6/1

Questions :

(1) Is every continuous function differentiable ? No

(2) Does every vector space consist of a basis ? Yes

(3) Will it rain tomorrow ? Yes or No → Weather predict

(4) Is it safe to release a tablet in the market ? Yes or No

(5) Fix an Integer n, is it a prime ? Either Yes or No

Probability

→ A number in $[0,1]$ → Function

Closer to 1 : Event is more likely to appear

Closer to 0 : Event is not likely to appear

Classical Probability :

- Event : Possible outcome of an experiment
  ↳ Finite Set
- Sample Space : A sample space for an experiment is a set S containing each possible outcome of the experiment.
  ↳ An element of S is called a sample point.

  - Let's consider the experiment of tossing a coin. In this case,
    $S = \{head, tail\}$
  - Let's consider the experiment of throwing a dice. In this case,
    $S = \{1, 2, 3, 4, 5, 6\}$
  - Let's consider the experiment of choosing a real number b/w -1 and 1.
    $S = \{x \in \mathbb{R} \mid -1 < x < 1\}$
    ↳ Sample set is infinite. Classical probability cannot be defined.
    But there's another way to define the probability.

- Event : Any subset A of a sample space S is called an event.
- Classical Probability : The probability of an event A is

$$\boxed{P(A) = \frac{n(A)}{n(S)}}$$

$n(A)$: no. of elements in A

$n(S)$: no. of elements in S.

Properties :

Dependent of no. of elements in set A

Dependent of no. of elements in set C

Counting — Easy & Difficult
↳ Coin ex. ↳ Prime ex.

Ex. Let's consider the experiment of tossing a coin.
Here, $S = \{H, T\}$

Take $A = \{H\}$
$\Rightarrow P(A) = \frac{n(A)}{n(S)} = \frac{1}{2}$ [Easy counting]
↳ Mathematically

Probability → Not exactly 50–50 Practically
$\exists$ Slight error

Probability Model → Reduce the error as much as possible

Ex. Let's consider an experiment of tossing a dice.
Here, $S = \{1, 2, 3, 4, 5, 6\}$
Suppose, $A = \{1, 2\} \Rightarrow P(A) = 2/6 = 1/3$
Also, Suppose, $A = \{1, 2, 3\} \Rightarrow P(A) = 3/6 = 1/2$
↳ More likely to appear than above $[\because 1/2 > 1/3]$

Classical probability

S - Sample space, E = Event

$P(E) = \frac{n(A)}{n(S)}$

Axioms of probability

let S be a sample space A prob. function is funct^n P.
$P: S \to [0, 1]$ satisfy

(I) $P(S) = 1$ (II) $P(E) \geq 0 \; \forall \; E \subseteq S$

(III) $P(E_1 \cup E_2 \cup \dots \cup E_k) = P(E_1) + P(E_2) \dots P(E_k)$
where $E_1, E_2 \dots E_k$ are events with $E_1 \cap E_2 \dots E_k = 0$

counting technique for find probability

A Permutation is arrangement of objects in a definite order

example :– The number of permutation of n distinct objects used r at a time n denoted by

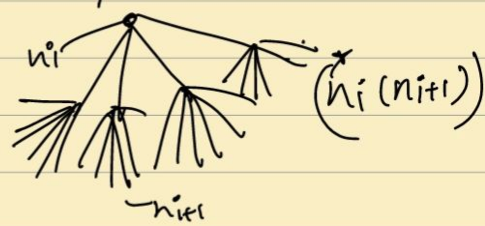
$$^{n}P_r = \frac{n!}{(n-r)!} \qquad \underbrace{\overset{n}{\square}\ \overset{(n-1)}{\square}\ \cdots\cdots\ \overset{n-(r-1)}{\square}}_{r\text{-position}}$$

goal :→ (putting n objects in r position)

$$^{n}P_r = n(n-1)\cdots\cdots(n-(r-1))$$

Multiplication principle.

Consider an experiment having k stages and the i^{th} stage can occur in n_i number of ways Then the experiment can occur in $n_1 \cdots\cdots n_r$ number of ways



Combinations $\{a_1, a_2, a_3\} \to$ arrang $\quad a_1 a_2 a_3$ — 6 $\quad \underline{3!}$

$a_3\, a_2\, a_1$

$a_2\, a_3\, a_1$

$\vdots \quad \vdots \quad \vdots$

combination $\{a_1, a_2, a_3\}$

$\{a_2, a_3, a_1\}$

$\{a_1\ a_3, a_2\}$

Definitn :→

A combination is a selection of an objects from a given set of n objects.

examples :– Select r objects from n objects $\qquad {}^{n}C_r = \dfrac{{}^{n}P_r}{r!}$

$${}^{n}C_r = \frac{n!}{(n-r)!\,r!}$$

Example:- A foundry ships a lot of 20 engine blocks of which five contain internal flaws. The purchaser will selects three blocks at random and test them for hardness. The lot will be accepted if no flaws are found. What is the probability that this lot will be accepted?

Soln :-

20 Total 5 flaws 3 select

$$= \frac{{}^{15}C_3}{{}^{20}C_3} = \frac{455}{1140} = 0.39$$

2) The contribution of the blood types in the united states is roughly 41% type A, 9% type B, 4% type AB and 46% type O. What is the P that a random person has type, A, B or AB?

Soln :- 41 — A, 9 — B, 4 — AB, 46 — O

A_1 is the event of a random guy has A type

A_2 ——— B type

A_3 ——— AB type

$$P(A_1 \cup A_2 \cup AB) = P(A_1) + P(A_2) + P(A_3).$$

8/1

$$P(A) = \frac{n(A)}{n(S)} \quad , \quad n(S) \neq 0$$

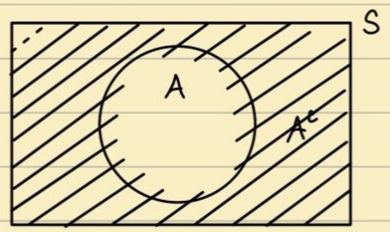
Properties :

(1) $P(\phi) = 0$

$$P(\phi) = \frac{n(\phi)}{n(S)} = \frac{0}{n(S)} = 0$$

(2) Let $A \subseteq S$ be an event , Then

$$P(A^c) = 1 - P(A)$$



(3) Let $A_1, A_2 \subseteq S$ be two events , Then

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$$

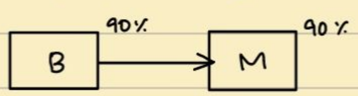
Proof:

$$P(A_1 \cup A_2) = \frac{n(A_1 \cup A_2)}{n(S)} = \frac{n(A_1) + n(A_2) - n(A_1 \cap A_2)}{n(S)}$$

$$= \frac{n(A_1)}{n(S)} + \frac{n(A_2)}{n(S)} - \frac{n(A_1 \cap A_2)}{n(S)}$$

$$= P(A_1) + P(A_2) - P(A_1 \cap A_2)$$

Ex. We have a propulsion system in series



Parallel

Find the probability that the engine component is operable

Sol: Let A_1 be the event that the main engine operates

Let A_2 be the event that the backup engine operates

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$$

$= 0.9 + 0.9 - P(A_1 \cap A_2)$

$= 0.9 + 0.9 - P(A_1) \cdot P(A_2)$

$= 0.9 + 0.9 - 0.81$

$= 0.99$

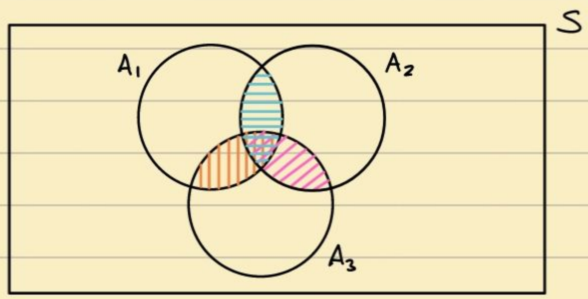
For Independent events :

