CSE400 – Fundamentals of Probability in Computing

Lecture 7: Expectation, CDFs, PDFs and Problem Solving

Dhaval Patel, PhD (Associate Professor)
January 27, 2025

Lecture Outline

- The Cumulative Distribution Function (CDF)
  - Definition
  - Properties
  - Examples
- The Probability Density Function (PDF)
  - Definition
  - Properties
  - Example
- Expectation of Random Variables (RVs)
  - Definition and Example
  - Expectation of a Function of a Random Variable
  - Linear Operation with Expectation
- n^{th} Moments and Central Moments of RVs: Variance, Skewness, Kurtosis

1 CDF and PDF: Intuition (Water Tank Analogy)

The lecture motivates the concept of CDF using a water tank analogy.

Interpretation

- The **height** h of water in the tank represents the value of a random variable.
- The **volume of water up to height** h, denoted $V(h)$, corresponds to the **cumulative probability** up to that value.

Volume Expression

The volume of water up to height h is:

$$V(h) = \int_0^h \pi R^2 \, dh = (\pi R^2)h$$

where πR^2 is constant for a cylindrical tank.

Analogy with PDF and CDF

- πR^2 is analogous to the **probability density function (PDF)** of a uniform distribution.
- The maximum volume of the tank is:

  $$V(H) = \pi R^2 H$$

  which is analogous to the total probability 1.

2 Cumulative Distribution Function (CDF)

2.1 Definition

Definition: The cumulative distribution function (CDF) of a random variable X is:

$$F_X(x) = \Pr(X \leq x), \qquad -\infty < x < \infty$$

Most of the information about the random experiment described by the random variable X is determined by the behavior of $F_X(x)$.

2.2 Properties of the CDF

The lecture states the following properties.

Property 1: Bounds

$$0 \leq F_X(x) \leq 1$$

Property 2: Limits at Infinity

$$F_X(-\infty) = 0, \qquad F_X(\infty) = 1$$

Property 3: Monotonicity

For $x_1 < x_2$,

$$F_X(x_1) \leq F_X(x_2)$$

Property 4: Interval Probability

For $x_1 < x_2$,

$$\Pr(x_1 < X \leq x_2) = F_X(x_2) - F_X(x_1)$$

3 CDF: Example #1 (Checking Validity of a CDF)

The lecture asks: **Find the valid CDF.**

Given candidate functions:

(a)
$$F_X(x) = \frac{1}{2} + \frac{1}{\pi}\tan^{-1}(x)$$

(b)
$$F_X(x) = \left[1 - e^{-x}\right] u(x)$$

(c)
$$F_X(x) = e^{-x^2}$$

(d)
$$F_X(x) = x^2 u(x)$$

Lecture Conclusion

From the lecture markings:

- (a) is a **valid CDF**.
- (b) is a **valid CDF**.
- (c) is **not** a valid CDF.
- (d) is **not** a valid CDF.

4 CDF: Example #2 (Using a Given CDF)

Suppose a random variable has a CDF:

$$F_X(x) = (1 - e^{-x})u(x)$$

Find the following quantities.

4.1 (1) $\Pr(X > 5)$

$$\Pr(X > 5) = 1 - \Pr(X \leq 5)$$

Using the CDF:

$$\Pr(X \leq 5) = F_X(5)$$

Thus:

$$\Pr(X > 5) = 1 - F_X(5)$$

Substitute:

$$F_X(5) = 1 - e^{-5}$$

Therefore:

$$\Pr(X > 5) = 1 - (1 - e^{-5}) = e^{-5}$$

4.2 (2) $\Pr(X < 5)$

From the CDF definition:

$$\Pr(X < 5) = \Pr(X \leq 5) = F_X(5)$$

Thus:

$$\Pr(X < 5) = 1 - e^{-5}$$

4.3 (3) $\Pr(3 < X < 7)$

Using the interval probability property:

$$\Pr(3 < X < 7) = F_X(7) - F_X(3)$$

Substitute:

$$F_X(7) = 1 - e^{-7}, \qquad F_X(3) = 1 - e^{-3}$$

Hence:

$$\begin{aligned}\Pr(3 < X < 7) &= (1 - e^{-7}) - (1 - e^{-3}) \\ &= -e^{-7} + e^{-3} \\ &= e^{-3} - e^{-7}\end{aligned}$$

4.4 (4) $\Pr(X > 5 \mid X < 7)$

Let:

$$A = \{X > 5\}, \qquad B = \{X < 7\}$$

Then:

$$\Pr(X > 5 \mid X < 7) = \Pr(A \mid B) = \frac{\Pr(A \cap B)}{\Pr(B)}$$

Now:

$$A \cap B = \{5 < X < 7\}$$

Thus:

$$\Pr(A \cap B) = \Pr(5 < X < 7) = F_X(7) - F_X(5)$$

Also:

$$\Pr(B) = \Pr(X < 7) = F_X(7)$$

Therefore:

$$\Pr(X > 5 \mid X < 7) = \frac{F_X(7) - F_X(5)}{F_X(7)}$$

5 Probability Density Function (PDF)

5.1 Definition (Continuous RVs)

The PDF of the random variable X evaluated at point x is:

$$f_X(x) = \lim_{\epsilon \to 0} \frac{\Pr(x < X < x + \epsilon)}{\epsilon}$$

5.2 PDF–CDF Relationship

For a continuous range:

$$\Pr(x < X < x + \epsilon) = F_X(x + \epsilon) - F_X(x)$$

Substituting into the PDF definition:

$$f_X(x) = \lim_{\epsilon \to 0} \frac{F_X(x + \epsilon) - F_X(x)}{\epsilon}$$

This is the derivative definition, so:

$$f_X(x) = \frac{dF_X(x)}{dx}$$

5.3 Key Conclusion

Hence, it is seen that:

- The **PDF** of a random variable is the **derivative** of its CDF.
- Conversely, the **CDF** of a random variable can be expressed as the **integral** of its PDF.