. Notice that X_i has geometric distribution with probability $\frac{n-i+1}{n}$.

Coupon Collector's Problem

Coupon Collector's Problem

If you shop at Auchan for groceries, you receive a gift packet for every 5000 forints you spend. The gift packet contains a character from the Harry Potter story. Suppose every character is equally likely to turn up, how much money should you spend to be able to collect all the 10 figures they are giving?

Solution

Let X_i be the number of gift packets opened after receiving the $i-1$ character and until the i^{th} character. Notice that X_i has geometric distribution with probability $\frac{n-i+1}{n}$.
So the expected number of packets we need, by Linearity of Expectation is $1+\frac{n}{n-1}+\frac{n}{n-2}+\ldots+\frac{n}{1}$

The End