

Lecture Scribe CSE 400



L5_A_S1 (2).pdf
PDF

Generate a lecture scribe intended strictly as exam-oriented reference material for CSE 400: Fundamentals of Probability in Computing – Lecture 5, covering Bayes' Theorem, Random Variables, and Probability Mass Function.

Use only the attached Lecture 5 slides/PDF and any explicitly relevant textbook content connected to this lecture.

Do not introduce any material, definitions, notation, explanations, intuition, examples, assumptions, or results that are not explicitly present in the provided context.

While writing the scribe, reason step by step through the lecture content exactly as presented, including:

Definitions and notation

Assumptions and conditions

Statements of results (including the Law of Total Probability and Bayes' Formula)

Proofs or derivations, with all intermediate steps

Worked examples (including all numerical computations and conditional probability steps)

Ensure that the logical flow mirrors the lecture sequence, including:

Weighted average interpretation of conditional probabilities

Learning-by-example applications of Bayes' Theorem

Formal derivation using the Law of Total Probability

Definition and interpretation of random variables

Distributions of discrete random variables

Definition and properties of the probability mass function (PMF)

PMF-based examples exactly as shown in the lecture

The reasoning steps must be explicit and complete, so that a student preparing for a reading-based / closed-notes exam can reconstruct all arguments and calculations solely from this scribe.

Maintain a formal academic tone, clear sectioning, and faithful adherence to the lecture's scope and structure. Do not add summaries, commentary, simplifications, alternative explanations, or material beyond what appears in the provided context.

Below is a **faithful, exam-oriented lecture scribe** for **CSE 400: Fundamentals of Probability in Computing – Lecture 5**, reconstructed **strictly from the provided Lecture 5 slides/PDF**, preserving the **exact logical flow, notation, results, and worked computations** as presented in the lecture.

No additional material beyond the given context is introduced.

All content is derived from the attached Lecture 5 slides .

CSE 400: Fundamentals of Probability in Computing

Lecture 5 — Bayes' Theorem, Random Variables, and Probability Mass Function

1. Bayes' Theorem

1.1 Weighted Average of Conditional Probabilities

Let A and B be events.

We may express the event A as

$$A = AB \cup AB^c$$

since, for an outcome to be in A , it must either be in both A and B , or be in A but not in B .

The events AB and AB^c are mutually exclusive.

By **Axiom 3**,

$$\Pr(A) = \Pr(AB) + \Pr(AB^c)$$

Using the definition of conditional probability,

$$\Pr(AB) = \Pr(A \mid B) \Pr(B)$$

$$\Pr(AB^c) = \Pr(A \mid B^c) \Pr(B^c)$$

Hence,

$$\begin{aligned} \Pr(A) &= \Pr(A \mid B) \Pr(B) + \Pr(A \mid B^c) \Pr(B^c) \\ &= \Pr(A \mid B) \Pr(B) + \Pr(A \mid B^c) [1 - \Pr(B)] \end{aligned}$$

Interpretation:

The probability of event A is a weighted average of conditional probabilities, where the weights are the probabilities of the events on which the conditioning is done.

1.2 Learning by Example — Example 3.1 (Part 1)

Problem Statement

An insurance company classifies people into two categories:

- Accident prone
- Not accident prone

Given:

- $\Pr(\text{Accident within 1 year} \mid \text{Accident prone}) = 0.4$
- $\Pr(\text{Accident within 1 year} \mid \text{Not accident prone}) = 0.2$
- $\Pr(\text{Accident prone}) = 0.3$

Find the probability that a new policyholder will have an accident within one year.

Definitions

Let:

- A_1 : policyholder has an accident within one year
- A : policyholder is accident prone

Solution

Condition on whether the policyholder is accident prone:

$$\Pr(A_1) = \Pr(A_1 \mid A) \Pr(A) + \Pr(A_1 \mid A^c) \Pr(A^c)$$

Substitute values:

$$\begin{aligned}\Pr(A_1) &= (0.4)(0.3) + (0.2)(0.7) \\ &= 0.12 + 0.14 \\ &= 0.26\end{aligned}$$

1.3 Learning by Example — Example 3.1 (Part 2)

Problem Statement

Suppose a new policyholder **has had an accident within one year**.

What is the probability that the policyholder is accident prone?

Solution

We seek:

$$\Pr(A \mid A_1)$$

By definition of conditional probability:

$$\Pr(A \mid A_1) = \frac{\Pr(A \cap A_1)}{\Pr(A_1)}$$

Using multiplication rule:

$$\Pr(A \cap A_1) = \Pr(A) \Pr(A_1 \mid A)$$

Thus,

$$\begin{aligned}\Pr(A \mid A_1) &= \frac{(0.3)(0.4)}{0.26} \\ &= \frac{0.12}{0.26} \\ &= \frac{6}{13}\end{aligned}$$

1.4 Formal Introduction: Law of Total Probability

Suppose:

- B_1, B_2, \dots, B_n are mutually exclusive events
- $\bigcup_{i=1}^n B_i = B$

Exactly one of the events B_1, \dots, B_n must occur.

