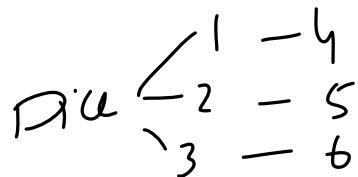


## (Algebraic vs Random) variable

- What are Algebraic Variables?

In Algebra a variable, like  $x$ , is an unknown value

$$x + 5 = 10 \Rightarrow x = 5$$



- What are Random Variables in Stats and Probability?

A Random Variable is a set of possible values from a random experiment.



$$\begin{array}{l} X = \{1, 0\} \\ \text{randomly} \\ H=1 \quad T=0 \end{array} \qquad \begin{array}{l} Y = \{1, 2, 3, 4, 5, 6\} \\ \text{sample space} \end{array}$$

$X \quad Y \quad Z$        $x, y, z$

- Types of Random Variables?

## Type of random variable

Discrete  
RV

$$\begin{cases} \{H, T\} \\ \{1, 2, 3, 4, 5, 6\} \end{cases}$$

↑ ↑ ↑

Continuous  
RV

$$X = \{0, 10\}$$

একটি class এ গোপ্য ক্ষেত্র  
বা কানুনী। e.g.  
(0~4) কৃষ্ণ ধা তিনি  
২৫ বাবু। কুকুর সতা  
 $\{1, 2, 3, 4, 5, 6\}$  কো ফিক্স

value কৃতি কৃতি পর্যন্ত value  
রয়ে দে।

গুরুত্ব List যোগান, random variable @ outcome শব্দ  
এবং তার probability খোজাবে।

### 1. What are Probability Distributions?

coin  
toss

A probability distribution is a list of all of the possible outcomes of a random variable along with their corresponding probability values.

coin toss	1 (H)	0 (T)
probab	$\frac{1}{2}$	$\frac{1}{2}$

dice  
খেল

1	2	3	4	5	6
$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

1	1	2	3	4	5	6	7
2	3	4	5	6	7	8	
3	4	5	6	7	8	9	
4	5	6	7	8	9	10	
5	6	7	8	9	10	11	
6	7	8	9	10	11	12	

Problem with Distribution!

In many scenarios, the number of outcomes can be much larger and hence a table would be tedious to write down. Worse still, the number of possible outcomes could be infinite, in which case, good luck writing a table for that.

Example - Height of people, Rolling 10 dice together

Solution - Function?

What if we use a mathematical function to model the relationship between outcome and probability?

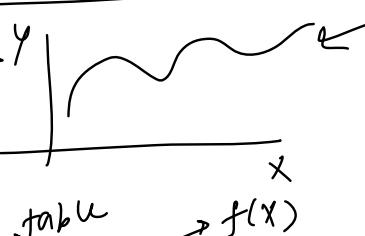
x → outcome

y → prob

1 (1/6)	2	3	4	5	6
$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

$$y = f(x)$$

PDF



Note - A lot of time Probability Distribution and Probability Distribution Functions are

2 dice →

2 3 4 5 6 7 8 9

২টা ইক্স নিম্নলিখিত  
এবং যোগ (২, ১) পথে  
ইয়েস্ট Probabilit distribution

2 → $\frac{1}{36}$
3 → $\frac{2}{36}$
4 → $\frac{3}{36}$
5 → $\frac{4}{36}$
6 → $\frac{5}{36}$