$P(A_1 \cap A_2) = P(A_1) \cdot P(A_2)$

- Additive Property :

Let $A_1, A_2, A_3, \ldots\ldots, A_n$ be n events, Then

$$P(A_1 \cup A_2 \cup A_3 \cup \ldots\ldots \cup A_n)$$
$$= P(A_1) + P(A_2) + P(A_3) + \ldots\ldots P(A_n)$$
$$- P(A_1 \cap A_2) - P(A_1 \cap A_3) - \ldots\ldots - P(A_1 \cap A_n) - P(A_2 \cap A_3) - \ldots\ldots - P(A_i \cap A_k)$$
$$+ \sum_{i<j<k} P(A_i \cap A_j \cap A_k) - \sum_{i<j<k<l} P(A_i \cap A_j \cap A_k \cap A_l) + \ldots\ldots$$
$$+ (-1)^n P(A_1 \cap A_2 \cap \ldots\ldots A_n)$$



Shelldon Ross
(Classical Prob.)

Prob. with R

Proof - H.W

9/1

Prove :

$$n(A_1 \cup A_2 \cup A_3 \ldots\ldots \cup A_n)$$
$$= \sum_{i=1}^{n} n(A_i) - \sum_{i<j} n(A_i \cap A_j) + \sum_{i<j<k} n(A_i \cap A_j \cap A_k) - \sum_{i<j<k<l} n(A_i \cap A_j \cap A_k \cap A_l)$$
$$+ (-1)^n n(A_1 \cap A_2 \cap \ldots\ldots \cap A_n)$$

By Induction hypothesis,

$N = \{1, 2, 3, \ldots\ldots\}$

$S_1, S_2, S_3, S_4, S_5, \ldots\ldots$

S_i : i^{th} statement

If S_1 is true & S_{n+1} is true, assuming S_n is true, Then, all the statements $S_1, S_2, S_3, \ldots\ldots S_{n-1}$ are true.

$F = \{m \mid S_m \text{ is false}\}$

$F \neq \emptyset$ and $F \subseteq N$

$\hookrightarrow$ F should have a minimum element k

$\Rightarrow k-1 \notin F \Rightarrow S_{k-1}$ is true $\Rightarrow S_k$ is true $\therefore$ Contradicts

$S_1 : n(A_1) = n(A_1)$

$S_2 : n(A_1 \cup A_2) = n(A_1) + n(A_2) - n(A_1 \cap A_2)$

Assume that ⊛ is true

$S_{n+1} : n(A_1 \cup A_2 \cup \ldots\ldots A_n \cup A_{n+1}$

$= n((A_1 \cup A_2 \cup \ldots\ldots A_n) \cup (A_{n+1}))$

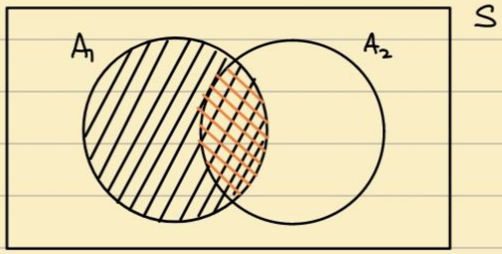
$= n(A_1 \cup A_2 \cup \ldots\ldots A_n) + n(A_{n+1}) - n((A_1 \cup A_2 \cup \ldots\ldots A_n) \cap (A_{n+1}))$

$= n(A_1 \cup A_2 \cup \ldots\ldots A_n) + n(A_{n+1}) - n((A_1 \cap A_{n+1}) \cup (A_2 \cap A_{n+1}) \cup \ldots\ldots \cup (A_n \cap A_{n+1}))$

→ Conditional Probability :

$P(A_1 | A_2)$

Assuming $A_2 \longrightarrow$ Happened , $P(A_1) = ?$



$$P(A_1 | A_2) = \frac{P(A_1 \cap A_2)}{P(A_2)}$$

→ Probability of A_1 given A_2

Ex. We consider the data from a software development company, where in 200 programs were written each week, in either C++ or Java with the following frequencies

	Compiles on first run	Does not compile on first run
C++	72	48
Java	64	16

(a) What is the probability that a random program compiles on the first run if it was written in C++.

Sol : Let E be the event that a random program run on the first compilation.

Let C++ be the event that a random program is written in C++.

$$P(E | C{+}{+}) = \frac{P(E \cap C{+}{+})}{P(C{+}{+})} = \frac{72/200}{120/200} = \frac{72}{120} = 9/15$$

(a) What is the probability that a random program compiles on the first run if it was written in Java

Sol : Let E be the event that a random program run on the first compilation.

Let Java be the event that a random program is written in Java.

$$P(E \mid Java) = \frac{P(E \cap Java)}{P(Java)} = \frac{64/200}{80/200} = \frac{64}{80} = 4/5$$

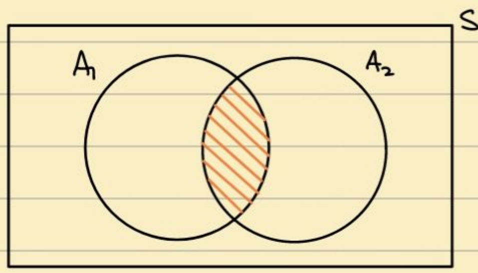
Let A_1, A_2 be two events. We say that the events are independent,
if $P(A_1 \mid A_2) = P(A_1)$
(or) $P(A_2 \mid A_1) = P(A_2)$

13/1

Conditional Probability:

The conditional Probability of an event A_2 given an event A_1 is

$$P(A_2 \mid A_1) = \frac{P(A_2 \cap A_1)}{P(A_1)}$$



$$P(A_2 \mid A_1) = \frac{n(A_2 \cap A_1)}{n(A_1)}$$

$$= \frac{n(A_2 \cap A_1)/n(S)}{n(A_1)/n(S)}$$

$$= \frac{P(A_2 \cap A_1)}{P(A_1)}$$

Independent events:

Two events A_1 and A_2 are independent if $P(A_2 \mid A_1) = P(A_2)$ [OR] $P(A_1 \mid A_2) = P(A_1)$

Note:

$$P(A_2 \mid A_1) = \frac{P(A_2 \cap A_1)}{P(A_1)}$$

$$\Rightarrow P(A_2) \cdot P(A_1) = P(A_2 \cap A_1)$$

Then, $$P(A_1 \mid A_2) = \frac{P(A_1 \cap A_2)}{P(A_2)}$$

$$= \frac{P(A_2) \cdot P(A_1)}{P(A_2)}$$

$$= P(A_1)$$

Ex:

As part of a promotional campaign, a store is offering a free Fitbit with the purchase of the new iPhone model. If 5% iPhones are faulty and 10% of the Fitbits are faulty, What is the probability that a customer gets both faulty iPhone and a faulty Fitbit ?