Writing:

$$A = \bigcup_{i=1}^n AB_i$$

and noting that the events AB_i are mutually exclusive, we have:

$$\Pr(A) = \sum_{i=1}^n \Pr(AB_i)$$

Using conditional probability:

$$\Pr(AB_i) = \Pr(A \mid B_i) \Pr(B_i)$$

Hence,

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

This is the **Law of Total Probability (Formula 3.4)**.

1.5 Bayes' Formula (Proposition 3.1)

Using:

$$\Pr(AB_i) = \Pr(B_i \mid A) \Pr(A)$$

and substituting into the law of total probability, we obtain:

$$\Pr(B_i \mid A) = \frac{\Pr(A \mid B_i) \Pr(B_i)}{\sum_{j=1}^n \Pr(A \mid B_j) \Pr(B_j)}$$

Where:

- $\Pr(B_i)$ is the **a priori probability**
 - $\Pr(B_i \mid A)$ is the **posterior probability**
-

1.6 Learning by Example — Example 3.2 (Cards Problem)

Setup

Three cards:

- One card: red-red (RR)
- One card: black-black (BB)
- One card: red-black (RB)

A card is randomly selected and placed down.

The upper side is observed to be red.

Find the probability that the other side is black.

Definitions

Let:

- RR, BB, RB : type of selected card
- R : upturned side is red

Solution

$$\Pr(RB \mid R) = \frac{\Pr(R \mid RB) \Pr(RB)}{\Pr(R \mid RR) \Pr(RR) + \Pr(R \mid RB) \Pr(RB) + \Pr(R \mid BB) \Pr(BB)}$$

Substitute values:

$$= \frac{(1/2)(1/3)}{(1)(1/3) + (1/2)(1/3) + (0)(1/3)}$$

Simplify:

$$\begin{aligned}
&= \frac{1/6}{1/3 + 1/6} \\
&= \frac{1/6}{1/2} \\
&= \frac{1}{3}
\end{aligned}$$

2. Random Variables

2.1 Motivation and Concept

When an experiment is performed, interest often lies in a function of the outcome rather than the outcome itself.

Examples:

- Tossing dice: sum of values
- Tossing coins: number of heads

These real-valued functions defined on the sample space are called **random variables**.

A random variable assigns:

- A real number to each outcome
 - Probabilities to possible values
-

2.2 Definition

A random variable X on a sample space Ω is a function:

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

that assigns to each sample point $\omega \in \Omega$ a real number $X(\omega)$.

In this lecture, attention is restricted to **discrete random variables**, whose values form a finite or countably infinite set.

2.3 Distribution of a Random Variable

Two key components:

1. The set of values the random variable can take
2. The probabilities with which it takes those values

For a value a ,

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

is an event, denoted $\{X = a\}$.

The probability:

$$\Pr[X = a]$$

is defined via the probability of the corresponding event.

The collection of these probabilities over all possible values constitutes the **distribution** of X .

2.4 Example — Tossing 3 Fair Coins

Let Y be the number of heads.

Possible values:

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

Probabilities:

$$\Pr(Y = 0) = \Pr(t, t, t) = \frac{1}{8}$$

$$\Pr(Y = 1) = \Pr(t, t, h), (t, h, t), (h, t, t) = \frac{3}{8}$$

$$\Pr(Y = 2) = \Pr(t, h, h), (h, t, h), (h, h, t) = \frac{3}{8}$$

$$\Pr(Y = 3) = \Pr(h, h, h) = \frac{1}{8}$$

Since Y must take one of these values:

$$\sum_{i=0}^3 \Pr(Y = i) = 1$$

3. Probability Mass Function (PMF)

3.1 Concept

A random variable that takes at most a countable number of possible values is called **discrete**.

Let X be a discrete random variable with range:

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

The function:

$$p_X(x_k) = \Pr(X = x_k)$$

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

Since X must take one of its possible values:

$$\sum_k p_X(x_k) = 1$$

3.2 Example — Two Independent Tosses of a Fair Coin

Sample space:

$$\Omega = \{(H, H), (H, T), (T, H), (T, T)\}$$

Let:

X = number of heads

PMF:

$$p_X(x) = \begin{cases} \frac{1}{4}, & x = 0 \text{ or } x = 2 \\ \frac{1}{2}, & x = 1 \\ 0, & \text{otherwise} \end{cases}$$

3.3 Example — Given PMF

Given:

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

Since:

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

we have:

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

Using:

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

Thus:

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

Hence:

$$\Pr(X = 0) = c = e^{-\lambda}$$

$$\Pr(X > 2) = 1 - [\Pr(X = 0) + \Pr(X = 1) + \Pr(X = 2)]$$

End of Lecture 5 Scribe



L5_A_S1 (2).pdf
PDF

Generate a lecture scribe intended strictly as exam-oriented reference material for CSE 400: Fundamentals of Probability in Computing – Lecture 5, covering Bayes' Theorem, Random Variables, and Probability Mass Function.

