CSE 400: Fundamentals of Probability in Computing

Lecture 4: Joint & Conditional Probability

1. Lecture Overview

Lecture Title: Joint & Conditional Probability
This lecture introduces:

- Joint probability distributions of discrete random variables
- Marginal probability distributions
- Conditional probability mass functions
- Independence of random variables
- Bayes' Rule
- Law of Total Probability

2. Random Variables (Setup)

Let:

- (X) and (Y) be **discrete random variables**

Assume:

- (X) takes values in the set ($\mathcal{X}$)
- (Y) takes values in the set ($\mathcal{Y}$)

Each random variable maps outcomes of the sample space to real numbers.

3. Joint Probability Mass Function (Joint PMF)

Definition

The **joint probability mass function** of discrete random variables (X) and (Y) is defined as:

$$p_{X,Y}(x,y) = \Pr(X = x \text{ and } Y = y)$$

for all:

$$x \in \mathcal{X}, \quad y \in \mathcal{Y}$$

Properties of the Joint PMF

1. **Non-negativity**

$$p_{X,Y}(x,y) \geq 0 \quad \forall x, y$$

2. **Normalization**

$$\sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p_{X,Y}(x,y) = 1$$

4. Joint Distribution Table

A **joint distribution** of two discrete random variables may be represented in tabular form:

- Rows correspond to values of (X)
- Columns correspond to values of (Y)
- Each cell contains ($p_{X,Y}(x, y)$)

The sum of all entries in the table equals **1**.

5. Marginal Probability Mass Functions

Marginal PMF of (X)

The **marginal probability mass function** of (X) is obtained by summing the joint PMF over all values of (Y):

$$p_X(x) = \sum_{y \in \mathcal{Y}} p_{X,Y}(x, y)$$

Marginal PMF of (Y)

Similarly, the marginal PMF of (Y) is:

$$p_Y(y) = \sum_{x \in \mathcal{X}} p_{X,Y}(x, y)$$

Properties of Marginals

$$\sum_x p_X(x) = 1, \quad \sum_y p_Y(y) = 1$$

Marginal distributions describe the probability behavior of one random variable independent of the other.

6. Example: Joint and Marginal Distribution

(As shown in the lecture slides)
Given a joint PMF table:

- ($p_X(x)$) is computed by summing across rows
- ($p_Y(y)$) is computed by summing down columns

All intermediate summation steps are shown explicitly in the lecture.

7. Conditional Probability Mass Function

Let (X) and (Y) be discrete random variables with joint PMF ($p_{X,Y}(x, y)$).
Assume:

$$p_Y(y) > 0$$

Conditional PMF of (X) Given (Y = y)

$$p_{X|Y}(x \mid y) = \Pr(X = x \mid Y = y)$$

$$= \frac{p_{X,Y}(x, y)}{p_Y(y)}$$

for all ($x \in \mathcal{X}$) such that ($p_Y(y) > 0$).

Properties

1. **Non-negativity**

$$p_{X|Y}(x \mid y) \geq 0$$

2. **Normalization**

$$\sum_{x \in \mathcal{X}} p_{X|Y}(x \mid y) = 1$$

Conditional PMF of (Y) Given (X = x)

Assume:

$$p_X(x) > 0$$

$$p_{Y|X}(y \mid x) = \frac{p_{X,Y}(x, y)}{p_X(x)}$$

8. Relationship Between Joint, Marginal, and Conditional PMFs

From the definition of conditional probability:

$$p_{X,Y}(x, y) = p_{X|Y}(x \mid y), p_Y(y)$$

Similarly,

$$p_{X,Y}(x, y) = p_{Y|X}(y \mid x), p_X(x)$$

Consistency Check

$$\sum_x p_{X,Y}(x, y) = \sum_x p_{X|Y}(x \mid y), p_Y(y)$$

$$= p_Y(y) \sum_x p_{X|Y}(x \mid y) = p_Y(y)$$

9. Example: Conditional Probability Computation

(As presented in the lecture)
Steps:

1. Compute ($p_Y(y)$)
2. Use

$$p_{X|Y}(x \mid y) = \frac{p_{X,Y}(x, y)}{p_Y(y)}$$

3. Verify normalization

All numerical steps are shown explicitly in the slides.