Solⁿ:

Let A_1 be the event that a customer gets a faulty Fitbit and
Let A_2 be the event that a customer gets a faulty iPhone .

$$
\begin{aligned}
P(A_1 \cap A_2) &= P(A_1) \cdot P(A_2) \\
&= 0.05 \times 0.10 \\
&= 0.005
\end{aligned}
$$

" Assumption is more important than the conclusion " - Dr. Subramani Sir [2025]

→ **Baye's Rule :**

If a sample S can be partitioned into k mutually exclusive and exhaustive events ; $A_1, A_2, \ldots A_k$, Then,

$$P(A_i|E) = \frac{P(E|A_i) \cdot P(A_i)}{P(A_1) \cdot P(E|A_1) + P(A_2) \cdot P(E|A_2) + \ldots + P(A_k) \cdot P(E|A_k)}$$

Total Probability Theorem : [Prerequisite]

If a sample S can be partitioned into k mutually exclusive and exhaustive events ; $A_1, A_2, \ldots A_k$, Then,

$$P(E) = P(A_1) \cdot P(E|A_1) + P(A_2) \cdot P(E|A_2) + \ldots + P(A_k) \cdot P(E|A_k)$$

$$
\begin{aligned}
P(E) &= P(E \cap S) \\
&= P(E \cap (A_1 \cup A_2 \cup A_3 \ldots \cup A_k)) \\
&= P((E \cap A_1) \cup (E \cap A_2) \cup \ldots \cup (E \cap A_k)) \\
&= P(E \cap A_1) + P(E \cap A_2) + \ldots + P(E \cap A_k)
\end{aligned}
$$

[∵ They are Mutually Exclusive Events ,]

$$= P(A_1) \cdot P(E|A_1) + P(A_2) \cdot P(E|A_2) + \ldots + P(A_k) \cdot P(E|A_k)$$

Now, coming to Baye's Theorem ,

$$P(A_i|E) = \frac{P(E|A_i) \cdot P(A_i)}{P(A_1) \cdot P(E|A_1) + P(A_2) \cdot P(E|A_2) + \ldots + P(A_k) \cdot P(E|A_k)}$$

$$LHS = P(A_i|E) = \frac{P(A_i \cap E)}{P(E)}$$

$$= \frac{P(E|A_i) \cdot P(A_i)}{P(E)}$$

$P(E) \rightarrow P(A_1) \cdot P(E|A_1) + P(A_2) \cdot P(E|A_2) + \ldots + P(A_k) \cdot P(E|A_k)$

→ Recall:

$f : S \rightarrow \mathbb{R}$

S → Sample Space

$\mathbb{R}$ → Real set

Look at functions on a set S to study the set.
& to understand our sample space.

f → Random variable

Random variable:

- Discrete Random variables ↓ Range of f = $1, 2, 3, 4, \ldots, n$ / $1, 2, 3, 4, \ldots$
- Continuous Random variable ↓ Range of f = $[a, b]$ / (a, b)

15/1

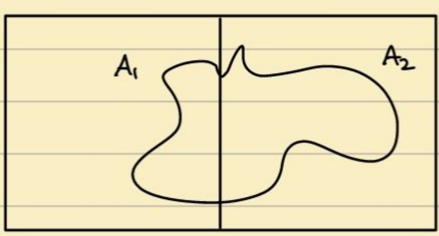
→ Example:

Recall our previous example

	Compiles on first run	Does not compile on first run
C++	72	48
Java	64	16

Find the probability that a random program compiles on the first run.

Sol^n:



S = 200

Let A_1 be the event that a random program is in C++ and let
A_2 be the event that a random program is in Java, &
Let E be the event that the program compiles on the first run.
By TLP,

$$P(E) = P(A_1)\, P(E|A_1) + P(A_2)\, P(E|A_2)$$

$$= \frac{120}{200} \times \frac{72}{120} + \frac{80}{200} \times \frac{64}{80}$$

$$= \frac{72}{200} + \frac{64}{200}$$

$$= \frac{136}{200} = \frac{17}{25}$$

$$\therefore P(E) = 0.68$$

- Example:

Inquiries to an online computer system arrive on five communication lines. The percentages of messages recieved through the different lines are:

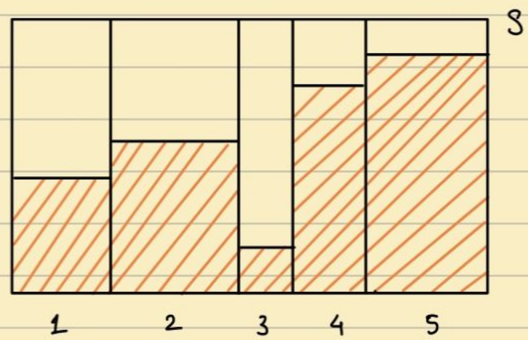
Line	1	2	3	4	5
% recieved	20	30	10	15	25

From past experiences, it is known that the % of messages exceeding 100 charecters are

Line	1	2	3	4	5
% exceeding 100 chars	40	60	20	80	90

What is the probability of messages exceeding 100 characters?

Sol^n:



Let E be the event that a random message exceeds 100 characters and A_i be the event that messaged recieved by the i^{th} channel.
$[i = 1, 2, 3, 4, 5]$

By TLP,

$$P(E) = P(A_1)\cdot P(E|A_1) + P(A_2)\cdot P(E|A_2) + P(A_3)\cdot P(E|A_3) + P(A_4)\cdot P(E|A_4) + P(A_5)\cdot P(E|A_5)$$

$$= \frac{20}{100} \times \frac{40}{100} + \frac{30}{100} \times \frac{60}{100} + \frac{10}{100} \times \frac{20}{100} + \frac{15}{100} \times \frac{80}{100} + \frac{25}{100} \times \frac{90}{100}$$

$$\therefore P(E) = 0.625$$

→ Example :

In a University department, 50% of documents are written in word, 30% in LaTeX and 20% in HTML.

From the past experience, it is known that,

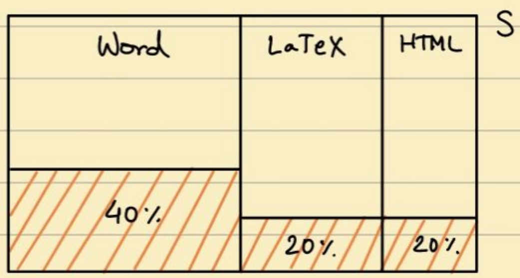
40% of the Word documents $\geq$ 10 pages

20% of the LaTeX documents $\geq$ 10 pages

20% of the HTML documents $\geq$ 10 pages

What is the probability that a random document is in LaTeX & has more than 10 pages ?

Sol^n :



Let E be the event that a random document has more than 10 pages.

$$P(LaTeX|E) = \frac{P(E|LaTeX)\cdot P(LaTeX)}{P(E)}$$

$$= \frac{P(E|LaTeX)\cdot P(LaTeX)}{P(E|LaTeX)\cdot P(LaTeX) + P(E|Word)\cdot P(Word) + P(E|HTML)\cdot P(HTML)}$$

$$= \frac{\frac{20}{100} \times \frac{30}{100}}{\frac{40}{100} \times \frac{50}{100} + \frac{40}{100} \times \frac{50}{100} + \frac{20}{100} \times \frac{20}{100}}$$

$$= \frac{0.06}{0.3} = 0.2$$

20/1

Random variables :

$X : S \rightarrow R$ is a random variable.

Random variable

Discrete Random variable | Continuous Random variable

- <u>Discrete random variable</u>:

A random variable X is discrete if it takes the values $a_1, a_2, \ldots\ldots, a_n$ where the set $\{a_1, a_2, \ldots\ldots, a_n\}$ is discrete.

