CS 107, Probability, Spring 2020
Lecture 21

Michael Poghosyan
mpoghosyan@aua.am

AUA

13 March 2020

Content

- The Cumulative Distribution Function: Examples
- Discrete Random Variables

Last Lecture ReCap

Let X be a rv.

- We want to be able to calculate Probabilities of the form

  $$\mathbb{P}(X \in A), \qquad A \subset \mathbb{R},$$

  in particular, we want to be able to calculate

  $$\mathbb{P}(a \leq X \leq b) = \mathbb{P}(X \in [a, b]), \qquad \forall a < b.$$

- The CDF of X is defined by

  $$F(x) = \mathbb{P}(X \leq x) = \mathbb{P}(X \in (-\infty, x]), \qquad x \in \mathbb{R}.$$

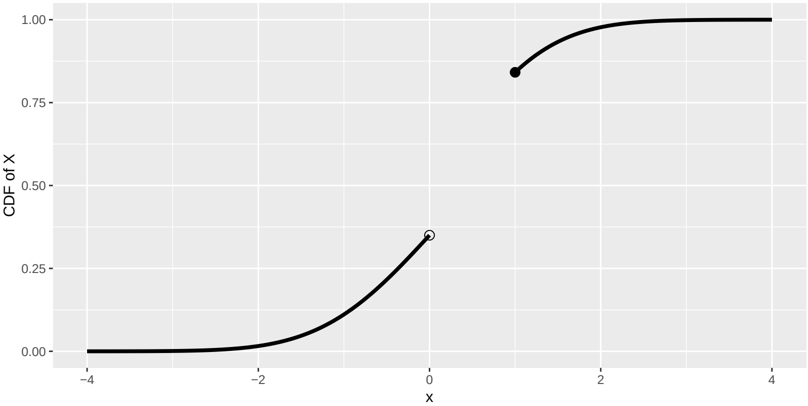
Last Lecture ReCap

- Characterization of CDFs: A function F is a CDF of some r.v. if and only if it satisfies the following properties:
  - $0 \leq F(x) \leq 1$, for any $x \in \mathbb{R}$;
  - $F(-\infty) = 0$ and $F(+\infty) = 1$;
  - F is an increasing function, i.e., if $x_1 \leq x_2$, then $F(x_1) \leq F(x_2)$;
  - F is right-continuous at every point, i.e. $F(x_0+) = F(x_0)$ at any $x_0 \in \mathbb{R}$
- If F is the CDF of r.v. X, then
  - $\mathbb{P}(X = a) = F(a) - F(a-)$;
  - $\mathbb{P}(a < X \leq b) = F(b) - F(a)$:

Examples

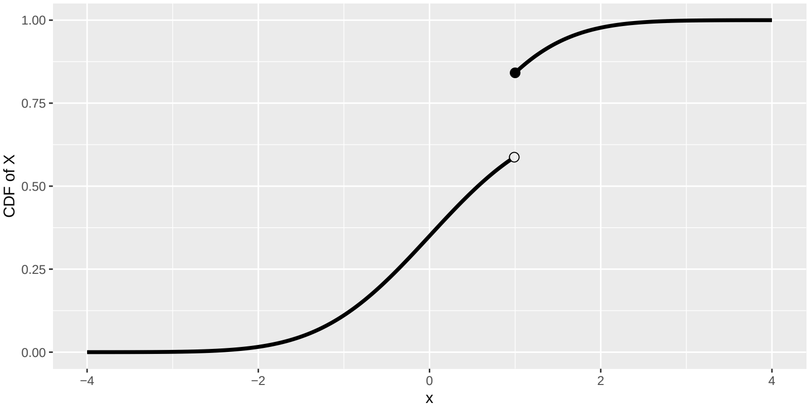
Reading Probabilities from the graph of CDF

Example 21.1: Below is the graph of some function $F(x)$. Is it the CDF of some r.v. X? If yes, calculate $\mathbb{P}(X = 0.2)$, $\mathbb{P}(X = 1)$, $\mathbb{P}(0 < X \leq 3)$, $\mathbb{P}(X \leq 1)$, $\mathbb{P}(X < 1)$, $\mathbb{P}(X > 0)$:



Reading Probabilities from the graph of CDF

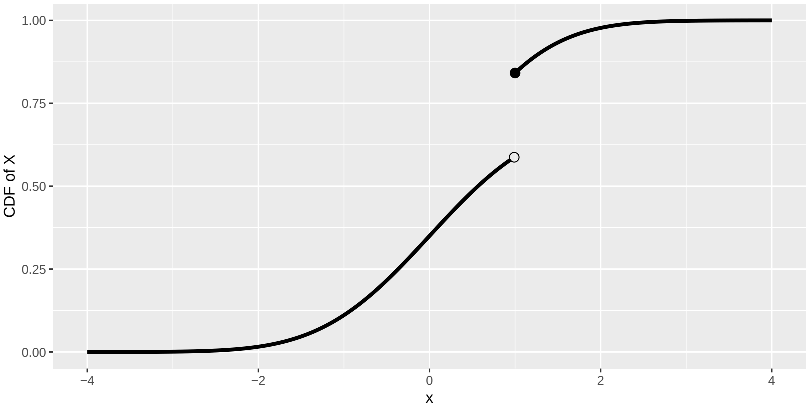
Example 21.2: Below is the graph of some function $F(x)$.



a. Is it the CDF of some r.v. X?

Reading Probabilities from the graph of CDF

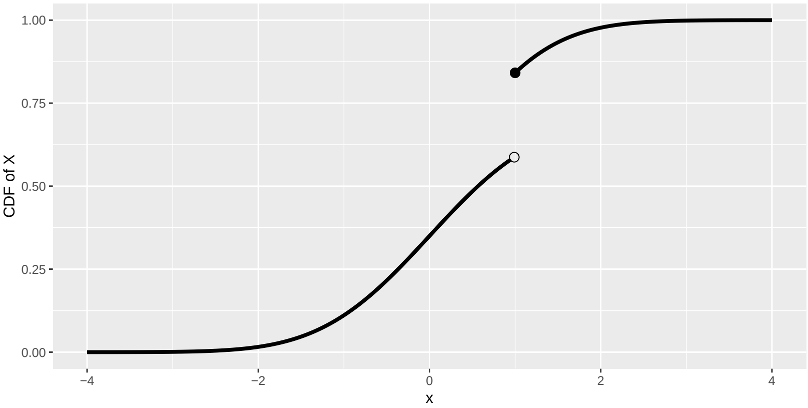
Example 21.2: Below is the graph of some function $F(x)$.



b. Calculate $\mathbb{P}(X = 0.2)$;

Reading Probabilities from the graph of CDF

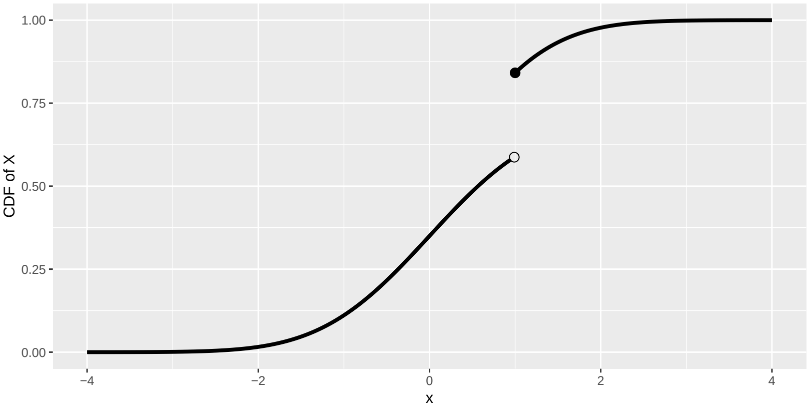
Example 21.2: Below is the graph of some function $F(x)$.



c. Calculate $\mathbb{P}(X = 1)$;

