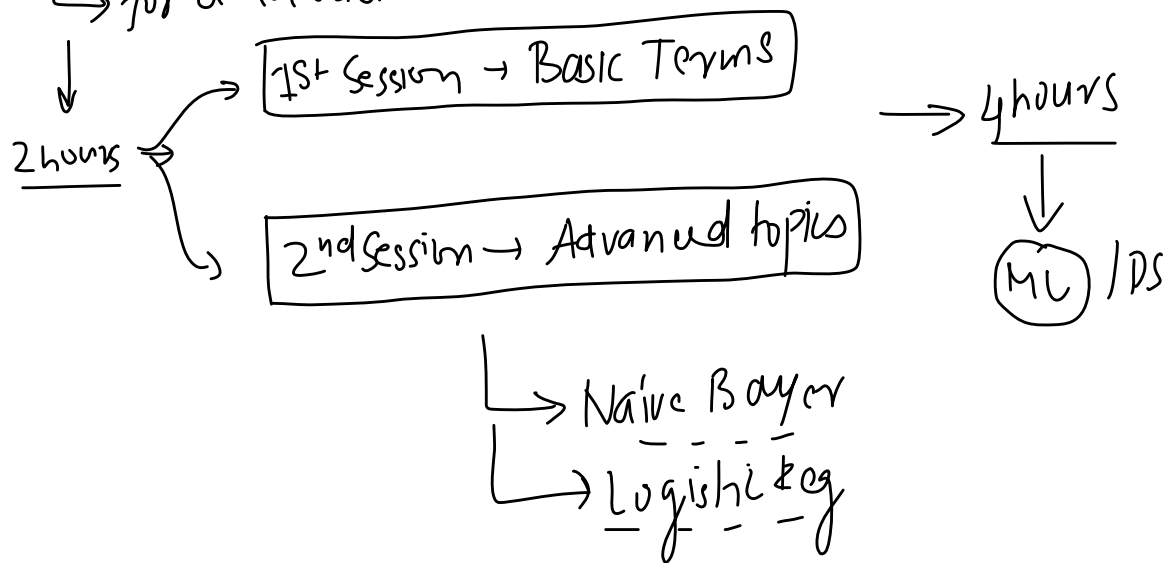


Plan of Attack

16 June 2023 19:46

Crash course on Probability

↳ for data science



Terminology

15 June 2023 18:20

⑤ terms

1. Random Experiment

An experiment is called random experiment if it satisfies the following two conditions:

- (i) It has more than one possible outcome. ✓
- (ii) It is not possible to predict the outcome in advance. ✓

2. Trial

Trial refers to a single execution of a random experiment. Each trial produces an outcome.

3. Outcome

Outcome refers to a single possible result of a trial.

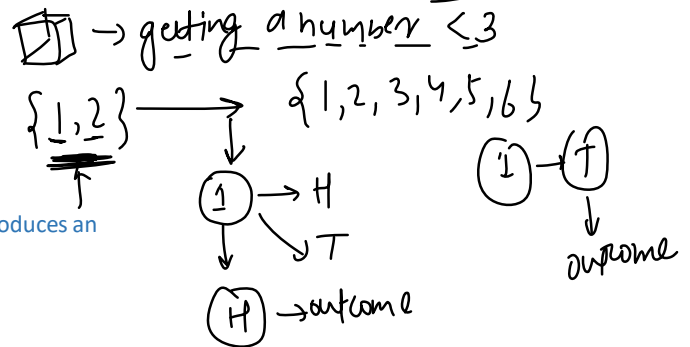
4. Sample Space

Sample Space of a random experiment is the set of all possible outcomes that can occur. Generally, one random experiment will have one set of sample space.