Use only the attached Lecture 5 slides/PDF and any explicitly relevant textbook content connected to this

lecture.

Do not introduce any material, definitions, notation, explanations, intuition, examples, assumptions, or results that are not explicitly present in the provided context.

While writing the scribe, reason step by step through the lecture content exactly as presented, including:

Definitions and notation

Assumptions and conditions

Statements of results (including the Law of Total Probability and Bayes' Formula)

Proofs or derivations, with all intermediate steps

Worked examples (including all numerical computations and conditional probability steps)

Ensure that the logical flow mirrors the lecture sequence, including:

Weighted average interpretation of conditional probabilities

Learning-by-example applications of Bayes' Theorem

Formal derivation using the Law of Total Probability

Definition and interpretation of random variables

Distributions of discrete random variables

Definition and properties of the probability mass function (PMF)

PMF-based examples exactly as shown in the lecture

The reasoning steps must be explicit and complete, so that a student preparing for a reading-based / closed-

notes exam can reconstruct all arguments and calculations solely from this scribe.

Maintain a formal academic tone, clear sectioning, and faithful adherence to the lecture's scope and structure. Do not add summaries, commentary, simplifications, alternative explanations, or material beyond what appears in the provided context.

Below is a **strictly exam-oriented lecture scribe** for

CSE 400: Fundamentals of Probability in Computing – Lecture 5, reconstructed **faithfully and exclusively** from the attached Lecture 5 slides/PDF.

No material beyond the provided context is introduced.

All definitions, notation, derivations, and examples follow the **exact lecture order and content**.

Source: Lecture 5 slides .

CSE 400: Fundamentals of Probability in Computing

Lecture 5 — Bayes' Theorem, Random Variables, and Probability Mass Function

1. Bayes' Theorem

1.1 Weighted Average of Conditional Probabilities

Let A and B be events.

An event A can be expressed as:

$$A = AB \cup AB^c$$

since for an outcome to be in A , it must either be in both A and B , or be in A but not in B .

The events AB and AB^c are **mutually exclusive**.

By **Axiom 3** of probability:

$$\Pr(A) = \Pr(AB) + \Pr(AB^c)$$

Using the definition of conditional probability:

$$\Pr(AB) = \Pr(A \mid B) \Pr(B)$$

$$\Pr(AB^c) = \Pr(A \mid B^c) \Pr(B^c)$$

Hence,

$$\Pr(A) = \Pr(A \mid B) \Pr(B) + \Pr(A \mid B^c) \Pr(B^c)$$

Since $\Pr(B^c) = 1 - \Pr(B)$,

$$\Pr(A) = \Pr(A \mid B) \Pr(B) + \Pr(A \mid B^c)[1 - \Pr(B)]$$

Interpretation (as stated in the lecture):

The probability of event A is a **weighted average of conditional probabilities**, where the weights are the probabilities of the events on which the conditioning is done.

1.2 Learning by Example — Example 3.1 (Part 1)

Problem Statement

An insurance company divides people into two classes:

- Accident prone
- Not accident prone

Given:

- Probability that an accident-prone person has an accident within one year: 0.4
- Probability that a non-accident-prone person has an accident within one year: 0.2
- 30% of the population is accident prone

Find the probability that a new policyholder will have an accident within one year.

Definitions

Let:

- A_1 : the policyholder has an accident within one year
- A : the policyholder is accident prone

Solution

Condition on whether the policyholder is accident prone:

$$\Pr(A_1) = \Pr(A_1 \mid A) \Pr(A) + \Pr(A_1 \mid A^c) \Pr(A^c)$$

Substitute the given values:

$$\begin{aligned}\Pr(A_1) &= (0.4)(0.3) + (0.2)(0.7) \\ &= 0.12 + 0.14 \\ &= 0.26\end{aligned}$$

1.3 Learning by Example — Example 3.1 (Part 2)

Problem Statement

Suppose a new policyholder **has an accident within one year**.

What is the probability that the policyholder is accident prone?