Any finite set is a discrete set.

e.g. $1, 2, 3, \ldots\ldots, 10$

$$|a_i - a_j| \geq 1$$

Take $x_1, x_2, \ldots\ldots, x_n$

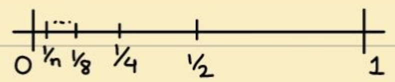
$$r = \min\{|x_i - x_j| \mid 1 \leq i, j \leq n\} \qquad [r : \text{non-negative}]$$

Then, $|x_i - x_j| \geq r$

e.g. $1, 2, 3, 4, 5, 6, 7, 8$

$$|a_i - a_j| \geq 1$$

e.g. $1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots\ldots, \frac{1}{n} \longrightarrow$ Not discrete



Let $\{a_n\}_{n=1}^{\infty}$ be a convergent sequence, Then set $\{a_1, a_2, \ldots\ldots, a_n\}$ is not discrete.

- <u>Continous random variable</u>:

A random variable X is said to be continuous if it takes the values $[a, b]$, $[a, b)$, $(a, b]$ OR (a, b)

Continuous set $\longrightarrow$ Connected set, i.e. no breaks

Let X be a discrete random variable & Assume that it takes the values $a_1, a_2, \ldots\ldots a_n$

- <u>Probability Density Function</u> (PDF) [OR] <u>Probability Mass Function</u> (PMF):

The PDF of X is a function

$$\rho : \mathbb{R} \longrightarrow \mathbb{R}_{\geq 0}$$

So that

$$\rho(x) = P(X = x) \quad [\text{if } x = a_i \text{ for some } i]$$

$$\rho(x) = P(X = x) = 0, \text{ otherwise}$$

Ex. Three balls to be randomly selected without replacement from an urn containing 20 balls numbered 1 through 20

Here, $S = \{(a, b, c) \mid 1 \leq a, b, c \leq 20 \ \& \ a \neq b \neq c\}$

I am interested in the largest number in the tuple.

Define :

$$X : S \longrightarrow \mathbb{R}$$

$$X((a, b, c) = \max\{a, b, c\}$$

Clearly, X is a random variable.

The possible values of X are $3, 4, 5, \ldots\ldots, 20$

As the set $\{3, 4, 5, \ldots\ldots, 20\}$ is discrete,

the random variable X is discrete.

What is the PDF of X ?

$$p(x) = \begin{cases} \dfrac{\binom{i-1}{2}}{\binom{20}{3}}, & \text{if } x = i, \\ & \quad \text{where } i = 3, 4, 5, \ldots\ldots 20 \\ 0, & x \neq 3, 4, 5, \ldots\ldots, 20 \end{cases} \qquad \begin{matrix} [p(X = x)] \\ \\ [p(X = x) = 0] \end{matrix}$$

$$\left[\begin{matrix} p(X = 3) & p(X = 4) & p(X = 5) \\ = \dfrac{\binom{2}{2}}{\binom{20}{3}} & = \dfrac{\binom{3}{2}}{\binom{20}{3}} & = \dfrac{\binom{4}{2}}{\binom{20}{3}} \end{matrix}\right]$$

- Cumulative Distribution Function (CDF)

  The CDF of X is a function

  $$F : \mathbb{R} \longrightarrow \mathbb{R}_{\geq 0}$$

  and is defined as

  $$F(X) = P(X \leq x)$$

– Expectation:

The expectation of X is

$$E[X] = \sum_{x} x \cdot P(x)$$

– Variance:

The variance of X is $Var(X) = E[(X - \mu)^2]$

i.e. $Var(X) = E\left[(X - E[X])^2\right]$

$= E\left[X^2 - 2 \cdot X \cdot E[X] + E^2[X]\right]$

20/1

Let X be a discrete random variable,

Then, X^2 is also a discrete random variable

As, if X takes the values $a_1, a_2, \ldots\ldots, a_n$,

X^2 takes the values $a_1^2, a_2^2, \ldots\ldots, a_n^2$

$$\left[\begin{aligned}(f+f)(x) &= f(x) + f(x)\\ f^2(x) &= (f \cdot f)(x) = f(x) \cdot f(x)\\ &= (f(x))^2\end{aligned}\right]$$

In general, take a positive integer n,

X^n is a discrete random variable

Further, if

$$h(t) = a_m t^m + a_{m-1} t^{m-1} + a_{m-2} t^{m-2} + \ldots\ldots + a_1 t + a_0,$$

where $a_0, a_1, a_2, \ldots\ldots, a_m \in \mathbb{R}$

$$h(X) = a_m X^m + a_{m-1} X^{m-1} + a_{m-2} X^{m-2} + \ldots\ldots + a_1 X + a_0$$

is also a discrete random variable.

In General,

$$S \xrightarrow{X} \mathbb{R} \xrightarrow{f} \mathbb{R}$$

Then $f \circ X$ is a discrete random variable.

→ Classification of Discrete Random variables:

- Bernoulli Random Variable:

A random variable X is said to be Bernoulli if it takes $0, 1$ with PDF.

$$p(0) = \beta, \qquad \beta \in [0,1]$$
$$p(1) = 1 - \beta$$

e.g. Let's toss a coin.

Let X be the no. of heads in the outcome

$S = \{H, T\}$

$X : S \to \mathbb{R}$ with values $0, 1$

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

$\therefore$ X is a Bernoulli Random variable.

- Expectation of a Bernoulli Random variable:

Let X be a Bernoulli Random variable

$$\begin{aligned}E[X] &= \sum_x x \cdot P(x)\\ &= 0 \cdot p(0) + 1 \cdot p(1)\\ &= 1 - \beta\end{aligned}$$

- Variance of Bernoulli Random variable

$= E[(X-\mu)^2]$

$= E[X^2 - 2X\mu + \mu^2]$

<u>H.W</u>

Let X and Y be two discrete random variables,

Then, $E[X+Y] = E[X] + E[Y]$

<u>21/1</u>

Let X and Y be two discrete random variables,

Then, $E[aX+bY] = aE[X] + bE[Y]$, $a, b \in \mathbb{R}$

<u>Sol^n</u>:

$X, Y : S \longrightarrow \mathbb{R}$

$aX, bY : S \longrightarrow \mathbb{R}$

$X \rightsquigarrow x_1, x_2, x_3, \ldots$

$Y \rightsquigarrow y_1, y_2, y_3, \ldots$

$aX + bY \rightsquigarrow aX_1 + bY_1, aX_2 + bY_2, \ldots$

$$E[aX+bY] = \sum (ax_i + by_i) \underset{ax+by}{P}(ax_i + by_i)$$

$$= a \sum_i x_i \cdot \underset{ax+by}{P}(ax_i + by_i) + b \sum_i y_i \cdot \underset{ax+by}{P}(ax_i + by_i)$$

Hence, original statement is not true.