5. Event

Event is a specific set of outcomes from a random experiment or process. Essentially, it's a subset of the sample space. An event can include a single outcome, or it can include multiple outcomes. One random experiments can have multiple events.

tossing a coin \rightarrow RE
 \downarrow
 $\rightarrow H, T$

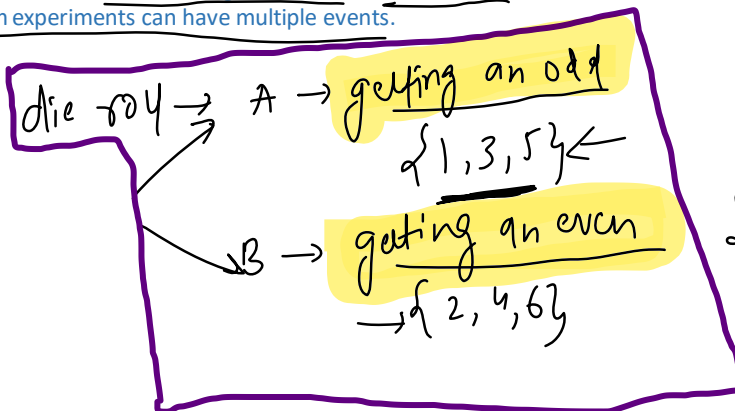


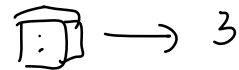
$\textcircled{1} \rightarrow H, T \rightarrow \{H, T\}$

die $\xrightarrow{SS} \{1, 2, 3, 4, 5, 6\}$

$\textcircled{1} \rightarrow \textcircled{H} \rightarrow \{H, T\}$
 \uparrow
 getting ahead $\{H\}$

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





1. Rolling a die
2. Tossing a coin twice
3. Titanic example

- 1) RE \rightarrow rolling a die
- 2) Trial \rightarrow rolling the die once
- 3) Outcome $\rightarrow \{3\}$
- 4) SS $\rightarrow \{1, 2, 3, 4, 5, 6\}$
- 5) Event
 - \rightarrow getting a 3 $\rightarrow \{3\}$
 - \rightarrow getting a number $> 4 \rightarrow \{5, 6\}$
 - \rightarrow getting a number odd $\rightarrow \{1, 3, 5\}$

(1)

- RE \rightarrow Tossing the coin twice
- Trial \rightarrow tossing the coin twice (once)
- Outcome $\rightarrow \{H, T\}$
- \rightarrow SS $\rightarrow \{(H, H), (H, T), (T, H), (T, T)\}$
- Event \rightarrow getting 2 heads $\{(H, H)\}$
- getting at least 1 head $\{(H, H), (H, T), (T, H)\}$
- 3 outcomes

1, 2, 3

Titanic \rightarrow 891 passengers \rightarrow Pclass

RE \rightarrow randomly drawing out a passenger

Trial \rightarrow and finding its Pclass

Outcome $\rightarrow \{1\}$

SS $\rightarrow \{1, 2, 3\}$

Event \rightarrow A \rightarrow the passenger is from Pclass=1 $\{1\}$

B \rightarrow not from Pclass=2 $\{1, 3\}$

Types of Events

16 June 2023 08:29

2, 4, 6

$\{2\}$ $\{1\}$
 $\{2, 4, 6\}$

1. **Simple Event:** Also known as an elementary event, a simple event is an event that consists of exactly one outcome.

For example, when rolling a fair six-sided die, getting a 3 is a simple event.

2. **Compound Event:** A compound event consists of two or more simple events.

For example, when rolling a die, the event "rolling an odd number" is a compound event because it consists of three simple events: rolling a 1, rolling a 3, or rolling a 5.

3. **Independent Events:** Two events are independent if the occurrence of one event does not affect the probability of the occurrence of the other event.

For example, if you flip a coin and roll a die, the outcome of the coin flip does not affect the outcome of the die roll.

① ②
 ↑ ↑
 spade

4. **Dependent Events:** Events are dependent if the occurrence of one event does affect the probability of the occurrence of the other event.

For example, if you draw two cards from a deck without replacement, the outcome of the first draw affects the outcome of the second draw because there are fewer cards left in the deck.

52
 ① → $P(\text{Ace}) = \frac{13}{52}$
 $P(\text{spade}) = \frac{13}{52}$

5. **Mutually Exclusive Events:** Two events are mutually exclusive (or disjoint) if they cannot both occur at the same time.