Solution

We seek:

$$\Pr(A \mid A_1)$$

By definition of conditional probability:

$$\Pr(A \mid A_1) = \frac{\Pr(A \cap A_1)}{\Pr(A_1)}$$

Using the multiplication rule:

$$\Pr(A \cap A_1) = \Pr(A) \Pr(A_1 \mid A)$$

Thus:

$$\begin{aligned}\Pr(A \mid A_1) &= \frac{(0.3)(0.4)}{0.26} \\ &= \frac{0.12}{0.26} \\ &= \frac{6}{13}\end{aligned}$$

2. Formal Introduction: Law of Total Probability and Bayes' Formula

2.1 Law of Total Probability

Suppose:

- B_1, B_2, \dots, B_n are mutually exclusive events
- $\bigcup_{i=1}^n B_i = B$

Exactly one of the events B_1, B_2, \dots, B_n must occur.

Write:

$$A = \bigcup_{i=1}^n AB_i$$

The events AB_i are mutually exclusive.

Hence:

$$\Pr(A) = \sum_{i=1}^n \Pr(AB_i)$$

Using conditional probability:

$$\Pr(AB_i) = \Pr(A \mid B_i) \Pr(B_i)$$

Therefore:

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

This result is known as the **Law of Total Probability (Formula 3.4)**.

2.2 Bayes' Formula (Proposition 3.1)

Using:

$$\Pr(AB_i) = \Pr(B_i \mid A) \Pr(A)$$

and substituting into the law of total probability, we obtain:

$$\Pr(B_i \mid A) = \frac{\Pr(A \mid B_i) \Pr(B_i)}{\sum_{j=1}^n \Pr(A \mid B_j) \Pr(B_j)}$$

Where:

- $\Pr(B_i)$ is the **a priori probability**
 - $\Pr(B_i \mid A)$ is the **posterior probability** of event B_i given A
-

2.3 Learning by Example — Example 3.2 (Cards)

Problem Statement

Three cards:

- One card: red–red (RR)
- One card: black–black (BB)
- One card: red–black (RB)

The cards are mixed, one card is randomly selected, and placed down.

If the **upper side is red**, find the probability that the **other side is black**.

Definitions

Let:

- RR, BB, RB : event that the chosen card is red–red, black–black, or red–black
- R : event that the upturned side is red

Solution

$$\Pr(RB \mid R) = \frac{\Pr(RB \cap R)}{\Pr(R)}$$

Using conditional probability:

$$\Pr(RB \mid R) = \frac{\Pr(R \mid RB) \Pr(RB)}{\Pr(R \mid RR) \Pr(RR) + \Pr(R \mid RB) \Pr(RB) + \Pr(R \mid BB) \Pr(BB)}$$

Substitute values:

$$= \frac{(1/2)(1/3)}{(1)(1/3) + (1/2)(1/3) + (0)(1/3)}$$

Simplify:

$$= \frac{1/6}{1/2} = \frac{1}{3}$$

3. Random Variables

3.1 Motivation and Concept

In many experiments, interest lies in a **function of the outcome** rather than the outcome itself.

Examples given:

- Dice tossing: sum of the two dice
- Coin flipping: number of heads

These real-valued functions defined on the sample space are called **random variables**.

- Values are determined by outcomes
 - Probabilities are assigned to possible values
-

3.2 Definition of a Random Variable

A random variable X on a sample space Ω is a function:

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

that assigns to each sample point $\omega \in \Omega$ a real number $X(\omega)$.

In this lecture, attention is restricted to **discrete random variables**, whose values form a finite or countably infinite subset of \mathbb{R} .

3.3 Distribution of a Discrete Random Variable

Two key components:

1. The set of values the random variable can take
2. The probabilities with which it takes those values

For a value a , the set:

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

is an event, denoted $\{X = a\}$.

The probability:

$$\Pr[X = a]$$

is obtained from the probability of this event.

The collection of these probabilities for all possible values of a is called the **distribution of X** .

3.4 Example — Tossing 3 Fair Coins

Let Y denote the number of heads.

Possible values:

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

Probabilities:

$$\Pr(Y = 0) = \Pr(t, t, t) = \frac{1}{8}$$

$$\Pr(Y = 1) = \Pr(t, t, h), (t, h, t), (h, t, t) = \frac{3}{8}$$

$$\Pr(Y = 2) = \Pr(t, h, h), (h, t, h), (h, h, t) = \frac{3}{8}$$

$$\Pr(Y = 3) = \Pr(h, h, h) = \frac{1}{8}$$

Since Y must take one of these values:

$$\sum_{i=0}^3 \Pr(Y = i) = 1$$

4. Probability Mass Function (PMF)

4.1 Concept

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

Let X be a discrete random variable with range:

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

The function:

$$p_X(x_k) = \Pr(X = x_k)$$

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

Since X must take one of the values x_k :

$$\sum_k p_X(x_k) = 1$$

4.2 Example — Two Independent Tosses of a Fair Coin

Sample space:

$$\Omega = \{(H, H), (H, T), (T, H), (T, T)\}$$

Let:

X = number of heads

PMF:

$$p_X(x) = \begin{cases} \frac{1}{4}, & x = 0 \text{ or } x = 2 \\ \frac{1}{2}, & x = 1 \\ 0, & \text{otherwise} \end{cases}$$

4.3 Example — Given PMF

Given:

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

where $\lambda > 0$.

