CSE 400 — Lecture 7 Scribe

Fundamentals of Probability in Computing

January 27, 2025

1 Cumulative Distribution Function (CDF)

1.1 Intuition: Water Tank Analogy

A water tank analogy is used to build intuition for the cumulative distribution function.

Let the height of water in the tank be denoted by h.

The volume of water up to height h is denoted by $V(h)$.

The height h is analogous to the value of a random variable.

The volume $V(h)$ is analogous to the cumulative probability up to that value.

Assume:

- The tank has constant radius R.
- Cross-sectional area is πR^2.

The volume up to height h is:

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

Key analogies stated explicitly in the lecture:

- πR^2 is analogous to a 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 total probability $= 1$.

Thus:

- CDF corresponds to accumulated volume.
- PDF corresponds to rate of accumulation.

1.2 Definition of the CDF

Definition. Let X be a random variable. The cumulative distribution function (CDF) of X is defined as

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

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

1.3 Properties of the CDF

1. **Bounds**

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

2. **Limits at Infinity**

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

3. **Monotonicity**

   For $x_1 < x_2$,

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

4. **Interval Probability**

   For $x_1 < x_2$,

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

1.4 Example 1: Validity of CDFs

Task. Determine whether each given function is a valid CDF.

(a) As $x \to -\infty$, $F_X(x) \to 0$. As $x \to \infty$, $F_X(x) \to 1$. The function is monotonically increasing.
Conclusion: Valid CDF.

(b) The unit step function ensures

$$F_X(x) = \begin{cases} 0, & x < 0 \\ 1 - e^{-x}, & x \geq 0 \end{cases}$$

All bounds, monotonicity, and limits are satisfied.
Conclusion: Valid CDF.

(c)

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

Since $\tan^{-1}(x) \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$, we have $F_X(x) \in (0, 1)$.
As $x \to \pm\infty$, boundary conditions are violated as stated in the lecture.
Conclusion: Not a valid CDF.

(d) The function grows unbounded for large x, violating boundedness.
Conclusion: Not a valid CDF.

1.5 Example 2: Computing Probabilities from a Given CDF

Given

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

(a)

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

(b)

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

(c)

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

(d) Let $A = \{X > 5\}$ and $B = \{X < 7\}$. Then

$$\Pr(A \mid B) = \frac{F_X(7) - F_X(5)}{F_X(7)}$$

2 Probability Density Function (PDF)

2.1 Definition via Limit

For a continuous random variable X, the PDF at x is defined as

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

2.2 Relationship Between PDF and CDF

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

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

2.3 Fundamental Relationship

The PDF of a random variable is the derivative of its CDF. The CDF is the integral of its PDF.

3 Expectation of Random Variables

The lecture outline lists this section, but no definitions, derivations, or examples are present in the provided slides.