Random Variables And Distribution

$$X = \begin{cases} 0 & \text{if dice roll is odd} \\ 1 & \text{if dice roll is even} \end{cases}$$

$$\text{Swing} = \begin{cases} 1 & \text{if the voter changes party from last election} \\ 0 & \text{if no change} \end{cases}$$

You arrive at a bus-stop where the only bus comes every 30 minutes. You have exactly 10 minutes that you can wait for the bus. What is the probability you will get the bus?

Img src: unsplash

Instructors

Mousum Dutta
Chief Data Scientist, Spotle.ai
IIT Kharagpur

Dr Sourish Das
Assistant Professor,
Chennai Mathematical Institute
Common Wealth Rutherford Fellow,
University of Southampton

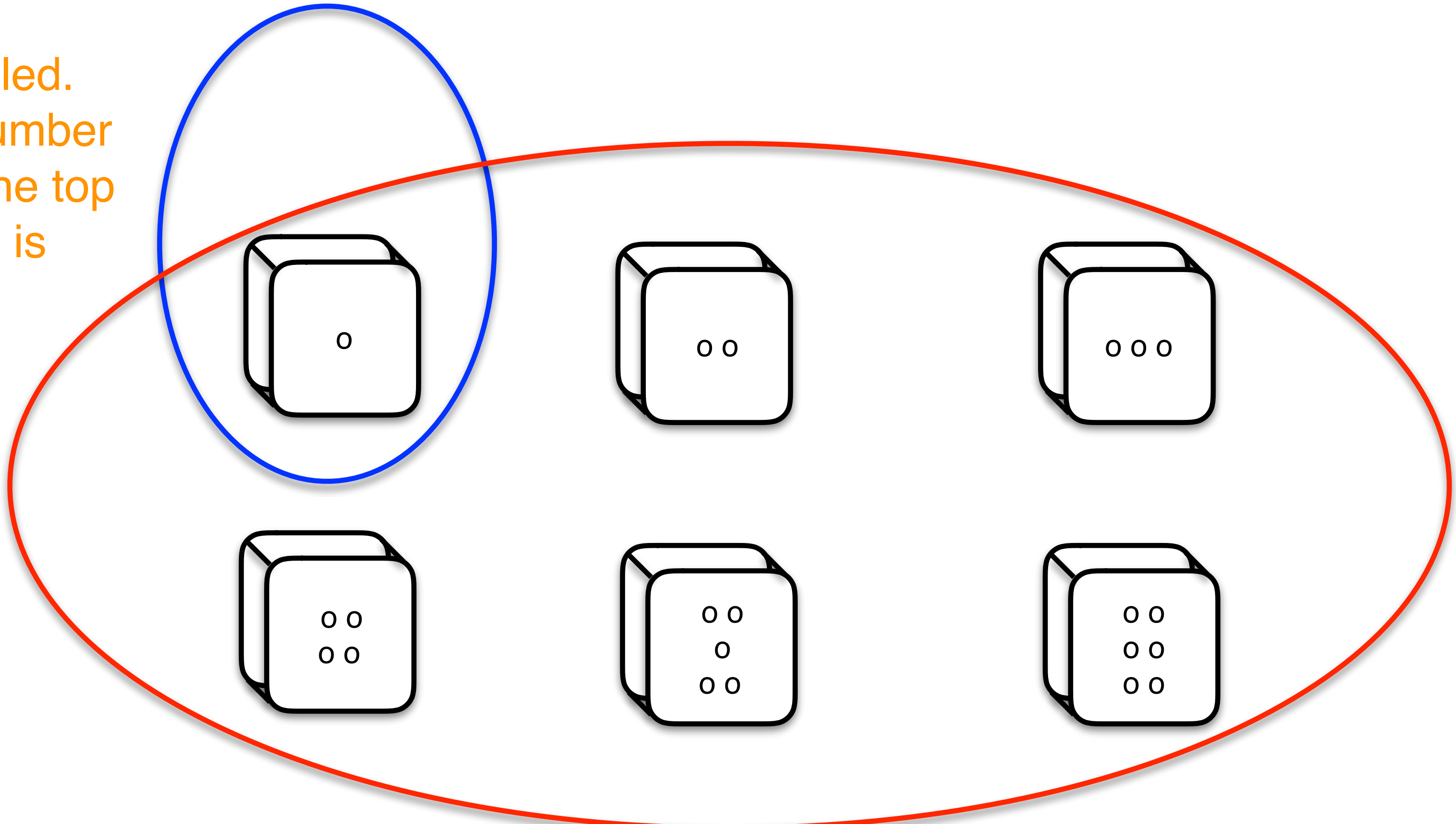
Random Variable

A **random variable** is a function that assigns values to each of an experiment's outcomes. Simply put, a random variable X is set of possible values of a random experiment.