Since:

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

We have:

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

Using:

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

Thus:

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

Hence:

$$\Pr(X = 0) = e^{-\lambda}$$

$$\Pr(X > 2) = 1 - [\Pr(X = 0) + \Pr(X = 1) + \Pr(X = 2)]$$

End of Lecture 5 — Exam-Oriented Scribe



L5_A_S1 (2).pdf
PDF

Generate a lecture scribe intended strictly as exam-oriented reference material for CSE 400: Fundamentals of Probability in Computing – Lecture 5, covering Bayes' Theorem, Random Variables, and Probability Mass Function.

Use only the attached Lecture 5 slides/PDF and any explicitly relevant textbook content connected to this lecture.

Do not introduce any material, definitions, notation,

explanations, intuition, examples, assumptions, or results that are not explicitly present in the provided context.

While writing the scribe, reason step by step through the lecture content exactly as presented, including:

Definitions and notation

Assumptions and conditions

Statements of results (including the Law of Total Probability and Bayes' Formula)

Proofs or derivations, with all intermediate steps

Worked examples (including all numerical computations and conditional probability steps)

Ensure that the logical flow mirrors the lecture sequence, including:

Weighted average interpretation of conditional probabilities

Learning-by-example applications of Bayes' Theorem

Formal derivation using the Law of Total Probability

Definition and interpretation of random variables

Distributions of discrete random variables

Definition and properties of the probability mass function (PMF)

PMF-based examples exactly as shown in the lecture

The reasoning steps must be explicit and complete, so that a student preparing for a reading-based / closed-notes exam can reconstruct all arguments and calculations solely from this scribe.

Maintain a formal academic tone, clear sectioning, and faithful adherence to the lecture's scope and structure. Do not add summaries, commentary, simplifications, alternative explanations, or material beyond what appears in the provided context.

Below is a **strictly exam-oriented lecture scribe** for

CSE 400: Fundamentals of Probability in Computing – Lecture 5, reconstructed **only** from the attached Lecture 5 slides/PDF.

The structure, notation, derivations, and examples **mirror the lecture sequence exactly**, with **no additions, omissions, or reinterpretations**.

Source: Lecture 5 slides .

CSE 400: Fundamentals of Probability in Computing

Lecture 5 — Bayes' Theorem, Random Variables, and Probability Mass Function

1. Bayes' Theorem

1.1 Weighted Average of Conditional Probabilities

Let A and B be events.

Event A can be expressed as

$$A = AB \cup AB^c$$

because, for an outcome to be in A , it must either be in both A and B , or be in A but not in B .

The events AB and AB^c are **mutually exclusive**.

By **Axiom 3** of probability,

$$\Pr(A) = \Pr(AB) + \Pr(AB^c)$$

Using the definition of conditional probability,

$$\Pr(AB) = \Pr(A \mid B) \Pr(B)$$

$$\Pr(AB^c) = \Pr(A \mid B^c) \Pr(B^c)$$

Hence,

$$\Pr(A) = \Pr(A \mid B) \Pr(B) + \Pr(A \mid B^c) \Pr(B^c)$$

Since $\Pr(B^c) = 1 - \Pr(B)$,

$$\Pr(A) = \Pr(A \mid B) \Pr(B) + \Pr(A \mid B^c)[1 - \Pr(B)]$$

Statement (as in the lecture):

The probability of event A is a **weighted average of the conditional probabilities**, with weights given by the probabilities of the events on which the conditioning is done.

2. Learning by Example (Bayes' Theorem)

Example 3.1 — Part 1

Problem Statement

An insurance company divides people into two classes:

- accident prone,
- not accident prone.

Given:

- An accident-prone person has an accident within a fixed 1-year period with probability 0.4.
- A non-accident-prone person has an accident within the same period with probability 0.2.
- 30% of the population is accident prone.

Find the probability that a new policyholder will have an accident within one year.

Definitions

Let:

- A_1 : event that the policyholder has an accident within one year,
- A : event that the policyholder is accident prone.

Solution

Condition on whether the policyholder is accident prone:

$$\Pr(A_1) = \Pr(A_1 \mid A) \Pr(A) + \Pr(A_1 \mid A^c) \Pr(A^c)$$

Substitute the given values:

$$\begin{aligned}\Pr(A_1) &= (0.4)(0.3) + (0.2)(0.7) \\ &= 0.12 + 0.14 \\ &= 0.26\end{aligned}$$

Example 3.1 — Part 2

Problem Statement

Suppose a new policyholder **has an accident within one year**.

What is the probability that the policyholder is accident prone?