7 →  $\frac{6}{36}$

8 →  $\frac{5}{36}$

9 →  $\frac{4}{36}$

10 →  $\frac{3}{36}$

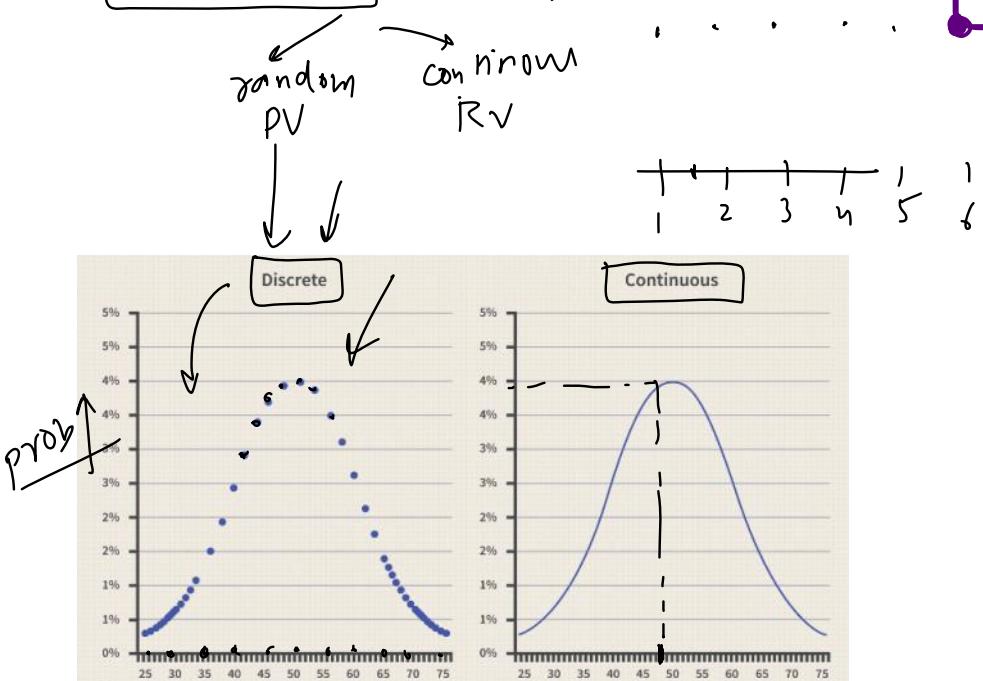
11 →  $\frac{2}{36}$

12 →  $\frac{1}{36}$

এখন, কথা রচনা, random vari.  
গুরুত্ব output এবং আনন্দ অন্তর  
যোগ করা (প্রাথমিক ১২টি)  
ক্ষেত্র - গোটা random  
experiment এ, তখন আমরা  
কি করবে? । কারণ, প্রযুক্তি  
Probability distribution  
বাসাগ সহজ হ। তবে,  
আশা হা, Probability distribution  
function ব্যবহার করব।

Note - A lot of time Probability Distribution and Probability Distribution Functions are used interchangeably.

### 1. Types of Probability Distributions (PDF)

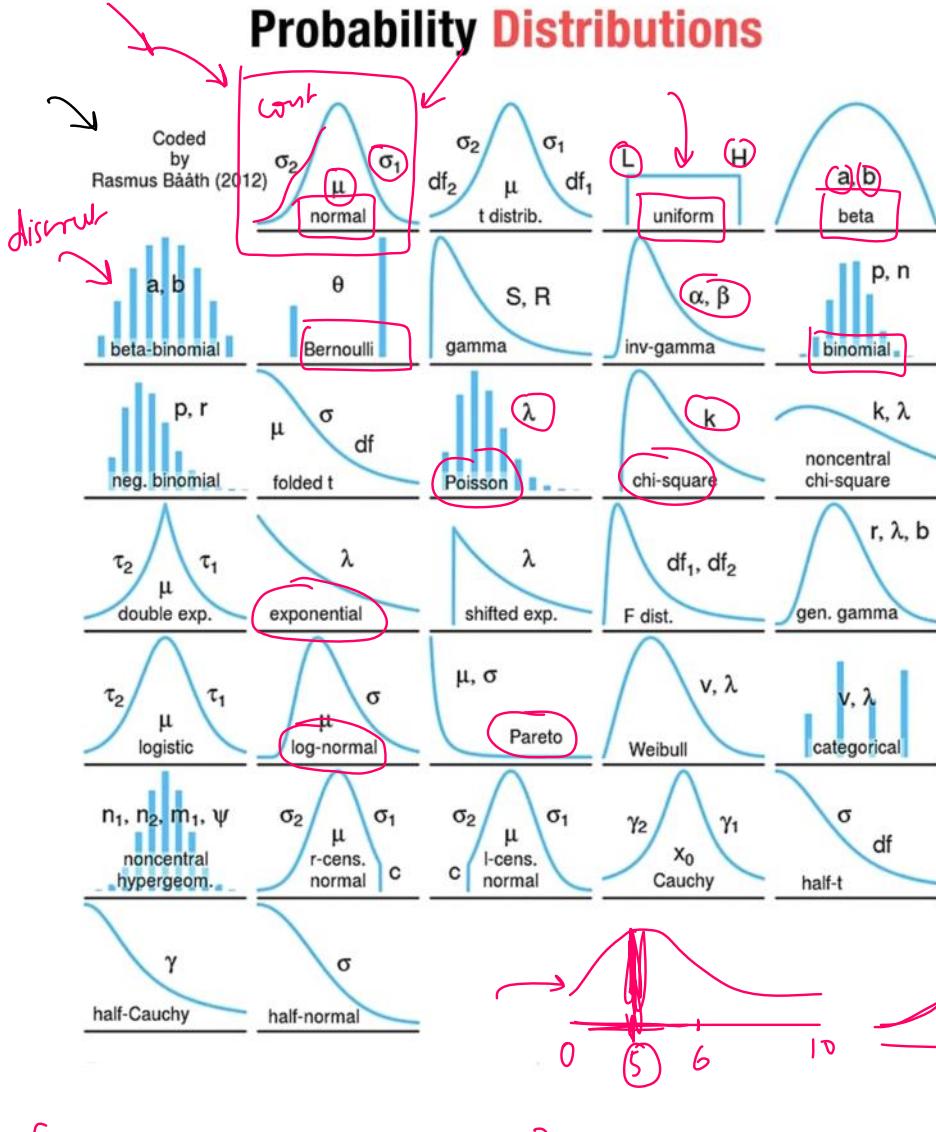


आनंद याप्तिः, Probability distribution & probability function सुरीदा एवं प्रार्थ्य-प्रयोग इति।



आमदृष्ट चर्तुणां घटे यान्हा वात्तन इस्तमालमें input and output निम्न, आमदृष्ट आनंद भवते, probability distribution function वात्तने गाहि। निम्न कठमाले famous probability distribution function देखा रही।

# Probability Distributions



## Why are Probability Distributions important?

- Gives an idea about the shape/distribution of the data.
- And if our data follows a famous distribution then we automatically know a lot about the data.