Reading Probabilities from the graph of CDF

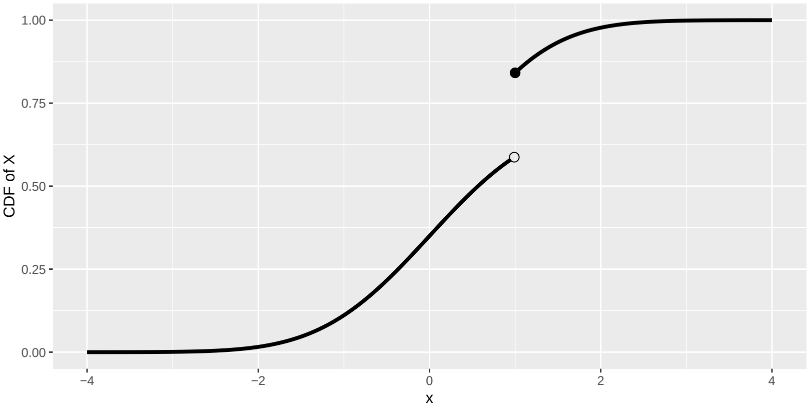
Example 21.2: Below is the graph of some function $F(x)$.



d. Calculate $\mathbb{P}(0 < X \leq 3)$

Reading Probabilities from the graph of CDF

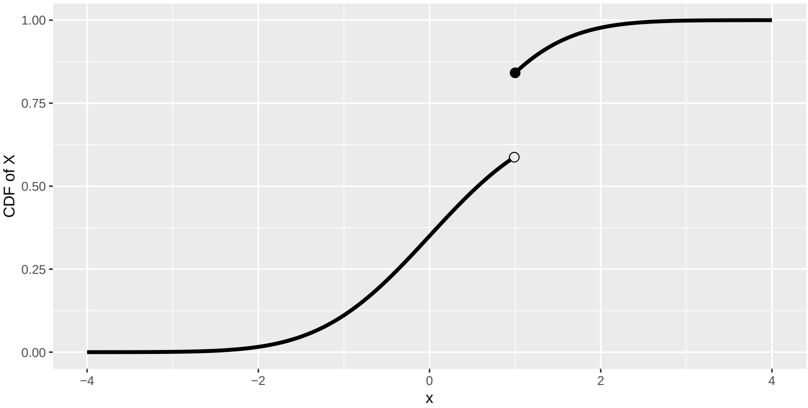
Example 21.2: Below is the graph of some function $F(x)$.



e. Calculate $\mathbb{P}(X \leq 1)$;

Reading Probabilities from the graph of CDF

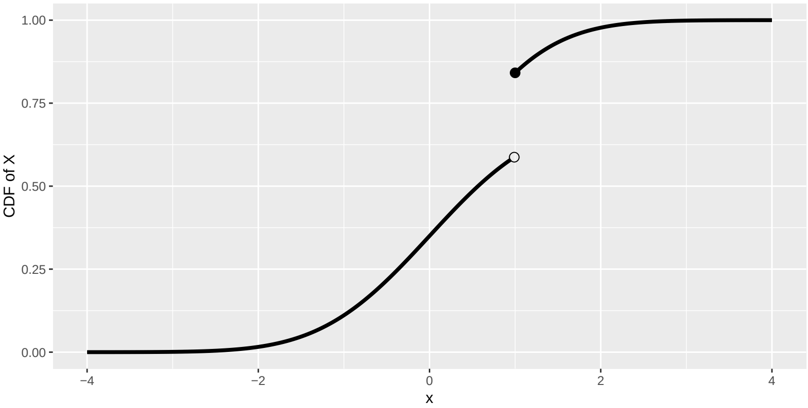
Example 21.2: Below is the graph of some function $F(x)$.



f. Calculate $\mathbb{P}(X < 1)$;