For example, when rolling a die, the events "roll a 2" and "roll a 4" are mutually exclusive because a single roll of the die cannot result in both a 2 and a 4.

roll in → odd
 → even

6. **Exhaustive Events:** A set of events is exhaustive if at least one of the events must occur when the experiment is performed.

For example, when rolling a die, the events "roll an even number" and "roll an odd number" are exhaustive because one or the other must occur on any roll.

⑦ → impossible
 $1 \leq x \leq 6$

7. **Impossible event and Certain Event**

What is Probability

15 June 2023

18:14

In simplest terms, probability is a measure of the likelihood that a particular event will occur. It is a fundamental concept in statistics and is used to make predictions and informed decisions in a wide range of disciplines, including science, engineering, medicine, economics, and social sciences.

Probability is usually expressed as a number between 0 and 1, inclusive:

- A probability of 0 means that an event will not happen.
- A probability of 1 means that an event will certainly happen.
- A probability of 0.5 means that an event will happen half the time (or that it is as likely to happen as not to happen).

$\boxed{1} \rightarrow$

Type of Probability:

Empirical Probability Vs Theoretical Probability

15 June 2023 18:15

Empirical Probability

Empirical probability, also known as experimental probability, is a probability measure that is based on observed data, rather than theoretical assumptions. It's calculated as the ratio of the number of times a particular event occurs to the total number of trials.

A. Suppose that, in our 100 tosses, we get heads 55 times and tails 45 times. What is the empirical probability of getting a head

B. Let's say you have a bag with 50 marbles. Out of these 50 marbles, 20 are red, 15 are blue, and 15 are green. You start to draw marbles one at a time, replacing the marble back into the bag after each draw. After 200 draws, you find that you've drawn a red marble 80 times, a blue marble 70 times, and a green marble 50 times. What is the empirical probability of getting a red marble?

Theoretical Probability

Theoretical (or classical) probability is used when each outcome in a sample space is equally likely to occur. If we denote an event of interest as Event A, we calculate the theoretical probability of that event as:

Theoretical Probability of Event A = Number of Favourable Outcomes (that is, outcomes in Event A) / Total Number of Outcomes in the Sample Space

A. Consider a scenario of tossing a fair coin 3 times. Find the probability of getting exactly 2 heads.

B. Consider a scenario of rolling 2 dice. What is the probability of getting a sum = 7

No. of trials → infinite tosses

empirical prob → theoretical

10 → 3 → 0.3
100 → 45H → 0.45
1000 → 470 → 0.47

Passenger → 70 M 30 F
100 → 70 M 30 F
3 males, $P(M) = 0.3$ → $P(M) = 0.7$

H → 55
Trials → 100

$$P(H) = \frac{55}{100}$$

20 red
80
200

1, 2, 3, 4, 5, 6

roll → 3

1
6 → $P(3)$

{H, T} → 1
2 → $P(4)$

যদি, আমাদের numbers of trial অনেক বেশি হয় তাহলে, আমাদের empirical probability তে আর Theoretical probability আরো আসন্ন হয়।

Implement this with python code.

Variable হালও ও কাজ function চরু মতো। সেইটা ফাংশনকে মতোই। i/p o/p দিয়ে কাজ করে। i/p on o/p কী ওয়ার্ড?

Random Variable

15 June 2023 18:16

misleading

function

In the context of probability theory, a random variable is a function that maps the outcomes of a random process (known as the sample space) to a set of real numbers.

Input: The input to the function is an outcome from the sample space of a random process. Input হলে all the sample space. প্রতিটা টার প্রতি করে S.S. input ও ফাংশন।

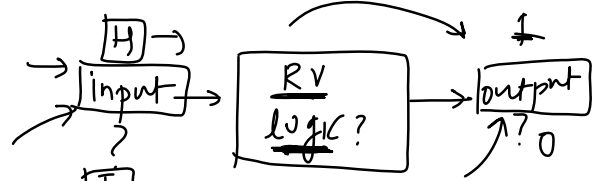
Output: The output of the function is a real number that we assign to each possible outcome.

