

CSE400 – Fundamentals of Probability in Computing

Lecture 6: Discrete Random Variables, Expectation and Problem Solving

Instructor: Dhaval Patel, PhD

January 22, 2025

L6_S2_A

1 Random Variables

1.1 Definition and Concept

Let Ω denote the sample space.

$\omega \in \Omega$ denote a sample point.

A random variable X is defined as a function:

$$X : \Omega \rightarrow \mathbb{R}$$

That assigns a real number to each sample point $\omega \in \Omega$:

$$X(\omega) \in \mathbb{R}$$

Thus, a random variable converts outcomes of an experiment into real numerical values.

1.2 Discrete Random Variables

The lecture restricts attention initially to discrete random variables, which satisfy:

They take values from a finite or countably infinite set.

The actual set of values:

$$\{X(\omega) : \omega \in \Omega\}$$

forms a discrete subset of the real numbers.

This means the values can be listed explicitly.

1.3 Visualization of Random Variable Distribution

The distribution of a discrete random variable can be visualized using a bar diagram:

The x-axis represents possible values of the random variable.

The height of each bar represents:

$$P(X = a)$$

Each probability corresponds to the probability of the associated event in the sample space.

2 Types of Random Variables

2.1 Discrete Random Variable

A random variable is discrete if:

- It has countable support.
- It is described using a Probability Mass Function (PMF).
- Probabilities are assigned to individual values.
- Each possible value has strictly positive probability.

2.2 Continuous Random Variable

A random variable is continuous if:

- It has uncountable support.
- It is described using a Probability Density Function (PDF).
- Probabilities are assigned over intervals.
- Each individual value has probability zero.

3 Example: Tossing Three Fair Coins

3.1 Experiment Description

An experiment consists of tossing 3 fair coins.

Define:

$$Y = \text{number of heads observed}$$

Possible values of Y :

$$Y \in \{0, 1, 2, 3\}$$

3.2 Computing Probabilities

The sample space contains 8 equally likely outcomes:

$$\{TTT, TTH, THT, THH, HTT, HTH, HHT, HHH\}$$

Each has probability:

$$\frac{1}{8}$$

Case 1: No heads

$$P(Y = 0) = P(TTT) = \frac{1}{8}$$

Case 2: One head

Possible outcomes:

$$\{TTH, THT, HTT\}$$

Thus:

$$P(Y = 1) = \frac{3}{8}$$

Case 3: Two heads

Possible outcomes:

$$\{THH, HTH, HHT\}$$

Thus:

$$P(Y = 2) = \frac{3}{8}$$

Case 4: Three heads

Possible outcome:

$$HHH$$

Thus:

$$P(Y = 3) = \frac{1}{8}$$

3.3 Validity Condition of PMF

Since Y must take exactly one of these values:

$$P(Y = 0) + P(Y = 1) + P(Y = 2) + P(Y = 3) = 1$$

Thus,

$$\frac{1}{8} + \frac{3}{8} + \frac{3}{8} + \frac{1}{8} = 1$$

This confirms the distribution is valid.

4 Probability Mass Function (PMF)

4.1 Definition

A random variable that can take at most countable values is called a discrete random variable.

Let the range of values be:

$$R_X = \{x_1, x_2, x_3, \dots\}$$

Define the function:

$$P_X(x_k) = P(X = x_k)$$

This function is called the Probability Mass Function (PMF) of X .

4.2 Properties of PMF

The PMF must satisfy:

Property 1: Non-negativity

$$P_X(x_k) \geq 0$$

Property 2: Total probability equals 1

Since the random variable must take one of its possible values:

$$\sum_{k=1}^{\infty} P_X(x_k) = 1$$

This follows because the possible values are mutually exclusive and collectively exhaustive events.

5 Example: Finding Unknown Constant in PMF

5.1 Given

The PMF is:

$$p(i) = c \frac{\lambda^i}{i!}, \quad i = 0, 1, 2, \dots$$

where $\lambda > 0$.

