ECE 302: Lecture 3.3 Cumulative Distribution Functions (discrete case)

Prof Stanley Chan

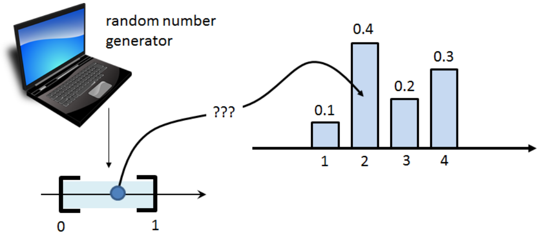
School of Electrical and Computer Engineering
Purdue University

Outline

Generating arbitrary random numbers

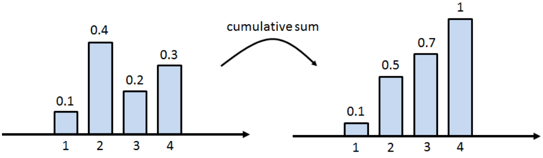
Question: How to generate random number from a PMF $p_X(k) = [0.1 \;\; 0.4 \;\; 0.2 \;\; 0.3]$?

Issue: A computer can only generate pre-define distributions, e.g., uniform.



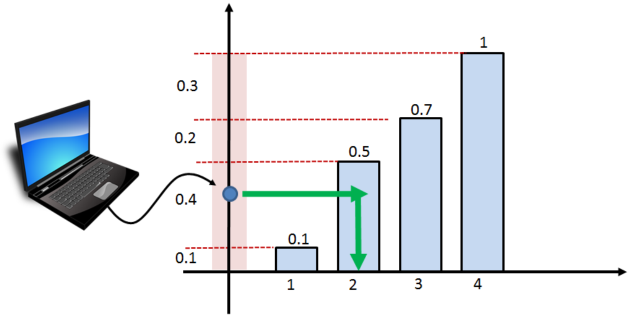
Generating arbitrary random numbers

Trick: Compute the **cumulative sum**.



Then what?

The power of cumulative sum



- Why study cumulative sum?
- Relationship to PMF?
- Properties of cumulative sum?

Outline

Cumulative Distribution Function

Definition

The **cumulative distribution function** (CDF) of a discrete random variable X is

$$F_X(x) \overset{\text{def}}{=} \mathbb{P}[X \leq x] = \sum_{x' \leq x} p_X(x').$$

Interpretation:

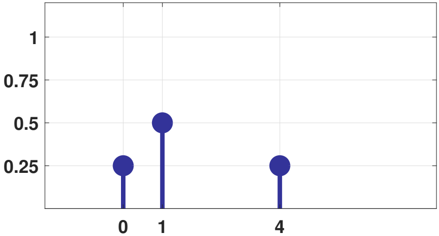
- CDF is the "integration" of PMF
- CDF is *well-defined* whereas PMF is not quite
- CDF works for both discrete and continuous random variables

Example 1

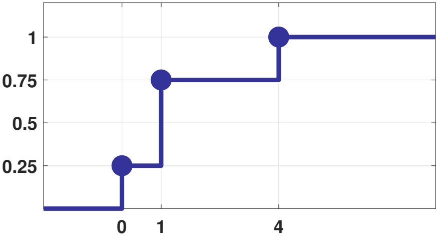
Example. Consider a random variable X with PMF

$$p_X(0) = \frac{1}{4}, \qquad p_X(1) = \frac{1}{2}, \qquad p_X(4) = \frac{1}{4}.$$

Find and sketch CDF.



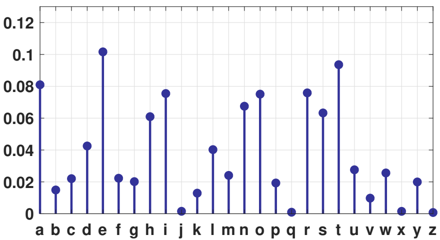
(a) PMF $p_X(k)$



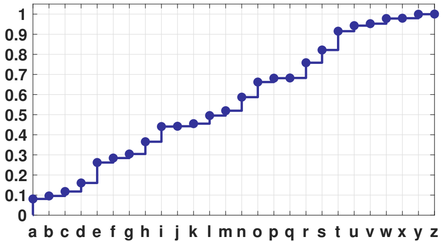
(b) CDF $F_X(k)$

Example 2

Example. English letters.



(a) PMF $p_X(k)$



(b) CDF $F_X(k)$

Outline

Properties of CDF

1. The CDF is a sequence of increasing
2. $F_X(+\infty) =$
3. $F_X(-\infty) =$
4. At positions where $p_X(x) > 0$, there is always a

Properties of CDF

5. The height of each jump is

6. The solid dot is always on the

Converting between PMF and CDF

Proposition

If X is a discrete random variable, then the PMF of X can be obtained from the CDF by

$$p_X(x_k) = F_X(x_k) - F_X(x_{k-1}). \tag{1}$$

A simpler version:

$$p_X(k) = F_X(k) - F_X(k-1). \tag{2}$$

Example

Example. If we are given the CDF

$$F_X(0) = \frac{1}{4}, \qquad F_X(1) = \frac{3}{4}, \qquad F_X(4) = 1,$$

how do we find the PMF? We know that the PMF will have non-negative values only at $x = 0, 1, 4$. For each of these x, we can show that

$$\begin{aligned} p_X(0) &= F_X(0) - F_X(-\infty) = \frac{1}{4} - 0 = \frac{1}{4}, \\ p_X(1) &= F_X(1) - F_X(0) = \frac{3}{4} - \frac{1}{4} = \frac{1}{2}, \\ p_X(4) &= F_X(4) - F_X(1) = 1 - \frac{3}{4} = \frac{1}{4}. \qquad \square \end{aligned}$$

Expressing PMF as delta functions

A PMF can technically be viewed as:

$$p_X(x) = \sum_{k \in X(\Omega)} \underbrace{p_X(k)}_{\text{PMF values}} \cdot \underbrace{\delta(x-k)}_{\text{delta function}}. \qquad (3)$$

For example,

$$p_X(x) = \frac{1}{4}\delta(x) + \frac{1}{2}\delta(x-1) + \frac{1}{4}\delta(x-2).$$

Then CDF is

$$F_X(x) = \frac{1}{4}u(x) + \frac{1}{2}u(x-1) + \frac{1}{4}u(x-2)$$

- $p_X(x)$ is not a function
- $F_X(x)$ is a function

Summary

- CDF = cum-sum of PDF
- CDF is always defined; some textbooks prefer to define PMF as derivatives of CDF
- CDF has several properties
- CDF of discrete random variables are stair case functions
- CDF can be used to generate random numbers of arbitrary distribution
- We will discuss CDF again when we talk about continuous random variables

Questions?