Q) Let X be a discrete random variables,

Then, $E[aX+b] = aE[X] + b$; $a, b \in \mathbb{R}$

<u>Sol^n</u>:

$$E[aX+b] = \sum_i (ax_i + b) \underset{ax+b}{P}(ax_i + b)$$

$$= a \sum_i x_i \cdot \underset{ax+b}{P}(ax_i + b) + b \underbrace{\sum_i \underset{ax+b}{P}(ax_i + b)}_{1}$$

$X \rightsquigarrow x_i$

$aX + b \rightsquigarrow ax_i + b$

$$= a \sum_i x_i \cdot \underset{ax+b}{P}(ax_i + b) + b$$

$$\boxed{\therefore E[aX+b] = a \cdot E[X] + b}$$

Hence Proved

- Variance of Bernoulli Distribution

$$= E[(X-\mu)^2]$$

$$= E[\underbrace{X^2 - 2X\mu}_{y} + \underbrace{\mu^2}_{\text{Constant}}]$$

$$= E[X^2 - 2X\mu] + E[\mu^2]$$

$$= \sum_i (x_i^2 - 2\mu x_i) \cdot \underset{X^2-2\mu X}{P}(x_i^2 - 2\mu x_i) + E[\mu^2]$$

$$= \sum_i x_i^2 \cdot \underset{X^2-2\mu X}{P}(x_i^2 - 2\mu x_i) - 2\mu \sum_i x_i \cdot \underset{X^2-2\mu X}{P}(x_i^2 - 2\mu x_i) + E[\mu^2]$$

$$= \sum_i x_i^2 \cdot \underset{X^2}{P}(x_i^2) - 2\mu \sum_i x_i \cdot \underset{X}{P}(x_i) + E[\mu^2]$$

$$= E[X^2] - 2\mu E[X] + E[\mu^2]$$

$$= E[X^2] - 2\mu^2 + \mu^2 \qquad [E[X] = \mu]$$

$$= E[X^2] - \mu^2$$

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

Variance of a Bernoulli Distribution = $\frac{1}{2}$

$\because$ We know that

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

$$E[X] = \frac{1}{2}$$

Now,

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

$$= 0^2 \cdot \underset{X^2}{P}(0^2) + 1^2 \cdot \underset{X^2}{P}(1^2) - (E[X])^2$$

$$= 0 + \underset{X^2}{P}(1) - \mu^2$$

$$= P_X(1) - \mu^2$$

$$= (1-p) - (1-p)^2$$

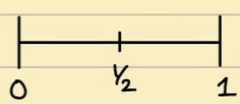
$$= 1 - p - (1 - 2p + p^2)$$

$$\therefore Var[X] = p(1-p)$$

$$\text{Standard deviation} = \sqrt{Var[X]}$$

$$= \sqrt{p(1-p)}$$

In case of tossing a coin, $p = \frac{1}{2}$



$\therefore$ Std. deviation = $\frac{1}{2}$ &

Variance = $\frac{1}{4}$

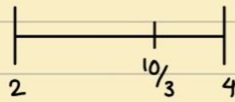
Q) $X \rightsquigarrow 2, 4$

$P(2) = 1/3$

$P(4) = 2/3$

X is Bernoulli ?

$E[X] = 2(1/3) + 4(2/3)$
$= 10/3$

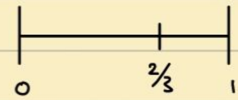


$X \rightsquigarrow 0, 1$

$P(0) = 1/3$

$P(1) = 2/3$

$E[X] = 0 \cdot 1/3 + 1 \cdot 2/3$
$= 2/3$



22/1

→ Binomial distribution:

A random variable X is said to be Binomial if X takes the values $0, 1, 2, \ldots\ldots, n$ where $n \in \mathbb{N}$ with the PDF

$$p(i) = P(X = i) = \binom{n}{i} \beta^i (1-\beta)^{n-i}, \quad \text{here } \beta \in [0, 1]$$

Here, we say that X is binomial with parameter (n, β)

E.g. Let's do an experiment with the possible outcomes Loss, Win. Suppose, we do that experiment n times & interested in the no. of wins. Let X be the no. of wins in the procedure (repeating the experiment n times)

Then, clearly X takes the values $0, 1, 2, \ldots\ldots, n$

$$p(0) = P(X = 0) = (1-\beta)^n$$
$$p(1) = P(X = 1) = \beta(1-\beta)^{n-1} \times {}^nC_1$$

Win $\rightarrow \beta$

Loss $\rightarrow 1-\beta$

$$\begin{array}{ccccc} \square & \square & \square \ldots\ldots & \square \\ \end{array}$$
$$\left.\begin{array}{ccccc} W & L & L & L & L \\ L & W & L & L & L \\ L & L & L & L & W \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{array}\right\} {}^nC_1$$

$$p(2) = P(X = 2) = {}^nC_2 \times \beta^2 (1-\beta)^{n-2}$$
$$p(i) = P(X = i) = \binom{n}{i} \beta^i (1-\beta)^{n-i}$$

$[i = 3, 4, 5, \ldots\ldots]$

Binomial Theorem:

$$(a+b)^n = \sum {}^nC_r a^r b^{n-r}$$
$$1 = 1^n = (\beta + (1-\beta))^n = \sum \underbrace{{}^nC_r \beta^r (1-\beta)^{n-r}}_{T_n}$$
$$\boxed{T_n = \binom{n}{i} \beta^i \cdot (1-\beta)^{n-i}}$$

H.W: Let X be a Binomial Distribution,

(1) Find CDF of X

(2) Find $E[X]$

(3) Find $Var[X]$

→ Poisson Random Variable:

A random variable is said to be Poisson random variable with parameter $\lambda > 0$, if it takes the values $0, 1, 2, \dots\dots, n$ with the PDF

$$p(i) = P(X=i) = \frac{e^{-\lambda} \cdot \lambda^i}{i!}$$

Remark:

Let X be a binomial distribution (n, β),

$$p(i) = \binom{n}{i} \beta^i (1-\beta)^{n-i} \qquad i = 0, 1, 2, \dots$$

$$= \frac{n!}{(n-i)!\, i!} \beta^i (1-\beta)^{n-i}$$

$[\lambda = n\beta]$

$$= \frac{n!}{(n-i)! \cdot i!} \left(\frac{\lambda}{n}\right)^i \left(1 - \frac{\lambda}{n}\right)^{n-i}$$

$$= \frac{n(n-1)(n-2) \dots\dots (n-(i-1)) \dots\dots 1}{n^i} \cdot \frac{\lambda^i}{i!} \cdot \frac{(1-\lambda/n)^n}{(1-\lambda/n)^i}$$

$$= (1-\lambda/n)^n \cdot \frac{\lambda^i}{i!} \cdot \frac{n(n-1)\cdot(n-2)\dots\dots(n-i+1)}{n^i \cdot (1-\lambda/n)^i}$$

$$= (1-\lambda/n)^n \cdot \frac{\lambda^i}{i!} \frac{(1-1/n)(1-2/n)\dots\dots\left(1-\frac{i+1}{n}\right)}{(1-\lambda/n)^i}$$

Apply $n \longrightarrow \infty$

$$= e^{-\lambda} \cdot \frac{\lambda^i}{i!} \cdot \frac{1}{1}$$

$$= \frac{e^{-\lambda} \cdot \lambda^i}{i!} \qquad \text{[Approximation]}$$

23/1

→ Recall:

A random variable is said to be Poisson random variable with parameter $\lambda > 0$, if it takes the values $0, 1, 2, \dots\dots, n$ with the PDF

$$p(i) = P(X=i) = \frac{e^{-\lambda} \cdot \lambda^i}{i!}$$

$$E[X] = \sum_x x \cdot P(x)$$

$$= \sum_{i=0}^{\infty} i \cdot P(i)$$

$$= \sum_{i=0}^{\infty} i \cdot \frac{e^{-\lambda} \cdot \lambda^i}{i!}$$

$$= e^{-\lambda} \cdot \sum_{i=1}^{\infty} \lambda \cdot \frac{\lambda^{i-1}}{(i-1)!}$$