## A note on Parameters (PDF)

Parameters in probability distributions are numerical values that determine the shape, location, and scale of the distribution.

Different probability distributions have different sets of parameters that determine their shape and characteristics, and understanding these parameters is essential in statistical analysis and inference.

Normal probability - distribution  
curve ( $\sigma_1, \mu, \sigma_2$ )  
parametrized.

अब जानिए कि कौन कौन सी परामित्र हैं विभिन्न प्रकार के PDF

# ✓ [Probability Distribution Functions] → PDF

15 March 2023 20:08

{ A probability distribution function (PDF) is a mathematical function that describes the probability of obtaining different values of a random variable in a particular probability distribution.

$$Y = f(x)$$

Random discrete variable

Probability Mass function (PMF)

Probability Density function (PDF)

Random continuous variable

CPF

Cumulative Distribution function

PMF → CPF

PDF → CDF

PMF or PDF  
মিলে CDF আসে

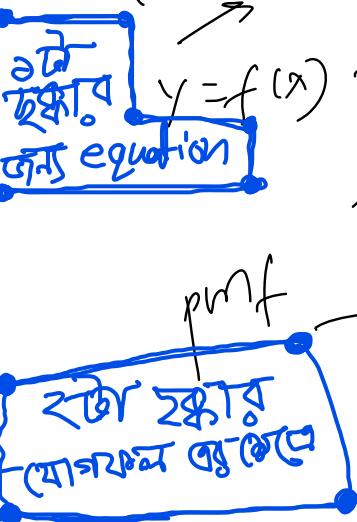
## Probability Mass Function (PMF)

15 March 2023 15:25

PMF stands for Probability Mass Function. It is a mathematical function that describes the probability distribution of a **discrete random variable**.

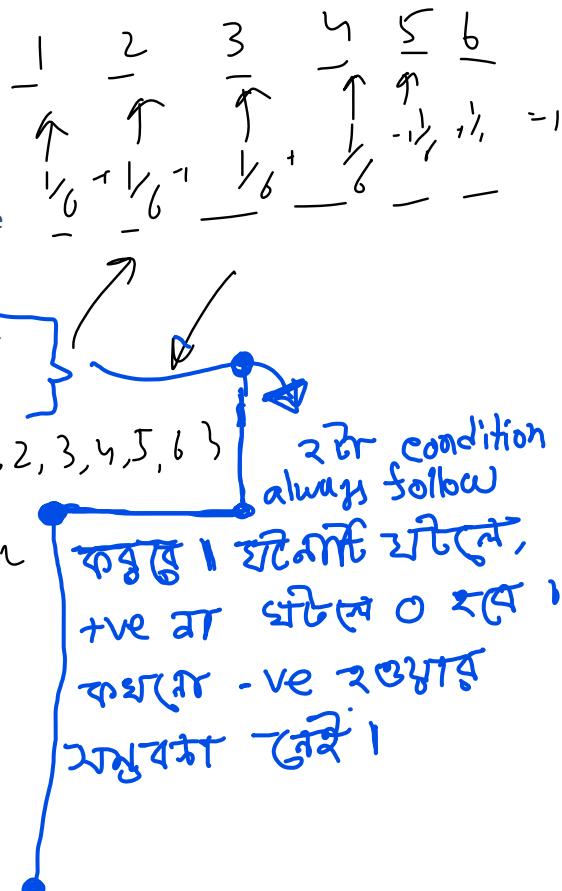
The PMF of a discrete random variable assigns a probability to each possible value of the random variable. The probabilities assigned by the PMF must satisfy two conditions:

- a. The probability assigned to each value must be non-negative (i.e., greater than or equal to zero).
- b. The sum of the probabilities assigned to all possible values must equal 1.



$$y = \begin{cases} \frac{1}{6} & \text{if } x \in \{1, 2, 3, 4, 5, 6\} \\ 0 & \text{otherwise} \end{cases}$$

$$y = \begin{cases} \frac{1}{36} & x \in \{2, 12\} \\ \frac{2}{36} & x \in \{3, 11\} \\ 0 & \text{otherwise} \end{cases}$$



Examples

[https://en.wikipedia.org/wiki/Bernoulli\\_distribution](https://en.wikipedia.org/wiki/Bernoulli_distribution)

[https://en.wikipedia.org/wiki/Binomial\\_distribution](https://en.wikipedia.org/wiki/Binomial_distribution)

Code example: २ दृष्टि एकाधारी प्रायिकता वर्णन करने के लिए Python में यह फलान्तरण करें।