We must find:

$$P(X = 0)$$

$$P(X > 2)$$

5.2 Step 1: Use Total Probability Condition

Since PMF must sum to 1:

$$\sum_{i=0}^{\infty} p(i) = 1$$

Substitute PMF:

$$\sum_{i=0}^{\infty} c \frac{\lambda^i}{i!} = 1$$

Factor out constant:

$$c \sum_{i=0}^{\infty} \frac{\lambda^i}{i!} = 1$$

Using known identity:

$$\sum_{i=0}^{\infty} \frac{\lambda^i}{i!} = e^{\lambda}$$

Thus:

$$ce^{\lambda} = 1$$

Solve for c :

$$c = e^{-\lambda}$$

5.3 Step 2: Find $P(X = 0)$

Substitute into PMF:

$$\begin{aligned} P(X = 0) &= c \frac{\lambda^0}{0!} \\ &= e^{-\lambda} \end{aligned}$$

5.4 Step 3: Find $P(X > 2)$

Use complement rule:

$$\begin{aligned}P(X > 2) &= 1 - P(X \leq 2) \\&= 1 - [P(X = 0) + P(X = 1) + P(X = 2)]\end{aligned}$$

Substitute:

$$= 1 - \left(e^{-\lambda} + \lambda e^{-\lambda} + \frac{\lambda^2 e^{-\lambda}}{2} \right)$$

Thus,

$$P(X > 2) = 1 - e^{-\lambda} \left(1 + \lambda + \frac{\lambda^2}{2} \right)$$

6 Bayes' Theorem

6.1 Conditional Probability Relationship

Using definition of conditional probability:

$$Pr(A|B) = \frac{Pr(B|A)Pr(A)}{Pr(B)}$$

This gives Bayes' Formula:

$$Pr(B_i|A) = \frac{Pr(A|B_i)Pr(B_i)}{\sum_{i=1}^n Pr(A|B_i)Pr(B_i)}$$

6.2 Terminology

Prior probability

$$Pr(B_i)$$

Probability before observing new information.

Posterior probability

$$Pr(B_i|A)$$

Probability after observing event A .

7 Bayes' Theorem Example: Auditorium

7.1 Problem Description

An auditorium has 30 rows.

Number of seats in rows:

Row 1 \rightarrow 11 seats

Row 2 \rightarrow 12 seats

Row 3 \rightarrow 13 seats

...

Row 30 \rightarrow 40 seats

Selection procedure:

Step 1: Select a row randomly (equal probability)

$$P(R_k) = \frac{1}{30}$$

Step 2: Select a seat randomly within that row.

7.2 Find $P(S_{15}|R_{20})$

Row 20 has:

$$10 + 20 = 30 \text{ seats}$$

Thus,

$$P(S_{15}|R_{20}) = \frac{1}{30}$$

7.3 Find $P(R_{20}|S_{15})$

Using Bayes' formula:

$$P(R_{20}|S_{15}) = \frac{P(S_{15}|R_{20})P(R_{20})}{\sum_{k=5}^{30} P(S_{15}|R_k)P(R_k)}$$

Since seat 15 exists only in rows 5 through 30.

Denominator:

$$\sum_{k=5}^{30} \frac{1}{k+10} \cdot \frac{1}{30}$$

Numerical evaluation from lecture:

$$P(R_{20}|S_{15}) \approx 0.0325$$

8 Summary of Key Results

Random Variable

$$X : \Omega \rightarrow \mathbb{R}$$

PMF

$$P_X(x_k) = P(X = x_k)$$

$$\sum P_X(x_k) = 1$$

Exponential Identity

$$\sum \frac{\lambda^i}{i!} = e^\lambda$$

Bayes' Theorem

$$P(B_i|A) = \frac{P(A|B_i)P(B_i)}{\sum P(A|B_i)P(B_i)}$$