Let $j = i-1$

$$= e^{-\lambda} \cdot \lambda \cdot \sum_{j=0}^{\infty} \frac{\lambda^j}{j!}$$

$$= e^{-\lambda} \cdot \lambda \cdot e^{\lambda}$$

$$= \lambda$$

$$\boxed{\therefore\ E[X] = \lambda}$$

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

$$\lambda\left(1 - \frac{\lambda}{n}\right)$$

In Binomial:

$E[X] = np$

$Var[X] = np(1-p)$

[$np = \lambda$ in case of Poisson]

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

$$E[X^2] = \sum_x x^2 \cdot P(x^2)$$

$$= \sum_{i=0}^{\infty} i^2 \cdot p(i^2)$$

$$= \sum_{i=0}^{\infty} i^2 \cdot P(i)$$

$$P(X^2 = i^2) = \frac{e^{-\lambda} \cdot \lambda^{i^2}}{(i^2)!} \Rightarrow P(i^2)$$

$$P(X = i) = \frac{e^{-\lambda} \cdot \lambda^i}{i!} \Rightarrow P(i)$$

$$= \sum_{i=0}^{\infty} i^2 \cdot \frac{e^{-\lambda} \cdot \lambda^i}{i!}$$

$$= e^{-\lambda} \cdot \sum_{i=1}^{\infty} i \cdot \frac{\lambda \cdot \lambda^{i-1}}{(i-1)!}$$

$$= \lambda \cdot e^{-\lambda} \cdot \sum_{i=2}^{\infty} \frac{\lambda \cdot \lambda^{i-2}}{(i-2)!} \cdot \frac{i}{i-1}$$

$$= \lambda^2 e^{-\lambda} \cdot \sum_{i=2}^{\infty} \frac{\lambda^{i-2}}{(i-2)!} \cdot \frac{i}{i-1}$$

$$= \lambda^2 e^{-\lambda} \cdot \sum_{i=2}^{\infty} \frac{\lambda^{i-2}}{(i-2)!} \cdot \left(1 + \frac{1}{i-1}\right)$$

$$= \lambda^2 e^{-\lambda} \cdot \sum_{i=2}^{\infty} \left(\frac{\lambda^{i-2}}{(i-2)!} + \frac{1}{i-1} \cdot \frac{\lambda^{i-2}}{(i-2)!}\right)$$

$$= \lambda^2 \cdot e^{-\lambda} \cdot e^{\lambda} + \lambda^2 e^{-\lambda} \cdot \sum_{i=2}^{\infty} \left(\frac{1}{i-1} \cdot \frac{\lambda^{i-2}}{(i-2)!}\right)$$

$$= \lambda^2 +$$

—— Will be proved later on ——

27/1

Let X be a binomial distribution with parameters (n, p), Then,

$$P(X=i) = {}^nC_i \cdot p^i (1-p)^{n-i}$$

$$= \frac{n(n-1)\cdots\cdots(n-i+1)}{n^2} \cdot \frac{\lambda^i}{i!} \cdot \frac{(1-\lambda/n)^n}{(1-\lambda/n)^i}$$

Ex. The no. of misprints in a group of pages of a book

Ex. The no. of people in a committee living upto 100 years of age

Ex. The no. of wrong telephone numbers that are dialed in a day.

Ex. The no. of packages of dog biscuits sold in a particular store each day.

Ex. The no. of customers entering a post office on a given day.

Ex. The no. of vacancies occuring during an year in the supreme court.

Ex. The no. of α-particles discharged in a fixed period of time from some radioactive material.

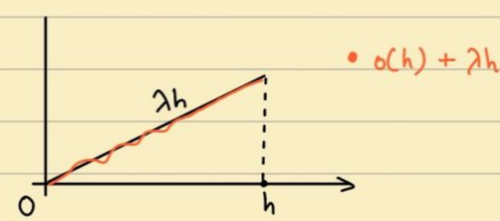
- Let us assume that the events are occuring at certain points of time.
  Assuming the following:
  (i) The probability that exactly 1 event occurs in a given interval of length h is $\lambda h + o(h)$, where $\lambda > 0$
  Here, $o(h) = f(h)$, with $\frac{f(h)}{h} \longrightarrow 0$ as $h \to 0$

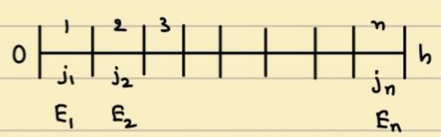
  (ii) The probability that 2 or more events occur in an interval of length h is 0.

  (iii) For any integer n, $J_1, J_2, \ldots\ldots J_n$, any set of n non overlapping intervals, let E_i be the event that exactly j_i events occur in the i^{th} subinterval, Then $E_1, E_2, \ldots\ldots, E_n$

E.g. The no. of earthquakes occuring during some fixed time span.
Let's take $h \to 3$ weeks
i.e. $h = 21$

$o(h) \longrightarrow$ error

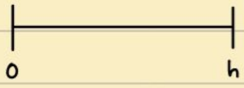




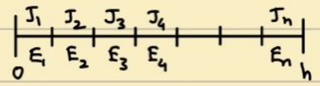
28/1

Recall:

(1) $\lambda h + o(h)$



(2) Probability ($2 \geq$ events) $= 0$

(3) 

$E_1, E_2, \ldots\ldots, E_n$ are independent.

- Let $N(h)$ be the no. of events occuring in that interval.

  Clearly, $N(h)$ takes the values $0, 1, 2, 3, 4, \ldots\ldots$

  We want to know, $k \in \mathbb{N} \cup \{0\}$,

$$P(N(h) = k) = P(k \text{ of the } n \text{ subintervals contains event and the other } n-k \text{ contain } 0 \text{ events}) + P(N(h) = k \text{ and atleast } 1 \text{ subinterval contains } 2 \text{ or more events})$$

Let B be the event that $N(h) = k$ and atleast one subinterval containing 2 or more events

$$P(B) \leq P(\text{atleast one subinterval containing 2 or more events})$$

$$= P\left(\bigcup_{i=1}^{n} \{i^{th} \text{ subinterval containing 2 or more events}\}\right)$$

$$\leq \sum P(i^{th} \text{ subinterval containing 2 or more events})$$

$$= \sum_{i=1}^{n} o(h/n)$$

$$= n \cdot o(h/n) = h\left(\frac{o(h/n)}{h/n}\right) \longrightarrow 0 \text{ as } n \longrightarrow \infty$$

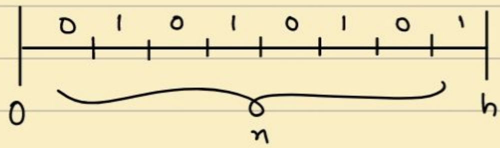
$$P(B) \longrightarrow 0 \text{ as } n \longrightarrow \infty$$

$$\left[\begin{array}{c} \text{As } n \to \infty,\ h/n \to 0,\ \text{Let } f = h/n \\ \because t \to 0,\ \therefore \frac{o(t)}{t} = \frac{f(t)}{t} \to 0 \end{array}\right] \Rightarrow P(B) \longrightarrow 0$$

We change the 2^{nd} assumption:

(ii) The probability that 2 or more events occur in an interval of length h is $o(h)$.