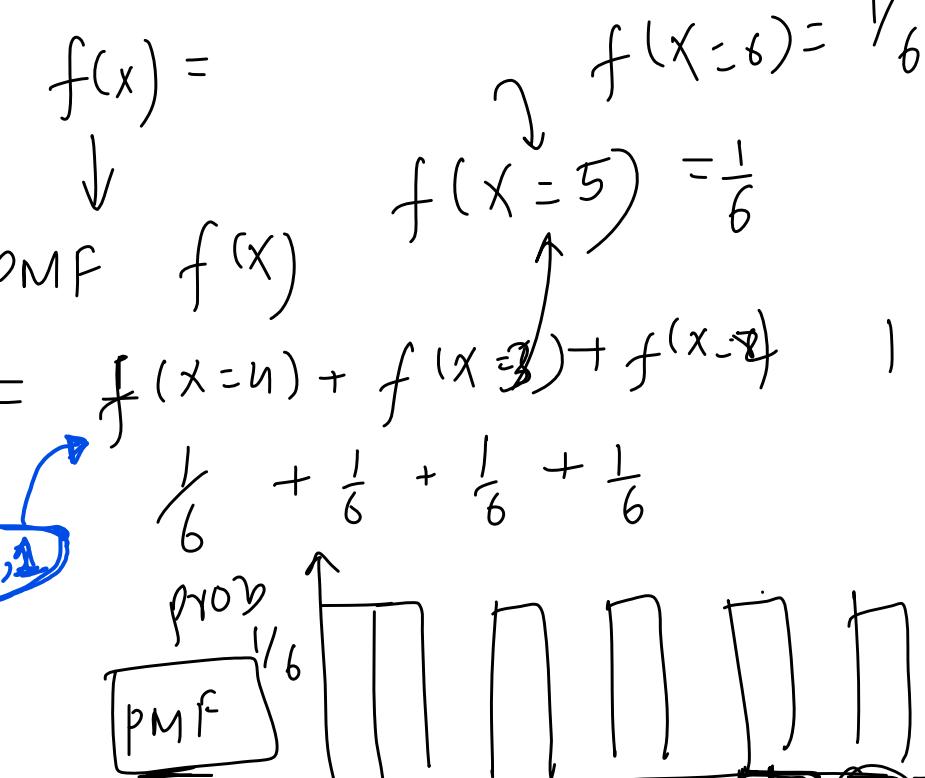
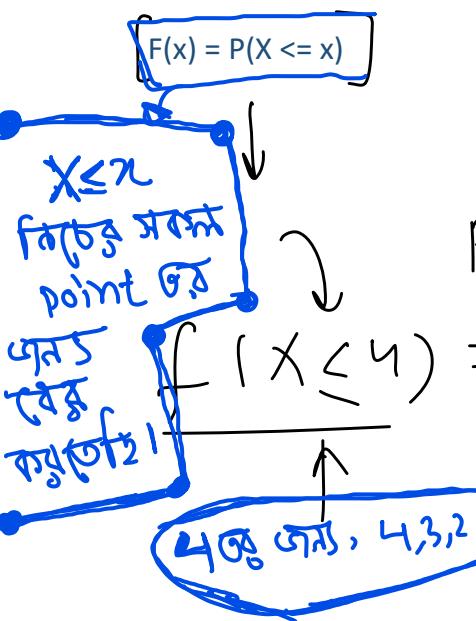


## Cumulative Distribution Function(CDF) of PMF

15 March 2023 20:09

এটি point দ্বারা অন্যত্রে কাটানো  
প্রয়োজন হবে point দ্বারা দুটি  
যত ধ্যান থাকবে সবচেয়ে অন্ত  
বেশ কাটা !