Reading Probabilities from the graph of CDF

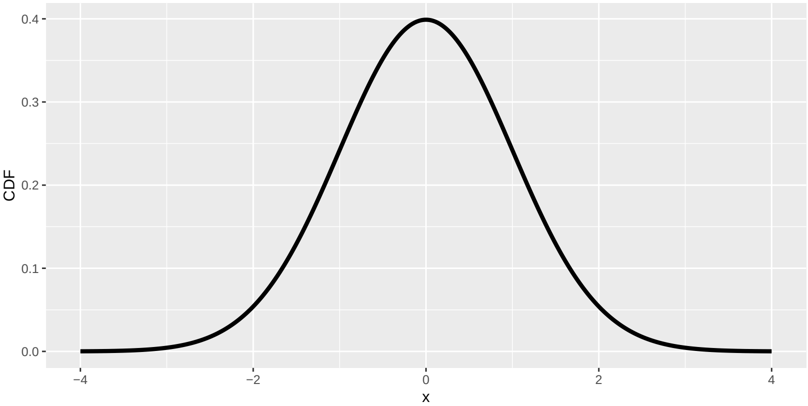
Example 21.2: Below is the graph of some function $F(x)$.



g. Calculate $\mathbb{P}(X > 0)$.

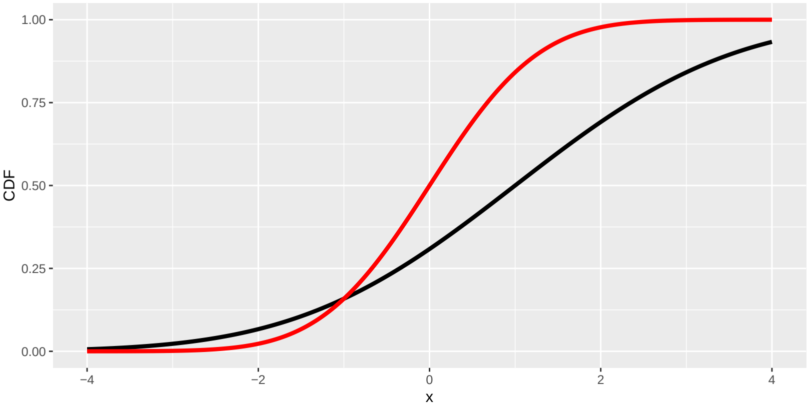
Reading Probabilities from the graph of CDF

Exercise 21.3: Is the following a graph of the CDF of some r.v. X? If yes, calculate $\mathbb{P}(X = 1)$, $\mathbb{P}(0 < X \leq 3)$, $\mathbb{P}(X \leq 1)$, $\mathbb{P}(X > 0)$:



Reading Probabilities from the graph of CDF

Example 21.4: Below you can find graphs of 2 CDFs: Red is for the r.v. X, and Black is for Y. Which one is larger: $\mathbb{P}(1 < X < 4)$ or $\mathbb{P}(1 < Y < 4)$?



Calculating Probabilities using CDF

Example 21.5: Assume X is a r.v. with the CDF

$$F(x) = \sigma(x) = sigmoid(x) = \frac{1}{1+e^{-x}}, \qquad x \in \mathbb{R}.$$

Calculate the Probabilities:

- $\mathbb{P}(X = 1.42)$;
- $\mathbb{P}(X \leq 0)$;
- $\mathbb{P}(X < 0)$;
- $\mathbb{P}(-2 < X < 2)$;
- $\mathbb{P}(X \in (-1, 0) \cup [3, 12])$.

Solution: OTB

Reading the info about the r.v. from CDF

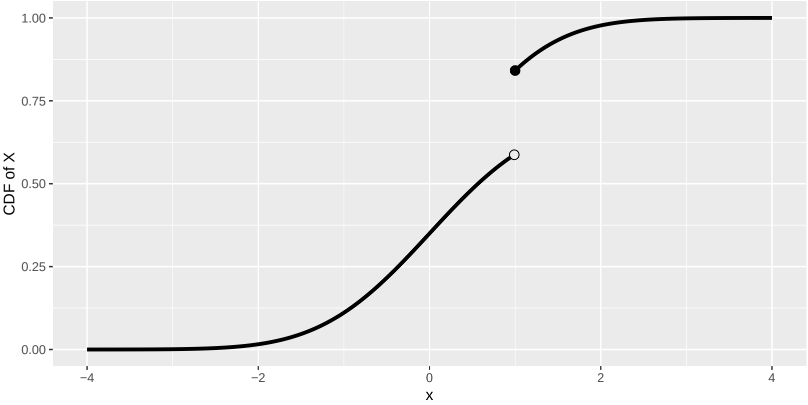
Example 21.6: Describe the r.v. X with the CDF