$P(A) = P($k subintervals containing exactly 1 event and the other n-k containing 0 events$)$



$$P(A) = \binom{n}{k} \cdot \left[\lambda \frac{h}{n} + o\left(\frac{h}{n}\right)\right]^k \left(1 - \frac{\lambda h}{n} - o\left(\frac{h}{n}\right)\right)^{n-k}$$

$$= \frac{n!}{k!\,(n-k)!} \left(\frac{\lambda h}{n} + o\left(\frac{h}{n}\right)\right)^k \left(1 - \frac{\lambda h}{n} - o\left(\frac{h}{n}\right)\right)^{n-k}$$

$$= \frac{n(n-1)\dots(n-k+1)}{k!} \cdot \frac{h^k}{nk} (\lambda h + o(1))^k \left(1 - \frac{\lambda h}{n} - o\left(\frac{h}{n}\right)\right)^{n-k}$$

$$= \left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\dots\left(n - \frac{k+1}{n}\right) \frac{h^k}{k!} (\lambda h + o(1))^k \left(1 - \frac{\lambda h}{n} - o\left(\frac{h}{n}\right)\right)^{n-k}$$

Cont. in next class

29/1

$$P(A) = \binom{n}{k} \cdot \left[\lambda \frac{h}{n} + o\left(\frac{h}{n}\right)\right]^k \left(1 - \frac{\lambda h}{n} - o\left(\frac{h}{n}\right)\right)^{n-k}$$

$$= \frac{n(n-1)(n-2)\dots(n-k+1)}{k!} \cdot \left[\lambda \frac{h}{n} + o\left(\frac{h}{n}\right)\right]^k \left(1 - \frac{\lambda h}{n} - o\left(\frac{h}{n}\right)\right)^{n-k}$$

$$= n\left[\lambda \frac{h}{n} + o\left(\frac{h}{n}\right)\right] \cdot (n-1)\left[\lambda \frac{h}{n} + o\left(\frac{h}{n}\right)\right] \cdot (n-2)\left[\lambda \frac{h}{n} + o\left(\frac{h}{n}\right)\right] \times \dots$$
$$\cdot (n-k+1)\left[\lambda \frac{h}{n} + o\left(\frac{h}{n}\right)\right] \cdot \left(1 - \frac{\lambda h}{n} - o\left(\frac{h}{n}\right)\right)^{n-k}$$

$$= \frac{(\lambda h)^k}{k!} e^{-\lambda h} \quad \text{as} \quad n \to \infty$$

- Suppose Earthquakes, Assumptions ①, ② & ③, Given $\lambda = 2/\text{week}$

(a) Find the probability that atleast 3 earthquakes occuring in the next 2 weeks

Sol:

$P(N(2) \geq 3)$

$= 1 - P(N(2) \leq 2)$

$= 1 - P(N(2) = 0) - P(N(2) = 1) - P(N(2) = 2)$

$= 1 - e^{-2\cdot 2} \cdot \frac{(2\cdot 2)^0}{0!} - e^{-4} \cdot \frac{4^1}{1!} - e^{-4} \cdot \frac{4^2}{2!}$

- **The Geometric Random variable:**

A random variable X is said to be a Geometric Random variable if it takes the values $1, 2, 3, 4, \ldots$ with the PDF

$$P(X=n) = (1-p)^{n-1} \cdot p \quad , \quad 0<p<1$$
$$n = 1, 2, 3, 4, \ldots$$

Suppose, we do independent trials with probability p, $0<p<1$, are performed until a success occurs.

Let X be the no. of trials required.

X takes the values $1, 2, 3, \ldots$

$$p(n) = P(X=n) = \underbrace{(1-p)^{n-1}}_{\text{Probability of failure}} \cdot \underbrace{p}_{\text{Probability of Success}}$$

$n = 1, 2, 3, 4, \ldots$

H.W

Q) Find the Expectation of a Geometric Random variable

Q) Find the Variance of a Geometric Random variable

Q) Find the CDF of a Geometric Random variable

- **The negative Binomial Random variable:**

Fix r

How many trials require until r success?

A random variable X is negative Binomial Random Variable if X takes the values $r, r+1, r+2, \ldots, r+n$ $(r \in \mathbb{N})$ with PDF

$$p(n) = P(X=n) = {}^{n-1}C_{r-1} (1-p)^{n-r} \cdot p^r$$
$$n = r, r+1, r+2, \ldots$$

Sol:

Let X be the no. of trials required to achieve r successes:

X takes the values $r, r+1, r+2, \ldots$

$$p(n) = P(X=n)$$
$$= \binom{n-1}{r-1} p^{r-1} (1-p)^{(n-1)-(r-1)}$$
$$= \binom{n-1}{r-1} p^{r-1} (1-p)^{n-r} \cdot \underbrace{p}_{\text{Success already fixed}}$$

Alternatively,

The Probability that X (no. of failures before the r^{th} success) takes a value x is :

$$P(X=x) = \binom{x+r-1}{r-1} p^r (1-p^x) \qquad [x = 0, 1, 2, \ldots]$$

- Expectation of a Negative Binomial Random variable :

$$E[X] = \frac{r(1-p)}{p}$$ (If X counts no. of failures before r successes)

$$E[X] = \frac{r}{p}$$ (If X counts total trials)

Proof:

$$E[X] = \sum_{r}^{\infty} n \cdot P(n)$$

$$\Rightarrow E[X] = \sum_{n=r}^{\infty} n \cdot \binom{n-1}{r-1} p^r (1-p)^{n-r}$$

$$X = Y_1 + Y_2 + \ldots + Y_r$$

$$E[Y_i] = 1/p \quad [\because Y_i \sim \text{Geometric}(p)]$$

$$\therefore E[X] = E[Y_1] + E[Y_2] + \ldots + E[Y_r]$$
$$= 1/p + 1/p + \ldots + 1/p$$
$$= r/p$$

- Variance of a Negative Binomial Random variable :

$$Var[X] = \frac{r(1-p)}{p^2}$$ (X represents the total no. of trials)

Proof:

$$E[X^2] = \sum_{n=r}^{\infty} n^2 P(n)$$

$$\Rightarrow E[X^2] = \frac{r(1-p)+r^2}{p^2}$$

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

$$= \frac{r(1-p)+r^2}{p^2} - \left(\frac{r}{p}\right)^2$$

$$= \frac{r(1-p)}{p^2}$$

- Hyper Geometric Random variable:

A random variable X is said to be a Hyper Geometric Random variable if it takes the values $0, 1, 2, 3, 4, \ldots$ with the PDF

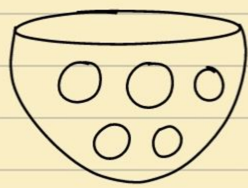
$$p(i) = P(X = i) = \frac{\binom{m}{i}\binom{n-m}{n-i}}{\binom{N}{n}}, \quad \text{for some } m, n, N \in \mathbb{N}$$
$$i = 0, 1, 2, \ldots, n$$

A random variable X is said to have zeta distribution if it takes the values $1, 2, 3, 4, \ldots$ with the PDF

$$p(k) = P(X = k) = \frac{c}{k^{\alpha+1}} \qquad k = 1, 2, 3, \ldots$$
for some $\alpha > 0$, $c \in \mathbb{R}$

$$\left[\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}, \; s \in \mathbb{C}\right]$$

30/1



N white balls
M black balls

Select a ball randomly with replacement one at a time, until a black one is obtained. What is the probability that exactly n draws are needed?