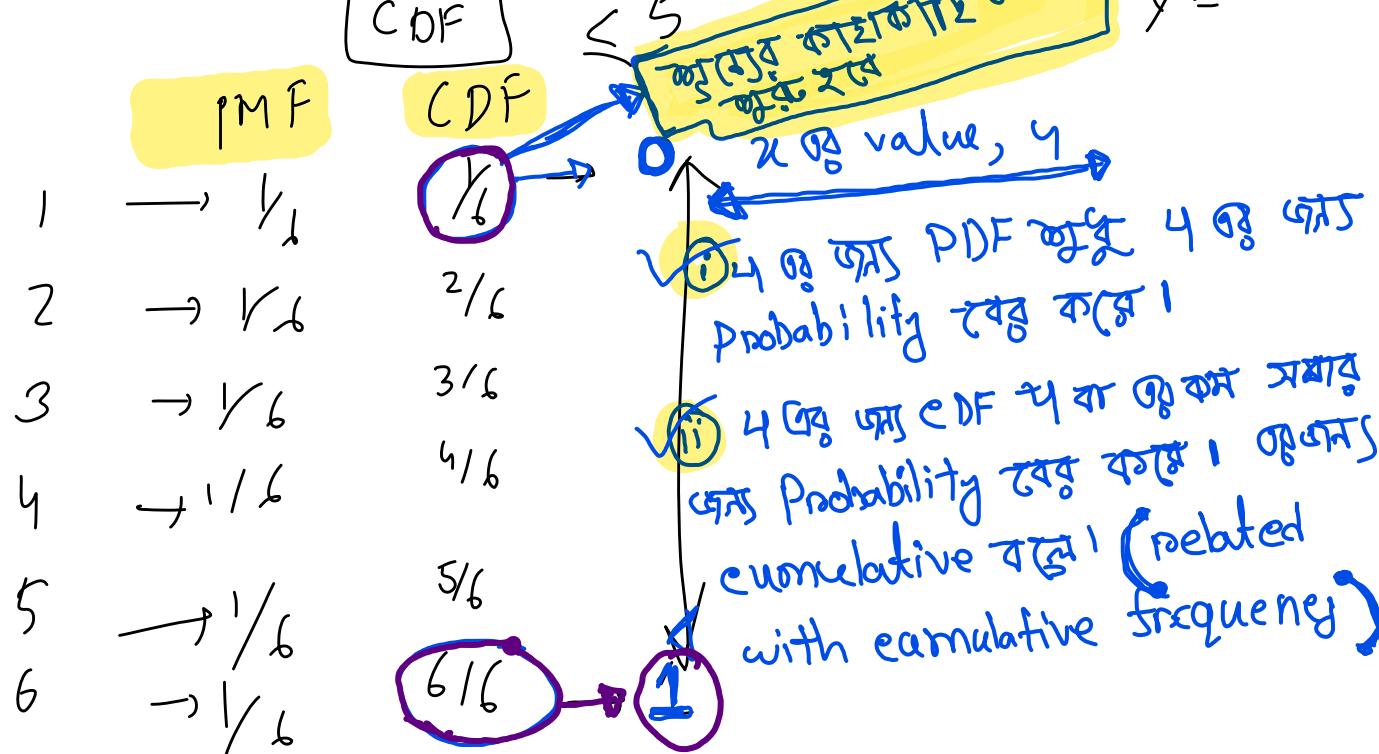
The cumulative distribution function (CDF)  $F(x)$  describes the probability that a random variable  $X$  with a given probability distribution will be found at a value less than or equal to  $x$



Examples:

[https://en.wikipedia.org/wiki/Bernoulli\\_distribution](https://en.wikipedia.org/wiki/Bernoulli_distribution)

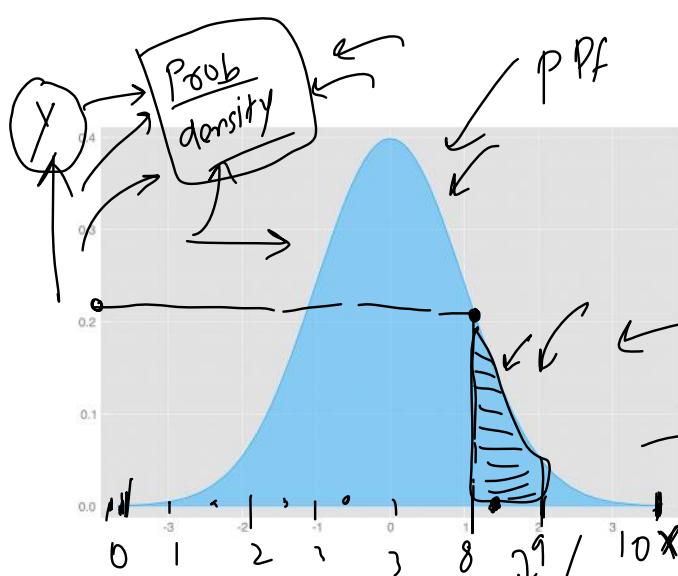
[https://en.wikipedia.org/wiki/Binomial\\_distribution](https://en.wikipedia.org/wiki/Binomial_distribution)



# Probability Density Function (PDF)

15 March 2023 15:25

PDF stands for Probability Density Function. It is a mathematical function that describes the probability distribution of a **continuous random variable**.



sample PDF

$$c_g p_a = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

$$P(0 \leq X \leq 1) = \int_0^1 f(x) dx$$

area

$$P(0 \leq X \leq 1) = 0.199$$

probability between 0 and 1

$$\int_8^9 f(x) dx$$

$8 \rightarrow 9$

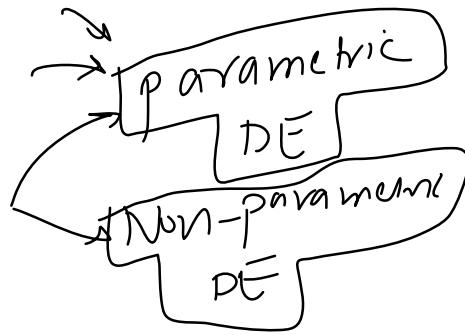
$$8 \rightarrow 8.00$$

- 1. Why Probability Density and why not Probability?
- 2. What does the area of this graph represents?
- 3. How to calculate Probability then?
- 4. Examples of PDF
  - a. [https://en.wikipedia.org/wiki/Normal\\_distribution](https://en.wikipedia.org/wiki/Normal_distribution)
  - b. [https://en.wikipedia.org/wiki/Log-normal\\_distribution](https://en.wikipedia.org/wiki/Log-normal_distribution)
  - c. [https://en.wikipedia.org/wiki/Poisson\\_distribution](https://en.wikipedia.org/wiki/Poisson_distribution)
- 5. How is graph calculated?