Solution

The desired probability is

$$\Pr(A \mid A_1)$$

By the definition of conditional probability,

$$\Pr(A \mid A_1) = \frac{\Pr(A \cap A_1)}{\Pr(A_1)}$$

Using the multiplication rule,

$$\Pr(A \cap A_1) = \Pr(A) \Pr(A_1 \mid A)$$

Thus,

$$\begin{aligned}\Pr(A \mid A_1) &= \frac{(0.3)(0.4)}{0.26} \\ &= \frac{0.12}{0.26} \\ &= \frac{6}{13}\end{aligned}$$

3. Formal Introduction: Law of Total Probability

Suppose that:

- B_1, B_2, \dots, B_n are mutually exclusive events, and
- $\bigcup_{i=1}^n B_i = B$.

Exactly one of the events B_1, B_2, \dots, B_n must occur.

Write

$$A = \bigcup_{i=1}^n AB_i$$

The events $AB_i, i = 1, \dots, n$, are mutually exclusive.

Therefore,

$$\Pr(A) = \sum_{i=1}^n \Pr(AB_i)$$

Using conditional probability,

$$\Pr(AB_i) = \Pr(A \mid B_i) \Pr(B_i)$$

Hence,

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

This result is known as the **Law of Total Probability (Formula 3.4)**.

4. Bayes' Formula (Proposition 3.1)

Using

$$\Pr(AB_i) = \Pr(B_i \mid A) \Pr(A)$$

and substituting into the Law of Total Probability, we obtain

$$\Pr(B_i \mid A) = \frac{\Pr(A \mid B_i) \Pr(B_i)}{\sum_{j=1}^n \Pr(A \mid B_j) \Pr(B_j)}$$

Where:

- $\Pr(B_i)$ is the **a priori probability**, and
 - $\Pr(B_i \mid A)$ is the **posterior probability** of event B_i given A .
-

5. Learning by Example — Example 3.2 (Cards)

Problem Statement

Three cards:

- one card with both sides red (RR),
- one card with both sides black (BB),
- one card with one side red and one side black (RB).

The cards are mixed, one card is randomly selected and placed on the ground.

If the **upper side is red**, find the probability that the **other side is black**.

Definitions

Let:

- RR, BB, RB : events that the chosen card is red-red, black-black, or red-black,
- R : event that the upturned side is red.

Solution

$$\Pr(RB \mid R) = \frac{\Pr(RB \cap R)}{\Pr(R)}$$

Using conditional probabilities,

$$\Pr(RB \mid R) = \frac{\Pr(R \mid RB) \Pr(RB)}{\Pr(R \mid RR) \Pr(RR) + \Pr(R \mid RB) \Pr(RB) + \Pr(R \mid BB) \Pr(BB)}$$

Substitute values:

$$= \frac{(1/2)(1/3)}{(1)(1/3) + (1/2)(1/3) + (0)(1/3)}$$

Simplify:

$$= \frac{1/6}{1/2} = \frac{1}{3}$$

6. Random Variables

6.1 Motivation and Concept

In many experiments, interest lies in a **function of the outcome** rather than the outcome itself.

Examples:

- In dice tossing, the focus may be on the sum of the dice.
- In coin flipping, the focus may be on the number of heads.

These real-valued functions defined on the sample space are called **random variables**.

- Values are determined by experiment outcomes.
 - Probabilities are assigned to possible values.
-

6.2 Definition

A random variable X on a sample space Ω is a function

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

that assigns a real number $X(\omega)$ to each $\omega \in \Omega$.

In this lecture, attention is restricted to **discrete random variables**, whose values form a finite or countably infinite subset of \mathbb{R} .

6.3 Distribution of a Discrete Random Variable

For a value a , the set

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

is an event, abbreviated as $\{X = a\}$.

The probability of this event is

$$\Pr[X = a]$$

The collection of probabilities $\Pr[X = a]$ for all possible values a is called the **distribution of the random variable**.

6.4 Example — Tossing 3 Fair Coins

Let Y denote the number of heads.

Possible values:

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

Probabilities:

$$\Pr(Y = 0) = \Pr(t, t, t) = \frac{1}{8}$$

$$\Pr(Y = 1) = \Pr(t, t, h), (t, h, t), (h, t, t) = \frac{3}{8}$$

$$\Pr(Y = 2) = \Pr(t, h, h), (h, t, h), (h, h, t) = \frac{3}{8}$$

$$\Pr(Y = 3) = \Pr(h, h, h) = \frac{1}{8}$$

Since Y must take one of these values,

$$\sum_{i=0}^3 \Pr(Y = i) = 1$$

7. Probability Mass Function (PMF)

7.1 Definition

A random variable that can take **at most a countable number of possible values** is said to be **discrete**.