10. Independence of Random Variables

Definition

Discrete random variables (X) and (Y) are **independent** if:

$$p_{X,Y}(x,y) = p_X(x), p_Y(y)$$

for all ($x \in \mathcal{X}$), ($y \in \mathcal{Y}$).

Equivalent Conditional Characterization

If (X) and (Y) are independent:

$$p_{X|Y}(x \mid y) = p_X(x)$$

$$p_{Y|X}(y \mid x) = p_Y(y)$$

whenever the conditioning probability is positive.

11. Example: Checking Independence

(As shown in the lecture)
Steps:

1. Compute ($p_X(x)$) and ($p_Y(y)$)
2. Compute products ($p_X(x)p_Y(y)$)
3. Compare with ($p_{X,Y}(x,y)$)

If equality holds for all entries → independent
Otherwise → dependent

12. Bayes' Rule

Assume:

$$p_X(x) > 0, \quad p_Y(y) > 0$$

Statement

$$p_{X|Y}(x \mid y) = \frac{p_{Y|X}(y \mid x), p_X(x)}{p_Y(y)}$$

Derivation

From joint probability:

$$p_{X,Y}(x,y) = p_{X|Y}(x \mid y), p_Y(y)$$

$$p_{X,Y}(x,y) = p_{Y|X}(y \mid x), p_X(x)$$

Equating:

$$p_{X|Y}(x \mid y), p_Y(y) = p_{Y|X}(y \mid x), p_X(x)$$

Dividing by ($p_Y(y)$):

$$p_{X|Y}(x \mid y) = \frac{p_{Y|X}(y \mid x), p_X(x)}{p_Y(y)}$$

13. Example: Bayes' Rule Application

(As presented in the lecture)

Steps:

1. Identify ($p_{Y|X}(y \mid x)$)
2. Identify ($p_X(x)$)
3. Compute ($p_Y(y)$) via marginalization
4. Substitute into Bayes' Rule

All intermediate computations are shown in the slides.

14. Law of Total Probability

Assume:

$$p_Y(y) > 0 \quad \forall y \in \mathcal{Y}$$

Statement

$$p_X(x) = \sum_{y \in \mathcal{Y}} p_{X|Y}(x \mid y), p_Y(y)$$

Derivation

Start from marginal probability:

$$p_X(x) = \sum_y p_{X,Y}(x,y)$$

Substitute:

$$p_{X,Y}(x,y) = p_{X|Y}(x \mid y), p_Y(y)$$

Thus:

$$p_X(x) = \sum_y p_{X|Y}(x \mid y), p_Y(y)$$

Symmetric Form

Assuming:

$$p_X(x) > 0$$

$$p_Y(y) = \sum_{x \in \mathcal{X}} p_{Y|X}(y \mid x), p_X(x)$$

15. Example: Law of Total Probability

(As shown in the lecture)

- Conditional probabilities are given
- Marginal probabilities are substituted
- Summation is computed term-by-term

All intermediate values are shown explicitly.

16. Summary of Core Identities

1. **Joint from conditional**

$$p_{X,Y}(x, y) = p_{X|Y}(x \mid y), p_Y(y)$$

2. **Bayes' Rule**

$$p_{X|Y}(x \mid y) = \frac{p_{Y|X}(y \mid x), p_X(x)}{p_Y(y)}$$

3. **Law of Total Probability**

$$p_X(x) = \sum_{y} p_{X|Y}(x \mid y), p_Y(y)$$

17. End of Lecture 4

This lecture covers:

- Joint PMFs
- Marginal PMFs
- Conditional PMFs
- Independence
- Bayes' Rule
- Law of Total Probability