# Density Estimation

16 March 2023 06:54

Density estimation is a statistical technique used to estimate the probability density function (PDF) of a random variable based on a set of observations or data. In simpler terms, it involves estimating the underlying distribution of a set of data points.



→ Density estimation can be used for a variety of purposes, such as hypothesis testing, data analysis, and data visualization. It is particularly useful in areas such as machine learning, where it is often used to estimate the probability distribution of input data or to model the likelihood of certain events or outcomes.

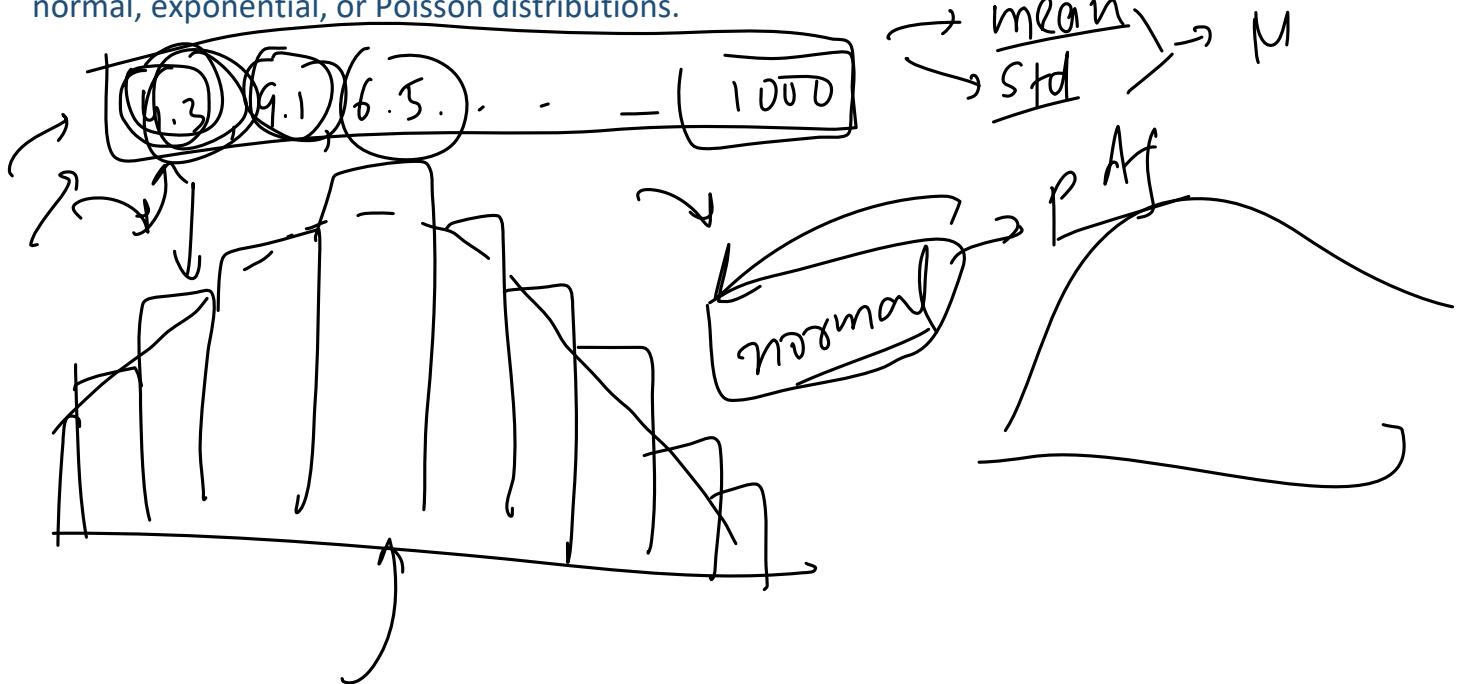
There are various methods for density estimation, including **parametric** and **non-parametric approaches**. Parametric methods assume that the data follows a specific probability distribution (such as a normal distribution), while non-parametric methods do not make any assumptions about the distribution and instead estimate it directly from the data.

Commonly used techniques for density estimation include **kernel density estimation (KDE)**, **histogram estimation**, and **Gaussian mixture models (GMMs)**. The choice of method depends on the specific characteristics of the data and the intended use of the density estimate.

## Parametric Density Estimation

16 March 2023 06:54

Parametric density estimation is a method of estimating the probability density function (PDF) of a random variable by assuming that the underlying distribution belongs to a specific parametric family of probability distributions, such as the normal, exponential, or Poisson distributions.



# Non-Parametric Density Estimation (KDE)

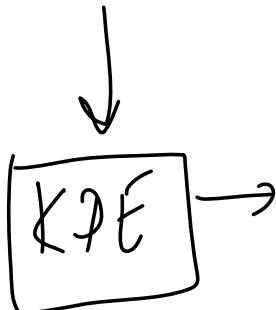
16 March 2023 06:55

But sometimes the distribution is not clear or it's not one of the famous distributions.