$$F(x) = \begin{cases} 1, & x \geq 0; \\ 0, & x < 0. \end{cases}$$

Solution: OTB

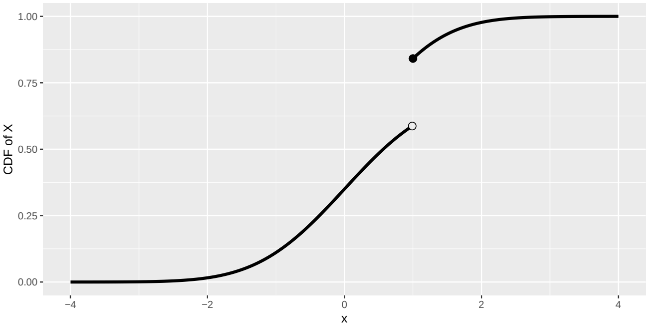
Reading the info about the r.v. from CDF

Example 21.7: For the r.v. X given through its CDF below, describe in some way which values are more probable, which values are less probable etc.



Remark

Remark: In fact, we will not consider r.v. X with the following type of CDFs (although they are completely legitimate):



Either we will deal with **piecewise constant** CDFs (CDFs of Discrete r.v.s) or with **continuous** CDFs (CDFs of Continuous r.v.s).

Discrete R.V.s

Random Variables Universe

So far we considered different examples of r.v.s. Some of them were taking discrete values, like the number of customers, and some were taking any real values from some interval, like the lifetime of a usb memory stick. In fact, the totality of all r.v.s can be classified into 4 groups:

- Discrete r.v.s;
- Continuous r.v.s;
- Singular r.v.s;
- Mixtures (sums) of the above r.v.s.

In our course we will consider only Discrete and Continuous r.v.s, leaving Singular ones and mixtures to advanced courses. In fact, any practical r.v. you will meet and use, will be either Discrete or Continuous (or, possibly, their mixture).

Discrete Random Variables

First, let us talk about Discrete r.v.s.
Assume X is a r.v. defined on Ω, i.e., $X : \Omega \to \mathbb{R}$.

Discrete Random Variable

We say that the r.v. X is **Discrete**, if the Range of X,

$$Range(X) = \{X(\omega) : \omega \in \Omega\} = \text{The set of all possible values of } X$$

is finite or countably infinite.

So if X is Discrete, then the Range of X can be written as

$$Range(X) = \{x_1, x_2, x_3, ...\},$$

where the set on the Right Hand Side (RHS) can be also finite.

Discrete Random Variables: Examples

Example 21.8: Which of the following r.v.s are Discrete?

- $X =$ the number of Heads when tossing 5 coins (or a coin 5 times).
- $Y =$ the number of children in the randomly chosen Armenian family, so if Ω is the set of all Armenian families, $\omega \in \Omega$ is a family, and $Y(\omega)$ is the number of children in the family ω.
- $Z =$ the number of car accidents today in Yerevan;
- $T =$ the claim size for today for one of our insurance companies;
- $U =$ the DJIA index closing value today;
- $V =$ the number of page clicks/search queries in Google;
- $W =$ the number of grammatical errors in my lecture slides;
- Can you give some more?

Mostly Discrete rvs are counting something: the number of goals, the number of webpage clicks, the number of app downloads, the number of infected persons, the number of cars that will pass through Baghramyan street today, the number of successes in the repeated trials experiment, ... But not only:

Example 21.9: Let us consider the Darts game. Let X be the points I will get when throwing a missile. Here the Sample Space Ω is the Dartboard, and an outcome is the point where my missile will hit the board. Ω itself is infinite (non-countable infinite). But the Range of X is finite:

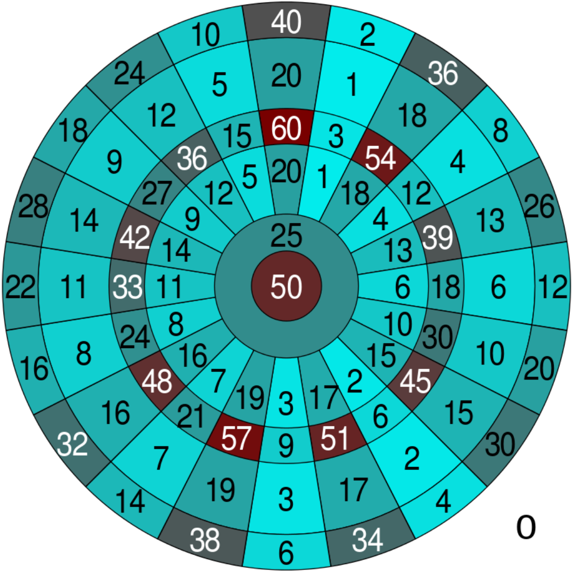


Figure: Darts game scores. By Cmglee, Wikipedia

Discrete Random Variables

In general,

- If Ω is Discrete (i.e., finite or countably infinite), then X will be Discrete;
- If Ω is not Discrete, then X CAN BE Discrete, as in the example with the Darts game.

OK, from this point on, for this lecture, we are forgetting about Ω, and considering only the Range of X, the set of all values of X.

Characterization of Discrete R.V.s

Now, assume X is a discrete r.v. with the Range (finite or countably infinite)

$$Range(X) = \{x_1, x_2, x_3, ...\}.$$

Then the function $p(x) = \mathbb{P}(X = x)$, $x \in \mathbb{R}$ is called the Probability Mass Function (PMF) of X.
In fact, PMF is non-zero only at points $x = x_k$, and we denote $p_k = \mathbb{P}(X = x_k)$, $k = 1, 2, 3, ...$. And in this case we write PMF in the table form:

Values of X	x_1	x_2	x_3	...
$\mathbb{P}(X = x)$	p_1	p_2	p_3	...

We will usually write PMF in the table form.

Characterization of PMF

Clearly, if X is a discrete r.v. with the PMF (in the table form):

Values of X	x_1	x_2	x_3	...
$\mathbb{P}(X = x)$	p_1	p_2	p_3	...

then

- $p_k \geq 0, \; k = 1, 2, 3, \ldots$
- $\sum_k p_k = 1$

Question: What is the difference between Discrete Probability Models and Discrete r.v.s?

PMF: Example

Example 21.10: Assume we have a box full of 10 red and 20 blue balls. We are picking at random a ball three times, with replacement. Let X be the number of red balls among this three balls:

$$X = \#(\text{red balls}).$$

Construct the PMF of X.
Solution: OTB

Question: What about the case when we are picking without replacements?

PMF: Example

Example 21.11: Again assume we have a box full of 10 red and 20 blue balls. We are picking at random a ball three times, with replacement. Now, let Y be the number of blue balls drawn minus the number of red balls drawn:

$$Y = \#(\text{blue balls}) - \#(\text{red balls}).$$

Construct the PMF of Y.
Solution: OTB

Question: What about the case when we are picking without replacements?

PMF and Calculation of Probabilities

Recall that, given a r.v. X, our aim was to be able to calculate the Probabilities of the type

$$\mathbb{P}(X \in A), \qquad A \subset \mathbb{R}.$$

Now, in the case we have a Discrete r.v. X given by its PMF:

Values of X	x_1	x_2	x_3	...
$\mathbb{P}(X = x)$	p_1	p_2	p_3	...

we can calculate

$$\mathbb{P}(X \in A) = \sum_{x_k \in A} p_k.$$

For example,

$$\mathbb{P}(a \le X \le b) = \sum_{a \le x_k \le b} p_k.$$

PMF and CDF for a Discrete r.v.

In particular, if the Discrete r.v. X is given by its PMF

Values of X	x_1	x_2	x_3	...
$\mathbb{P}(X = x)$	p_1	p_2	p_3	...

and if $F(x)$ is the CDF of X, then:

$$F(x) = \mathbb{P}(X \le x) = \sum_{x_k \le x} p_k.$$

Example

Example 21.12: Assume I am picking a point at random from $[0,5]$, uniformly. If my point is ≤ 2, I am paying $\$5$, if it is in $(2,3]$, then I am receiving $\$3$, in all other cases I am receiving $\$10$. Let X be my reward after playing this once. Construct the PMF and CDF for X.