Discrete RV
Continuous RV

$$X = \{1, 0\}$$

H T

input → logic → output



random experim
tossing a coin

$$SS = \{H, T\}$$

rolling a die

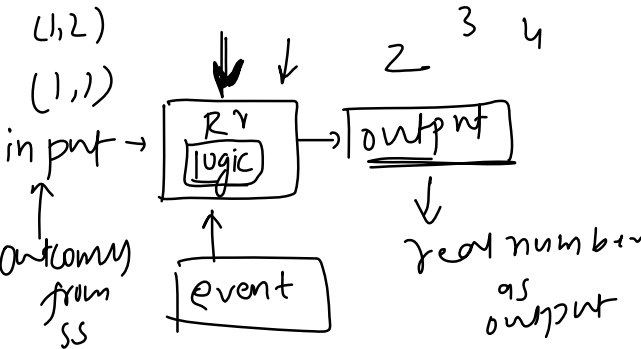
$$Y = \{1, 2, 3, 4, 5, 6\}$$

sample space {1, 2, 3, 4, 5, 6}

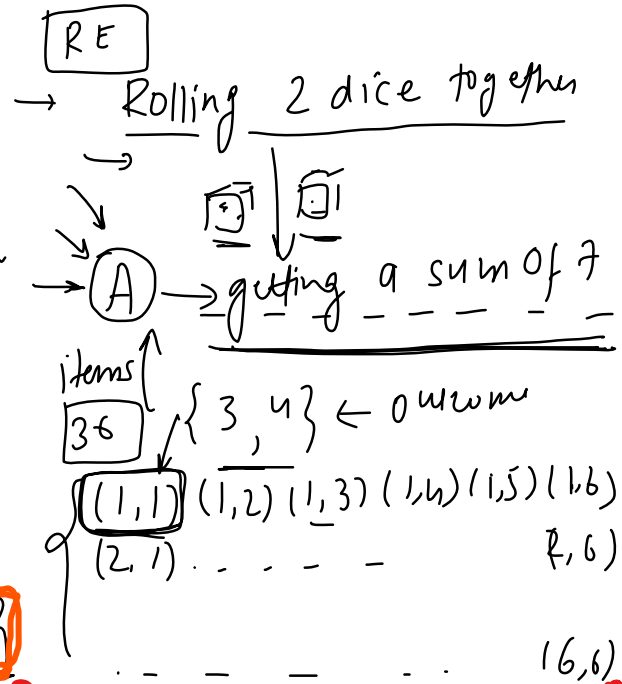
ই logic টার কী হবে।

The transformation from input to output in the function of a random variable is determined by how we choose to define the random variable. And the choice of how to define a random variable often depends on the specific aspects of the random process (or event) that you're interested in studying.

event থেকে logic বানাতে, event কী? কোনটা চাচ্ছি।

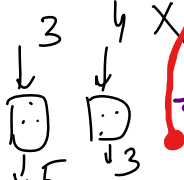


sample space
item → sum outcome



$$X = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$$

PMF →

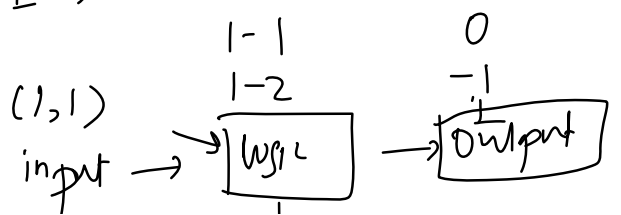


এটা হল rolling ৩ টার ফল থেকে, sample space (S.S) {3, 4} থেকে ৩ টার করে 3+6=9 করে, random variable বাবুসি

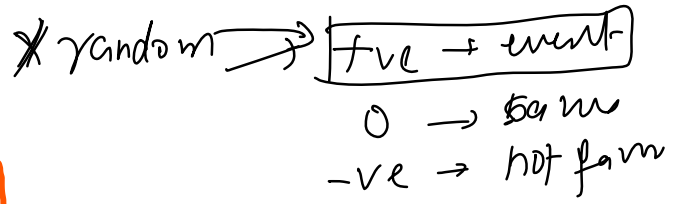
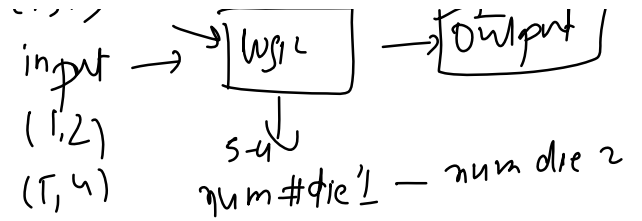
RE → rolling 2 dice