- Non-parametric density estimation is a statistical technique used to estimate the probability density function of a random variable without making any assumptions about the underlying distribution. It is also referred to as non-parametric density estimation because it does not require the use of a predefined probability distribution function, as opposed to parametric methods such as the Gaussian distribution.

The non-parametric density estimation technique involves constructing an estimate of the probability density function using the available data. This is typically done by creating a kernel density estimate

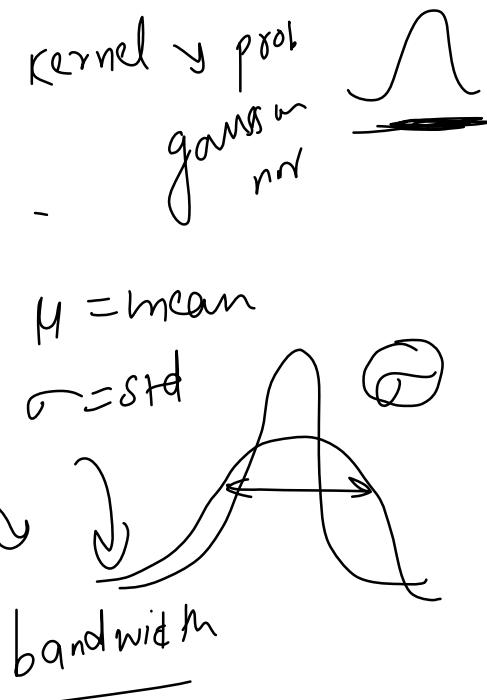
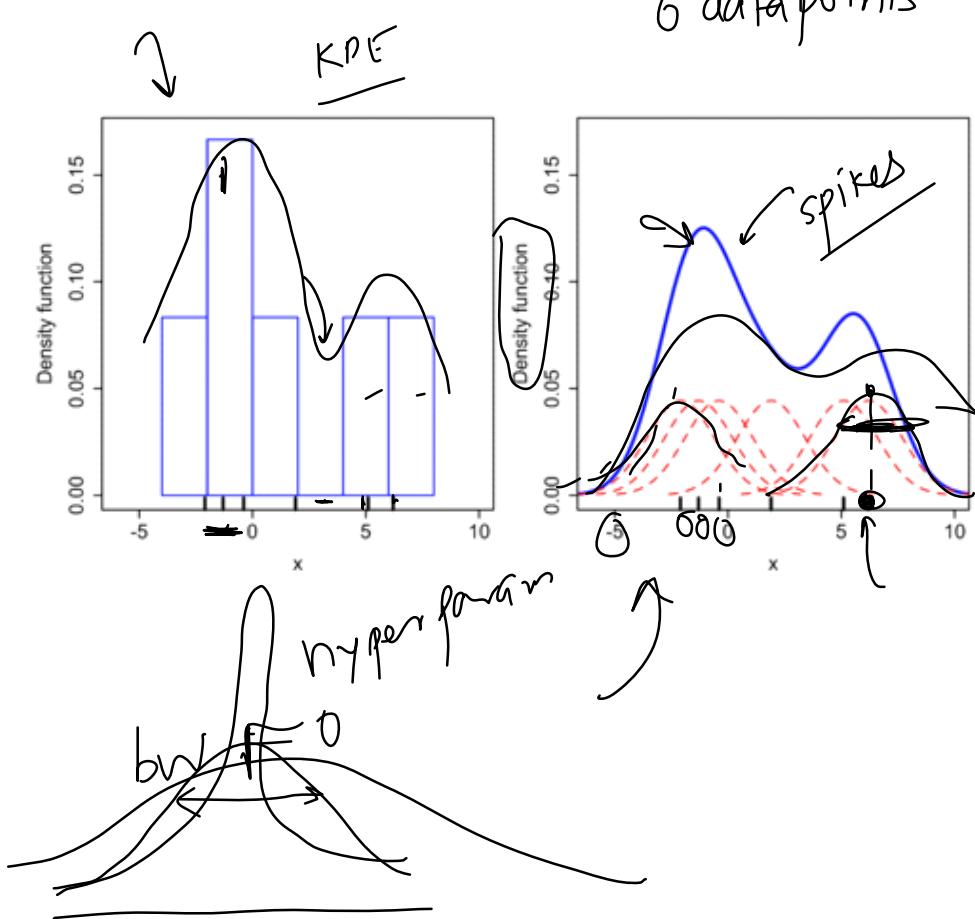
{ Non-parametric density estimation has several advantages over parametric density estimation. One of the main advantages is that it does not require the assumption of a specific distribution, which allows for more flexible and accurate estimation in situations where the underlying distribution is unknown or complex. However, non-parametric density estimation can be computationally intensive and may require more data to achieve accurate estimates compared to parametric methods.