- A random variable is denoted by a capital letter such as X, Y etc.
- A random variable is a function and not a probability. It can take any value, positive or negative.
- A Random Variable's set of values is the **Sample Space**. Random Variable **X** = "The number shown on the top face". **X** can be 1, 2, 3, 4, 5 or 6. Therefore the Sample Space is {1, 2, 3, 4, 5, 6}
- The probability associated with given value of a random variable is denoted by P(**X**=).

Random Variable - Example

A dice is rolled.
X = "The number shown on the top face”. What is P(X=1)?



P(X=1) = (Number of ways 1 can turn up) / (All possible outcomes of a dice roll)
= 1 / 6

Random Variable

Illustration:

- Flip a coin
- You get two outcomes: Head or Tail
- You receive Rs. 10/- if the flip results in a head else give Rs. 10/-
- State Space: {10, -10}
- Income is a random variable

Random Variable | Possible Values | Random Events

$$\mathbf{X} = \begin{cases} +10 \\ -10 \end{cases}$$

Random Variable - Example

Random Variable Possible Values Random Events

$$X = \begin{cases} +10 \\ -10 \end{cases}$$

what is the probability that you will earn Rs 20 in this experiment if we toss the coin 4 times consecutively following this rule? Can we earn Rs 20 if we revise this experiment and just toss the 3 times consecutively?

Watch the subject video for the answer.

Random Variable

Discrete Random Variable

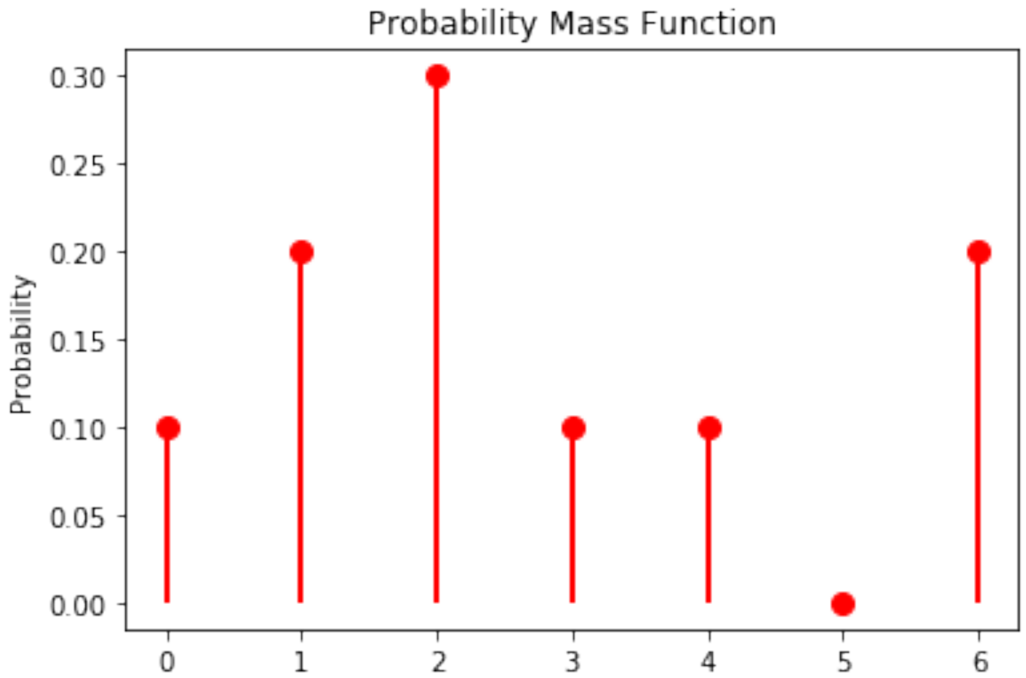
- A random variable that can take a countable number of values.
- For example – The number of words misspelled in a slide.