Let X be a discrete random variable with range

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

The function

$$p_X(x_k) = \Pr(X = x_k)$$

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

Since X must take one of the values x_k ,

$$\sum_k p_X(x_k) = 1$$

7.2 Example — Two Independent Tosses of a Fair Coin

Sample space:

$$\Omega = \{(H, H), (H, T), (T, H), (T, T)\}$$

Let

X = number of heads

PMF:

$$p_X(x) = \begin{cases} \frac{1}{4}, & x = 0 \text{ or } x = 2 \\ \frac{1}{2}, & x = 1 \\ 0, & \text{otherwise} \end{cases}$$

7.3 Example — Given PMF

Given:

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

where $\lambda > 0$.

Since

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

We have

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

Using

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

Thus,

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

Therefore,

$$\Pr(X = 0) = e^{-\lambda}$$

and

$$\Pr(X > 2) = 1 - [\Pr(X = 0) + \Pr(X = 1) + \Pr(X = 2)]$$

End of Lecture 5 — Exam-Oriented Scribe



L5_A_S1 (2).pdf
PDF

Generate a lecture scribe intended strictly as exam-oriented reference material for CSE 400: Fundamentals of Probability in Computing – Lecture 5, covering Bayes' Theorem, Random Variables, and Probability Mass Function.

Use only the attached Lecture 5 slides/PDF and any explicitly relevant textbook content connected to this lecture.

Do not introduce any material, definitions, notation, explanations, intuition, examples, assumptions, or results that are not explicitly present in the provided context.

While writing the scribe, reason step by step through the lecture content exactly as presented, including:

Definitions and notation

Assumptions and conditions

Statements of results (including the Law of Total Probability and Bayes' Formula)

Proofs or derivations, with all intermediate steps

Worked examples (including all numerical computations and conditional probability steps)

Ensure that the logical flow mirrors the lecture sequence, including:

Weighted average interpretation of conditional probabilities

Learning-by-example applications of Bayes' Theorem

Formal derivation using the Law of Total Probability

Definition and interpretation of random variables

Distributions of discrete random variables

Definition and properties of the probability mass function (PMF)

PMF-based examples exactly as shown in the lecture

The reasoning steps must be explicit and complete, so that a student preparing for a reading-based / closed-notes exam can reconstruct all arguments and calculations solely from this scribe.

Maintain a formal academic tone, clear sectioning, and faithful adherence to the lecture's scope and structure. Do not add summaries, commentary, simplifications, alternative explanations, or material beyond what appears in the provided context.

Below is a **strictly exam-oriented lecture scribe** for **CSE 400: Fundamentals of Probability in Computing – Lecture 5**, reconstructed **exclusively** from the attached Lecture 5 slides/PDF.

The order, notation, derivations, and worked examples **mirror the lecture exactly**, with **no additional material** introduced.

Source: Lecture 5 slides .

CSE 400: Fundamentals of Probability in Computing

Lecture 5 — Bayes' Theorem, Random Variables, and Probability Mass Function

1. Bayes' Theorem

1.1 Weighted Average of Conditional Probabilities

Let A and B be events.

We may express

$$A = AB \cup AB^c$$

since, for an outcome to be in A , it must either be in both A and B , or be in A but not in B .

The events AB and AB^c are mutually exclusive.

By **Axiom 3**,

$$\Pr(A) = \Pr(AB) + \Pr(AB^c).$$

Using conditional probability,

$$\Pr(AB) = \Pr(A \mid B) \Pr(B),$$

$$\Pr(AB^c) = \Pr(A \mid B^c) \Pr(B^c).$$

Hence,

$$\Pr(A) = \Pr(A \mid B) \Pr(B) + \Pr(A \mid B^c) \Pr(B^c) = \Pr(A \mid B) \Pr(B) + \Pr(A \mid B^c) [1 - \Pr(B)].$$

Statement (lecture):

The probability of event A is a **weighted average of conditional probabilities**, with weights given by the probabilities of the events on which the conditioning is done.

2. Learning by Example (Bayes' Theorem)

Example 3.1 — Part 1

Given

- Accident-prone person has an accident within one year with probability 0.4.
- Not accident-prone person has an accident within one year with probability 0.2.
- 30% of the population is accident prone.

Definitions

- A_1 : accident within one year.
- A : policyholder is accident prone.

Computation

Condition on whether the policyholder is accident prone:

$$\begin{aligned}\Pr(A_1) &= \Pr(A_1 \mid A) \Pr(A) + \Pr(A_1 \mid A^c) \Pr(A^c) \\ &= (0.4)(0.3) + (0.2)(0.7) = 0.12 + 0.14 = 0.26.\end{aligned}$$

Example 3.1 — Part 2

Question

Given that the policyholder has an accident within one year, find the probability that the policyholder is accident prone.