## Kernel Density Estimate(KDE)

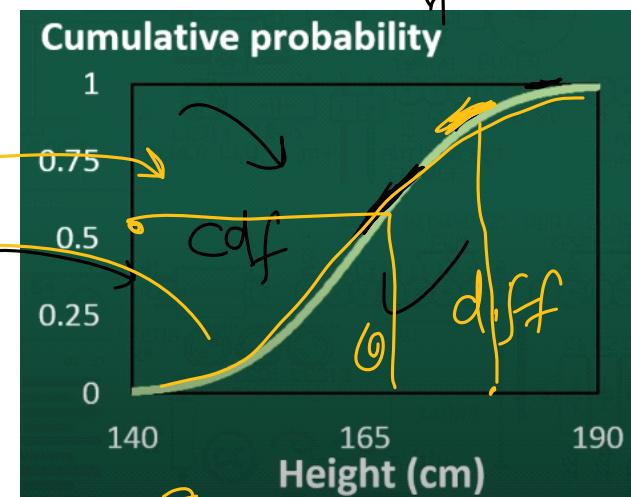
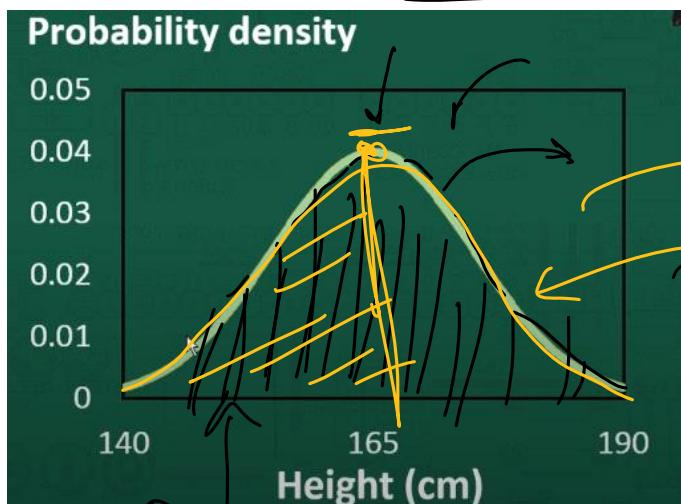
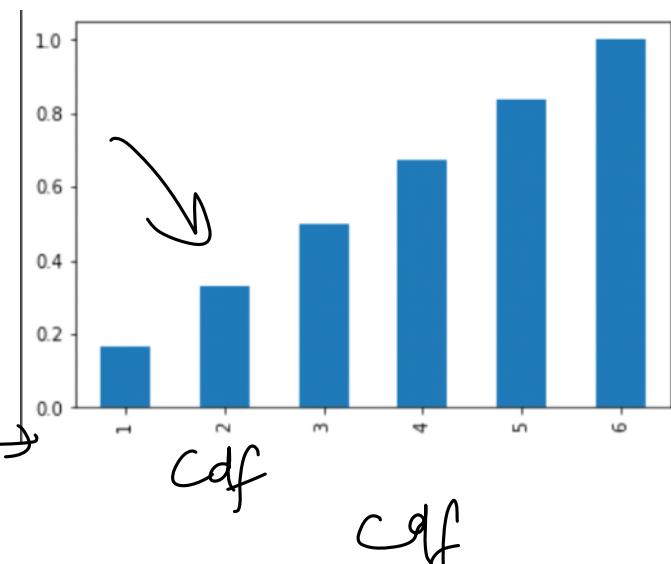
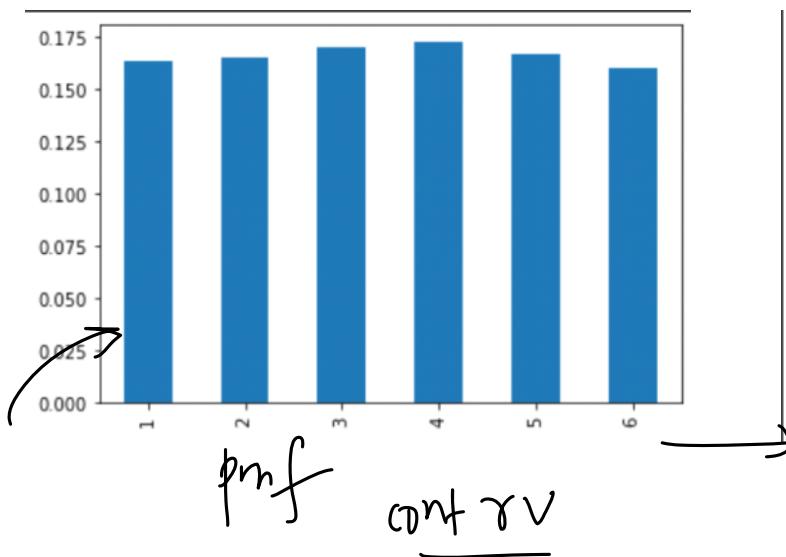
16 March 2023 16:08

The KDE technique involves using a kernel function to smooth out the data and create a continuous estimate of the underlying density function.



# Cumulative Distribution Function(CDF) of PDF

15 March 2023 15:25



integrate in

# How to use PDF and CDF in Data Analysis

15 March 2023 20:10

# 2D Probability Density Plots

16 March 2023 06:50