Continuous Random Variable

- A random variable that can assume any value along a given interval of a number line.
- For example – Time spent on a slide.

Probability Mass Function (PMF)

Probability Mass Function of a discrete random variable X denotes the set of probability values assigned to each of the values taken by X. The PMF is given by $P(X=x_i) = p_i$, $0 \le p_i \le 1 \text{ and } \sum_i p_i = 1$



Probability Mass Function - Example

Two coins are tossed.
If X is defined as
X = 0 if both the toss flips up head,
X = 1 if both the toss flips up tail and
X = 2 if both the flips of coin shows up a different result; that is Head and Tail or Tail and Head.
Then what is the probability of X = 0, X = 1 and X = 2?

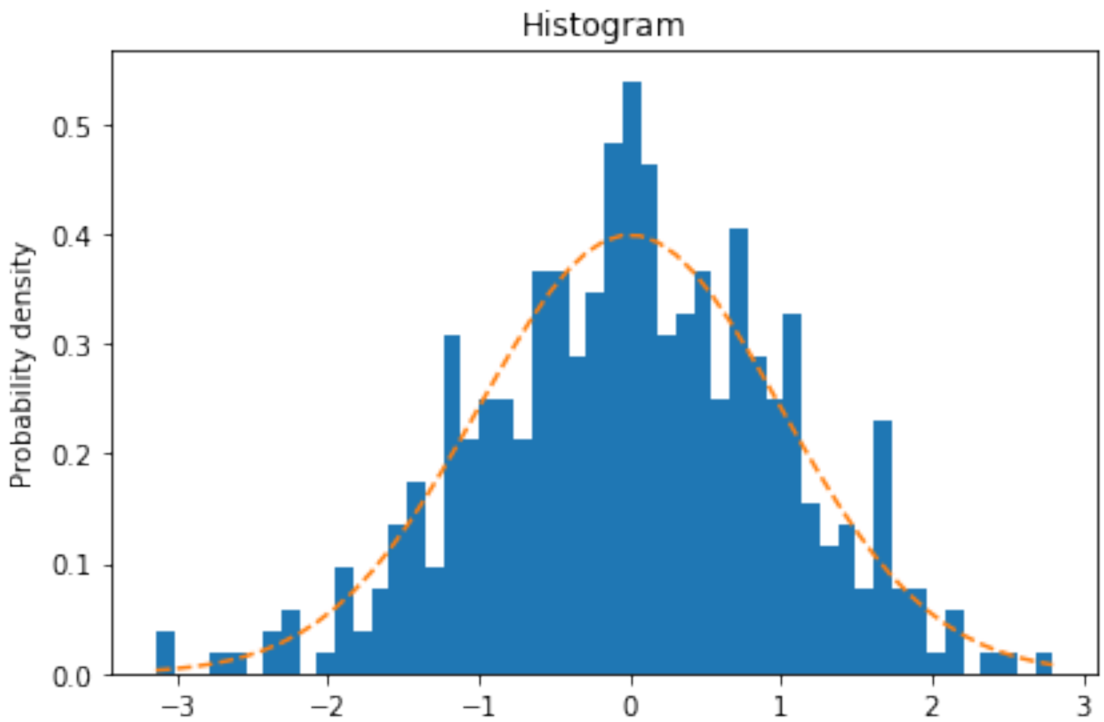
Watch the subject video for the answer.

Probability Density Function (PDF)

PDF denotes the probabilistic properties of a continuous random variable. It is the function whose value at any given point x_i in the sample space denotes the probability of $X = x_i$.

$$f(x) \geq 0$$

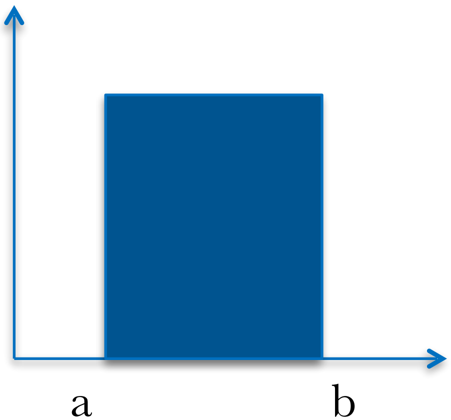
$$\int_x f(x)dx = 1$$



Rectangular Distribution

Uniform distribution or rectangular distribution is a probability density function that has equal probability for all values of the random variable in the sample space.