Computation

$$\begin{aligned}\Pr(A \mid A_1) &= \frac{\Pr(A \cap A_1)}{\Pr(A_1)} = \frac{\Pr(A) \Pr(A_1 \mid A)}{\Pr(A_1)} \\ &= \frac{(0.3)(0.4)}{0.26} = \frac{0.12}{0.26} = \frac{6}{13}.\end{aligned}$$

3. Formal Introduction: Law of Total Probability

Suppose B_1, B_2, \dots, B_n are mutually exclusive events such that

$$\bigcup_{i=1}^n B_i = B.$$

Exactly one of the events B_1, \dots, B_n must occur.

Write

$$A = \bigcup_{i=1}^n AB_i,$$

where the events AB_i are mutually exclusive.

Then

$$\Pr(A) = \sum_{i=1}^n \Pr(AB_i) = \sum_{i=1}^n \Pr(A \mid B_i) \Pr(B_i).$$

This is the **Law of Total Probability (Formula 3.4)**.

4. Bayes' Formula (Proposition 3.1)

Using

$$\Pr(AB_i) = \Pr(B_i \mid A) \Pr(A),$$

we obtain

$$\Pr(B_i \mid A) = \frac{\Pr(A \mid B_i) \Pr(B_i)}{\sum_{j=1}^n \Pr(A \mid B_j) \Pr(B_j)}.$$

Here,

- $\Pr(B_i)$ is the **a priori probability**,
 - $\Pr(B_i \mid A)$ is the **posterior probability** of B_i given A .
-

5. Learning by Example — Example 3.2 (Cards)

Setup

Three cards:

- red-red (RR),
- black-black (BB),
- red-black (RB).

One card is randomly selected and placed down. The upper side is red.

Definitions

- RR, BB, RB : type of selected card,
- R : upturned side is red.

Computation

$$\Pr(RB \mid R) = \frac{\Pr(RB \cap R)}{\Pr(R)} = \frac{\Pr(R \mid RB) \Pr(RB)}{\Pr(R \mid RR) \Pr(RR) + \Pr(R \mid RB) \Pr(RB) + \Pr(R \mid BB) \Pr(BB)}$$

Substitute values:

$$= \frac{(1/2)(1/3)}{(1)(1/3) + (1/2)(1/3) + (0)(1/3)} = \frac{1/6}{1/2} = \frac{1}{3}.$$

6. Random Variables

6.1 Motivation and Concept

Often we are interested in a function of the outcome rather than the outcome itself:

- Dice tossing: sum of dice.
- Coin flipping: number of heads.

These real-valued functions defined on the sample space are called **random variables**.

Values are determined by outcomes; probabilities are assigned to values.

6.2 Definition

A random variable X on a sample space Ω is a function

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

assigning a real number $X(\omega)$ to each $\omega \in \Omega$.

Attention is restricted to **discrete random variables**, whose values form a finite or countably infinite subset of \mathbb{R} .

6.3 Distribution of a Discrete Random Variable

For any value a ,

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

is an event, denoted $\{X = a\}$.

Its probability is $\Pr[X = a]$.

The collection of these probabilities over all possible values is the **distribution** of X .

6.4 Example — Tossing 3 Fair Coins

Let Y be the number of heads.

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

$$\Pr(Y = 0) = \Pr(t, t, t) = \frac{1}{8},$$

$$\Pr(Y = 1) = \Pr(t, t, h), (t, h, t), (h, t, t) = \frac{3}{8},$$

$$\Pr(Y = 2) = \Pr(t, h, h), (h, t, h), (h, h, t) = \frac{3}{8},$$

$$\Pr(Y = 3) = \Pr(h, h, h) = \frac{1}{8}.$$

Since Y must take one of these values,

$$\sum_{i=0}^3 \Pr(Y = i) = 1.$$

7. Probability Mass Function (PMF)

7.1 Definition

A random variable that can take at most a countable number of possible values is **discrete**.

Let X be a discrete random variable with range

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

The function

$$p_X(x_k) = \Pr(X = x_k)$$

is the **Probability Mass Function (PMF)** of X .

Since X must take one of the values x_k ,

$$\sum_k p_X(x_k) = 1.$$

7.2 Example — Two Independent Tosses of a Fair Coin

Sample space:

$$\Omega = \{(H, H), (H, T), (T, H), (T, T)\}.$$

Let X be the number of heads.

$$p_X(x) = \begin{cases} \frac{1}{4}, & x = 0 \text{ or } x = 2, \\ \frac{1}{2}, & x = 1, \\ 0, & \text{otherwise.} \end{cases}$$

7.3 Example — Given PMF

Given

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

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

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

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

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

Thus,

$$\Pr(X = 0) = e^{-\lambda},$$

$$\Pr(X > 2) = 1 - [\Pr(X = 0) + \Pr(X = 1) + \Pr(X = 2)].$$

End of Lecture 5 — Exam-Oriented Scribe