→ A → where the num on die 1 > num on die 2

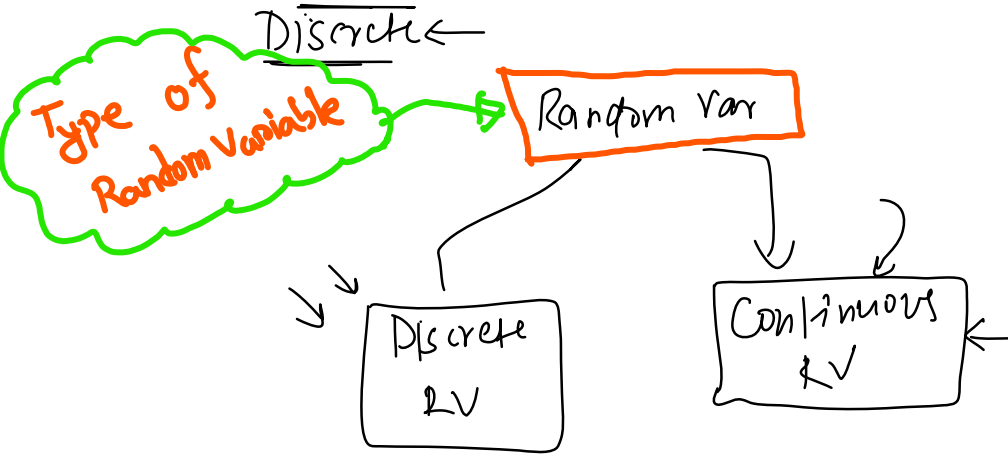
$$SS = \{(1,1), (1,2), (1,3), (1,4), (1,5), (1,6), (2,1), \dots, (6,6)\}$$



SS = $\{(1,1), (1,2), (1,3), (1,4), (1,5), (1,6), (2,1), (2,2), (2,3), (2,4), (2,5), (2,6), (3,1), (3,2), (3,3), (3,4), (3,5), (3,6), (4,1), (4,2), (4,3), (4,4), (4,5), (4,6), (5,1), (5,2), (5,3), (5,4), (5,5), (5,6), (6,1), (6,2), (6,3), (6,4), (6,5), (6,6)\}$



output \rightarrow 1.5
 1.32



RE \rightarrow cga \rightarrow 0-10

{ 0-10 }

$\{0-10\} \rightarrow$ **logic** \rightarrow **contin. den**
8.35 \rightarrow 8.35

Statistics - Lecture 03 (Discrete)

Probability Distribution of a Random Variable (Discrete)

15 June 2023 18:16

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

- 1) Toss
- 2) 1 die
- 3) 2 dice

<u>X</u>	<u>1</u>	<u>0</u>
<u>P(X)</u>	$\frac{1}{2}$	$\frac{1}{2}$

Probability dist of a random var

Toss a coin
 $\{H, T\}$

$\rightarrow RV \rightarrow X = \{1, 0\}$

$P(X=1) = \frac{1}{2}$ $P(X=0) = \frac{1}{2}$

Rolling a die
 $SS = \{1, 2, 3, 4, 5, 6\}$

$\rightarrow RV \rightarrow \{1, 2, 3, 4, 5, 6\}$

<u>X</u>	<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>5</u>	<u>6</u>
<u>P(X)</u>	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Probability dist of random variable (input)