P = P(X = x)



Rectangular distribution of random variable X, that takes values between a and b with equal probability p.

Since $\int_x f(x)dx = 1$

As sum of all probabilities = 1

Therefore area of rectangle = 1 or p*(b-a) = 1
Or p = 1/ b-a

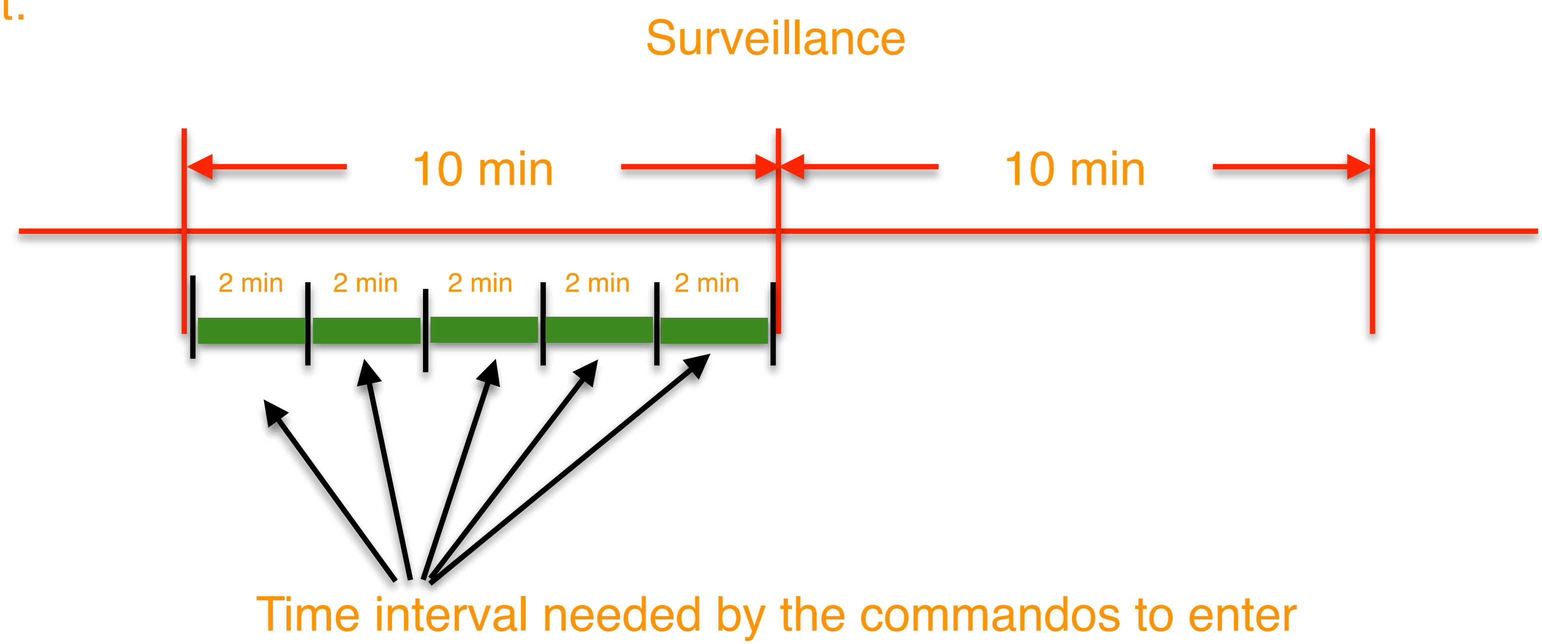
Rectangular Distribution - Example

Suppose a group of commandos need exactly 2 minutes to cross a highly secured area in the enemy camp. A surveillance happens every 10 minutes. The commandos do not know when the last surveillance happened. What is the probability that the commandos will get caught?