Solution: OTB

Example:

Example 21.13: Let X is the number of heads shown when tossing a fair coin 4 times. Then X is a r.v. with the PMF:

Values of X					
$\mathbb{P}(X = x)$					

Now,

- Calculate $\mathbb{P}(X = 2)$;
- Calculate $\mathbb{P}(X \leq 2.5)$;
- Calculate $\mathbb{P}(X \geq 1)$;
- Graph the PMF of X;
- Graph the CDF of X.

Example:

Example 21.14: Assume we are tossing a fair coin until it will land on Heads. Let X be the number of tosses.

- Construct the PMF of X;
- What is the Probability that we will have the first Head on the third toss?
- What is the Probability that the first 2 tosses will be Tails?
- What is the Probability that we will do even number of tosses?
- Graph the CDF of X.

The Graph of CDF of a Discrete r.v

In general, if the r.v. X is Discrete, given through its PMF:

Values of X	x_1	x_2	x_3	...
$\mathbb{P}(X = x)$	p_1	p_2	p_3	...

then the graph of the CDF $F(x)$ of X will be

- a piecewise constant function (step function);
- the jump points will be the points x_k, $k = 1, 2, ...$;
- the jump size at the point x_k will be p_k;
- F will be continuous from the right.

Graph on the board, pleeeese!

PMF from the CDF

Conversely, if we have the CDF $F(x)$ of a Discrete r.v. X, then:

- We can calculate the PMF by:

$$\mathbb{P}(X = x) = F(x) - F(x-) \qquad \forall x \in \mathbb{R}$$

In fact, this will be non-zero only at x_k, the values from the $Range(X)$:

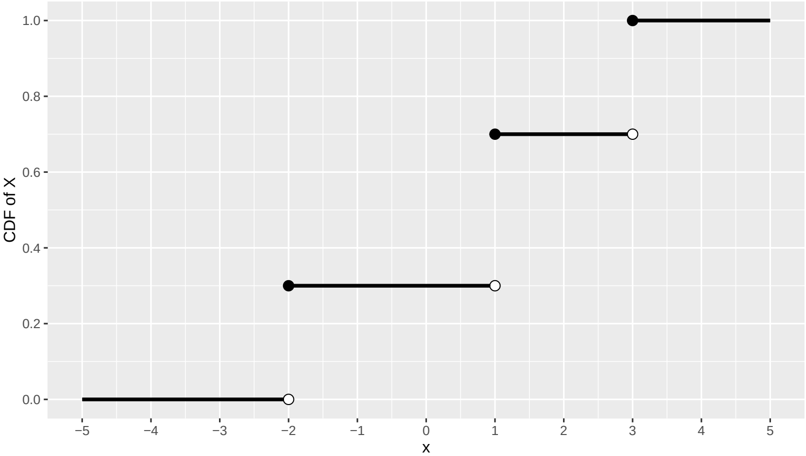
$$\mathbb{P}(X = x_k) = F(x_k) - F(x_k-), \qquad k = 1, 2, \ldots;$$

- Graphically, CDF will be a piecewise constant function
- Then the possible values x_k of X will be the jump points of F;
- The Probability $\mathbb{P}(X = x_k)$ will be the jump size at x_k

Give the graph, pleeeeese!

Reading Discrete r.v. from the CDF graph

Example 21.15: Below is the graph of the CDF of some r.v. X. Find the PMF of X.



Discrete r.v.s, summary

In this lecture we have learned that:

- R.v. X is Discrete, if the set of possible values of X is finite or countably infinite;
- We have two complete characteristics of a Discrete r.v. X: its CDF and PMF; having either of these, we can calculate $\mathbb{P}(X \in A)$ for any $A \subset \mathbb{R}$;
- The graph of the CDF of a Discrete r.v. is a piecewise constant (step) function; the jump (discontinuity) points of CDF are exactly the values of X, and the jump size at some point is the Probability of taking that particular value.

We will describe later some very important Discrete r.v.s: Bernoulli, Binomial, Geometric, Poisson, etc.