

CSE400 – Fundamentals of Probability in Computing

Lecture 7: Expectation, CDFs, PDFs, and Problem Solving

Source: Lecture slides by Dhaval Patel, PhD

January 27, 2025

1 The Cumulative Distribution Function (CDF)

1.1 Definition of the CDF

Let X be a random variable. The **Cumulative Distribution Function (CDF)** of X is defined as

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

For every real number x , the function $F_X(x)$ assigns the probability that the random variable X takes a value less than or equal to x . Thus, the CDF is a mapping from the real line to probabilities.

The lecture explicitly notes that **most of the information about the random experiment described by the random variable X is determined by the behavior of $F_X(x)$.**

1.2 Properties of the CDF

Property 1: Boundedness

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

Reasoning: Since $F_X(x) = \Pr(X \leq x)$ is a probability, by the axioms of probability it must lie between 0 and 1 inclusive.

Property 2: Limits at Infinity

$$F_X(-\infty) = 0, \quad F_X(\infty) = 1.$$

Reasoning:

- As $x \rightarrow -\infty$, the event $\{X \leq x\}$ becomes impossible, so its probability approaches 0.
- As $x \rightarrow \infty$, the event $\{X \leq x\}$ becomes certain, so its probability approaches 1.

Property 3: Monotonicity

For $x_1 < x_2$,

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

Reasoning: If $x_1 < x_2$, then the event $\{X \leq x_1\}$ is a subset of the event $\{X \leq x_2\}$. Since probabilities are monotone with respect to set inclusion, the inequality follows. Hence, a valid CDF must be non-decreasing.

Property 4: Probability Over an Interval

For $x_1 < x_2$,

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

Reasoning: The event $\{x_1 < X \leq x_2\}$ can be written as the difference between the events $\{X \leq x_2\}$ and $\{X \leq x_1\}$. Subtracting their probabilities yields the stated result.

1.3 Example: Validity of Candidate CDFs**Candidate 1**

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

- As $x \rightarrow -\infty$, $\tan^{-1}(x) \rightarrow -\frac{\pi}{2}$, hence $F_X(x) \rightarrow 0$.
- As $x \rightarrow \infty$, $\tan^{-1}(x) \rightarrow \frac{\pi}{2}$, hence $F_X(x) \rightarrow 1$.
- The function is monotone increasing.

Therefore, this function satisfies all CDF properties and is valid.

Candidate 2

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

where $u(x)$ is the unit step function.

- For $x < 0$, $u(x) = 0$, so $F_X(x) = 0$.
- For $x \geq 0$, $F_X(x) = 1 - e^{-x}$, which increases from 0 to 1.
- The function is non-decreasing and bounded between 0 and 1.

Hence, this is a valid CDF.

Candidate 3

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

- As $x \rightarrow -\infty$, $e^{-x^2} \rightarrow 0$.
- As $x \rightarrow \infty$, $e^{-x^2} \rightarrow 0$, not 1.

This violates the property $F_X(\infty) = 1$, so it is not a valid CDF.

2 The Probability Density Function (PDF)

2.1 Definition of the PDF

For a continuous random variable, the **Probability Density Function (PDF)** of X at point x is defined as

$$f_X(x) = \lim_{\epsilon \rightarrow 0} \frac{\Pr(x \leq X \leq x + \epsilon)}{\epsilon}.$$

This definition considers the probability that X lies in a small interval of width ϵ around x , normalized by the interval length.

2.2 Relationship Between PDF and CDF

For a continuous range,

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

Substituting into the PDF definition,

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

By the definition of the derivative,

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

2.3 Logical Dependency Between PDF and CDF

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

3 Expectation of Random Variables

3.1 Definition of Expectation

Expectation is a numerical summary associated with a random variable. It is defined with respect to the distribution of the random variable and is used to compute average or mean values.

3.2 Expectation of a Function of a Random Variable

The expectation of a function of X is computed by applying the same probabilistic weighting defined for X itself.

3.3 Linear Operations with Expectation

Expectation is linear under addition and scalar multiplication. This allows expectations of sums or scaled random variables to be computed systematically.

4 Moments and Central Moments of Random Variables

4.1 n -th Moments

The n -th moments are higher-order expectations that characterize the distribution of a random variable.

4.2 Central Moments

Central moments are defined relative to the mean and include:

- Variance
- Skewness
- Kurtosis

These quantities provide measures of spread, asymmetry, and tail behavior.

5 Summary

Lecture 7 establishes:

- The formal definition and properties of the CDF
- Methods to verify whether a function is a valid CDF
- The definition of the PDF and its derivation from the CDF
- The derivative–integral relationship between PDF and CDF
- The concept of expectation and higher-order moments