Rectangular Distribution - Example

Suppose a group of commandos need exactly 2 minutes to cross a highly secured area in the enemy camp. A surveillance happens every 10 minutes. The commandos do not know when the last surveillance happened. What is the probability that the commandos will get caught?

Hint:



Watch the subject video for the answer.

Cumulative Distribution Function

1. The cumulative distribution function (CDF) of random variable X is given by:

$$F_X(xi)=P(X \le x_i), \text{ for all } x_i \in R$$

2. For a discrete random variable, the CDF is given by:

$$P(X \le x) = F(x) = \sum_{k \le x} P(X = k)$$

3. For a continuous random variable, the CDF is given by:

$$P(X \le x) = F(x) = \int_{-\infty}^{x} f(y)dy$$

4. The probability that X takes a value between a and b, where b >= a is calculated as follows:

$$P(X \le b) = P(x \le a) + P(a < X \le b)$$

Therefore,

$$F_X(b) = F_X(a) + P(a < X \le b) \text{ or}$$

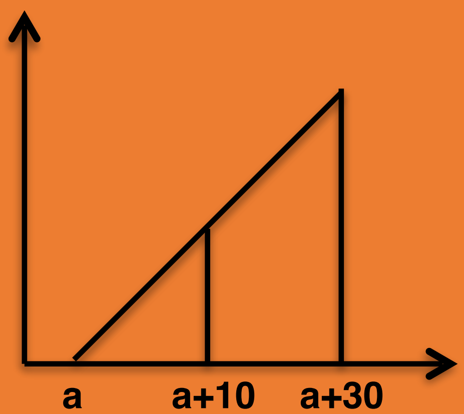
$$P(a < X \le b) = F_X(b) - F_X(a)$$

You arrive at a bus-stop where the only bus comes every 30 minutes. You have exactly 10 minutes that you can wait for the bus. What is the probability you will get the bus?

Hint: You can work out the problem using the CDF function.