Sample space

output

Rolling 2 dice

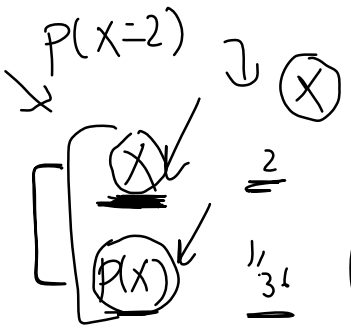
$SS = 36 \text{ items}$

$RV \rightarrow \text{sum}$

1 2

(a,b)	1	2	3	4	5	6
1	(1,1)	(2,1)	(3,1)	(4,1)	(5,1)	(6,1)
2	(1,2)	(2,2)	(3,2)	(4,2)	(5,2)	(6,2)
3	(1,3)	(2,3)	(3,3)	(4,3)	(5,3)	(6,3)
4	(1,4)	(2,4)	(3,4)	(4,4)	(5,4)	(6,4)
5	(1,5)	(2,5)	(3,5)	(4,5)	(5,5)	(6,5)
6	(1,6)	(2,6)	(3,6)	(4,6)	(5,6)	(6,6)

+	1	2	3	4	5	6
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



$X = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$

Probability Dist Function

100 dice roll

$100 \rightarrow 600$

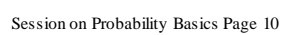
contin
RV

$X \rightarrow P(X)$

$\rightarrow (y, f(x))$

$P(X)$

$y = x^2 + 5x + 3$



আমরা Probability দিয়ে mean বের করতে পারি। এইটে দেখাও।

Mean of a Random Variable

15 June 2023 18:17

The mean of a random variable, often called the expected value is essentially the average outcome of a random process that is repeated many times. More technically, it's a weighted average of the possible outcomes of the random variable, where each outcome is weighted by its probability of occurrence.

rolling a die

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

mean

Mean of X

multiple

trials (10)

1000 die

avg

mean (X)

$$\frac{3+4+2+4+\dots}{1000} = \square$$

mean
বের করে ÷

$$\frac{1}{6} \times 5 + \frac{1}{6} \times 3 + \frac{1}{6} \times 4 + \frac{1}{6} \times 5 + \frac{1}{6} \times 3 + \frac{1}{6} \times 3 = \frac{5+3+4+5+3+3}{6} = \frac{23}{6}$$

$$\frac{2}{6} = \frac{1}{3}$$

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

$$= \left(\frac{1}{3} \right) (5) + \frac{1}{6}(4) + \frac{1}{2}(3) =$$

$$E[X] = \sum_{i=1}^n x_i P(x_i)$$

probability of that number

number

Variance of a Random Variable

15 June 2023 18:17

The variance of a random variable is a statistical measurement that describes how much individual observations in a group differ from the mean (expected value).

rolling
X

$E[X] = 3.5$

normal variance $\rightarrow Var(X)$

var = $\sum_{i=1}^n \frac{(x_i - \bar{x})^2}{n-1}$

mean

2, 4, 5, 6, 7, ... 10

Squared dist

mean

variance \rightarrow spread

avg

$(1-3.5)^2$ $(2-3.5)^2$

$(3-3.5)^2$

mean \rightarrow Expect

$(X - E[X])^2$

6 min

$Var(X) = E[(X - E[X])^2]$

$Var(X) = E[(X - E[X])^2]$

and given random variable X

$$Var(X) = E[(X^2 + (E[X])^2 - 2XE[X])]$$

$$= E[X^2] + E[(E[X])^2] - E[2XE[X]]$$

$$= E[X^2] + E[(E[X])^2] - \underbrace{E[2]}_2 \underbrace{E[X]}_{E[X]} \underbrace{E[E[X]]}_{E[X]}$$

$E[X] \rightarrow$ number \rightarrow constant

$E[X] \rightarrow$ 1, 2, 3

2×1

$$= E[X^2] + E[(E[X])^2] - 2E[X]E[X]$$

$$= E[X^2] + \underbrace{E[(E[X])^2]}_{(E[X])^2} - 2(E[X])^2$$

$$= E[X^2] + (E[X])^2 - 2(E[X])^2$$

$$= E[X^2] + \overset{\downarrow}{(E[X])^2} - 2(E[X])^2$$

$$\text{Var}(X) = E[X^2] - (E[X])^2$$

→ ✓

cont
discrete

$$\text{Var}(X) = E[(X - E[X])^2]$$

✓