Img src: unsplash

Spotle.ai Study Material
Spotle.ai/Learn

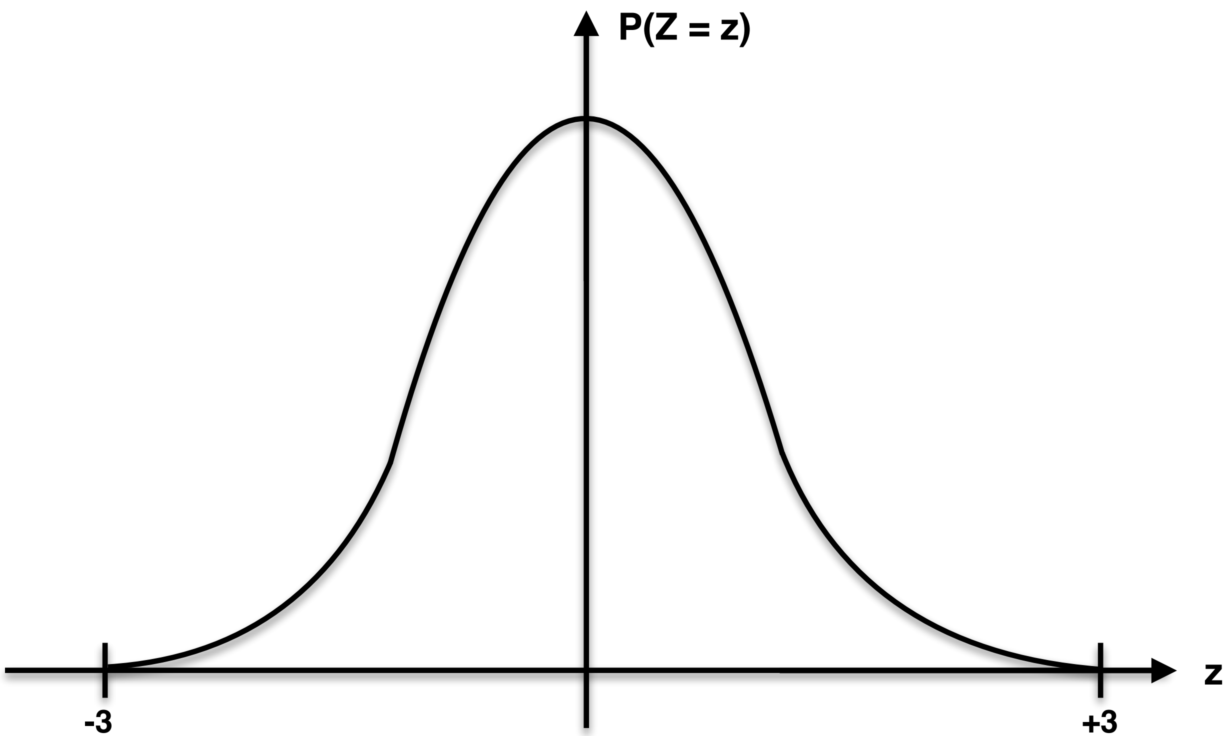


The cumulative probability that the bus arrives between a and a+10 where a is the point in time when you arrived is given by the cumulative probability at point a + 10 or by F(a+10). This works out to 1/3.

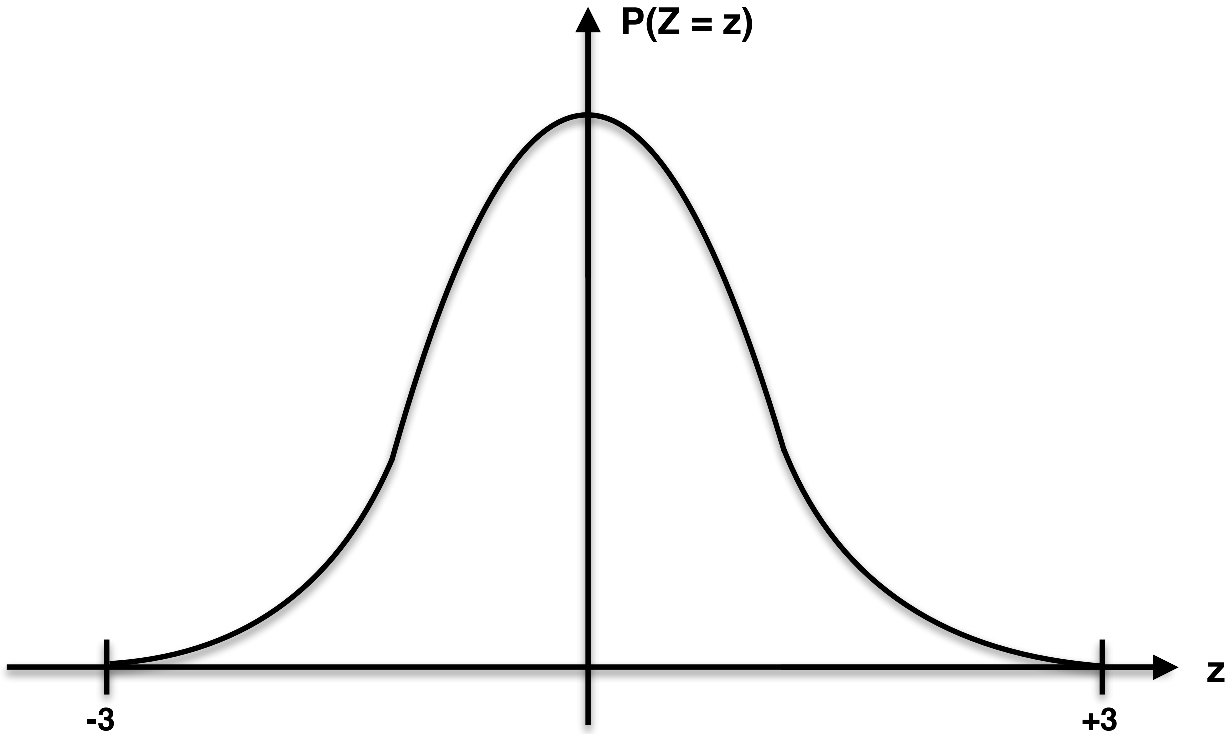
Img src: unsplash

The Normal Distribution

The probability function for the standard normal variable Z follows the bell-shaped symmetric curve or Standard Normal Distribution. Here the probability values are symmetrically spread around a central value without left or right bias.



The Normal Distribution - Think



Think of examples from real world that you think would follow Normal Distribution.

Watch the subject video for the answer.