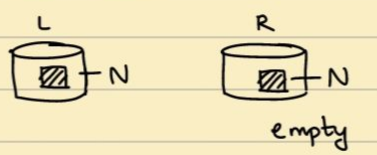
Sol^n:

Let X be the no. of drawns needed until a black ball is obtained. Clearly, X takes the values $1, 2, \ldots$

$$P(\text{obtaining a black ball}) = \frac{M}{M+N}$$

$$p(n) = p(X = n)$$
$$= \left(1 - \frac{M}{M+N}\right)^{n-1} \left(\frac{M}{M+N}\right)$$

The Banach match Problem:



Solⁿ Let X be the no. of choices until the R pocket is empty & the L pocket has k matches

$P(X = N+1+N-k)$

X takes the values $N, N+1, N+2, \dots$

$$= \binom{N+1+N-k+1}{N+1-1}\left(\frac{1}{2}\right)^{2N-k+1}$$

for $p = 1/2$, $r = N+1$

$n = 2N-k+1$

3/2

→ **Continuous random variable**

A random variable X is continous if $X : S \longrightarrow \mathbb{R}$

$R(X) = [a,b], (a,b), (a,b], [a,b)$

& there exists a function $f : \mathbb{R} \longrightarrow \mathbb{R}_{\geq 0}$

S.T.

$$P(X \in A) = \int_A f(x)\, dx \quad \text{[Riemann Integral]}$$

where $A \subseteq \mathbb{R}$

for us, A is always $[a,b], (a,b), [a,b), (a,b], (a,\infty)$, $[a,\infty), (-\infty,a)$ OR $(-\infty,a]$

Ex. Suppose X is a continuous random variable whose pdf is given by

$$f(x) = \begin{cases} c(4x-2x^2), & 0<x<2 \\ 0, & \text{otherwise} \end{cases}$$

(a) What is the value of c?

Solⁿ:

$$P(X \in (-\infty,\infty)) = \int_{-\infty}^{\infty} f(x)\,dx \overset{\text{set}}{=\!=} 1$$

$$= \int_{-\infty}^{0} 0\, dx + \int_0^2 c(4x-2x^2)\,dx + \int_2^\infty 0\,dx$$

$$= \int_0^2 c(4x-2x^2)\,dx$$

$$= c\left[\int_0^2 (4x)\,dx - \int_0^2 2x^2\,dx\right]$$

$$= c\left(2x^2\right]_0^2 - c\left(\frac{2}{3}x^3\right]_0^2$$

$$= 8c - 16/3\, c$$

$$\Rightarrow c\left(8-\frac{16}{3}\right) = 1$$

$$\Rightarrow \boxed{c = 3/8}$$

(b) what is Probability $P(X>1)$

$$\begin{aligned}
P(X>1) &= P(X \in (1,\infty)) \\
&= \int_1^{\infty} f(x)\,dx \\
&= \int_1^{2} c(4x-2x^2)\,dx + \int_2^{\infty} 0\,dx \\
&= c\left(2x^2 - \frac{2}{3}x^3\right]_1^2 \\
&= c\left(\left(8-16/3\right)-\left(2-2/3\right)\right) \\
&= c(6-14/3) \\
&= 4/3 \cdot c
\end{aligned}$$

As $c = 3/8$,

$$P(X>1) = 4/3 \cdot 3/8 = \underline{\underline{1/2}}$$

- Let X be a continuous random variable with the PDF f, Then the expectation of X is

$$E[X] = \int_{-\infty}^{\infty} x \cdot f(x)\,dx$$

& the variance is

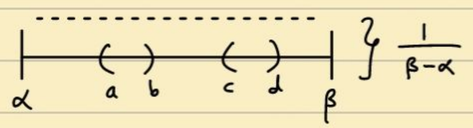
$$Var[X] = E[X^2] - (E[X])^2$$

4/1

(1) Uniform Continuous Random Variable:

A continuous random variable X is said to be a Uniform continuous random value if X takes the values in $[\alpha,\beta]$, $\alpha,\beta \in \mathbb{R}$ with the PDF

$$f(x) = \frac{1}{\beta-\alpha} \quad \text{if } \alpha \le x \le \beta$$



Let $A = (a,b)$

$$P(X \in A) = \int_A f(x)\,dx = \int_a^b \frac{1}{\beta-\alpha}\,dx = \frac{b-a}{\beta-\alpha}$$

Let $B = (c, d)$

$$P(X \in B) = \int_B f(x)\, dx = \int_c^d \frac{1}{\beta - \alpha}\, dx = \frac{d-c}{\beta - \alpha}$$

$\therefore\ P(X \in A) = P(X \in B)$

$$\Rightarrow \frac{b-a}{\beta - \alpha} = \frac{d-c}{\beta - \alpha} \quad \because \text{ Equal Intervals, i.e. } d-c = b-a$$

Expectation:

Let X be a continuous uniform random variable on $[\alpha, \beta]$

$$E[X] = \int_{-\infty}^{\infty} x \cdot f(x)\, dx$$

$$= \int_{\alpha}^{\beta} x \cdot \frac{1}{\beta - \alpha}\, dx$$

$$= \frac{1}{\beta - \alpha}\left(\frac{x^2}{2}\right]_{\alpha}^{\beta}$$

$$= \frac{\beta^2 - \alpha^2}{2} \cdot \frac{1}{\beta - \alpha}$$

$$\therefore\ E[X] = \frac{\alpha + \beta}{2}$$

Variance:

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

$$S \longrightarrow \mathbb{R} \xrightarrow[\text{cont.}]{f} \mathbb{R}_{\geq 0} \quad : \quad f \circ X : S \xrightarrow{\text{cont.}} \mathbb{R}_{\geq 0}$$

$$E[X^2] = \int_{-\infty}^{\infty} x^2 \cdot g(x)\, dx \qquad [\, g(x) \rightarrow \text{PDF of } X^2 \,]$$

$$= \int_{-\infty}^{\infty} x^2 \cdot \frac{1}{\beta - \alpha}\, dx \qquad \text{PDF of } x = \text{PDF of } x^2$$

$$= \frac{1}{\beta - \alpha} \cdot \frac{\beta^3 - \alpha^3}{3}$$

$$\therefore\ E[X^2] = \frac{\alpha^2 + \alpha\beta + \beta^2}{3}$$

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

$$Var[X] = \frac{\alpha^2+\alpha\beta+\beta^2}{3} - \frac{\alpha^2+2\alpha\beta+\beta^2}{4}$$

$$\boxed{\therefore \quad Var[X] = \frac{(\beta-\alpha)^2}{12}}$$

5/1

Let X be a continuous uniform random variable on (α, β), $\alpha, \beta \in \mathbb{R}$,
The PDF of X is

$$f(x) = \begin{cases} \frac{1}{\beta-\alpha} & \text{if } \alpha < x < \beta \\ 0 & \text{otherwise} \end{cases}$$

Then, X^2 is also a continuous random variable in (α, β)

Sol^n

$X \to$ Uniform	$X^2 \to$ Uniform
$X : S \to (\alpha, \beta)$	$X^2 : S \to (\alpha^2, \beta^2)$
$f(x) = \frac{1}{\beta-\alpha}$	$S \xrightarrow{X} (\alpha, \beta) \xrightarrow{X^2} (\alpha^2, \beta^2)$ or (β^2, α^2)

$$E[X^2] = \int_{\alpha}^{\beta} x^2 \cdot f(x)\, dx$$

$$\left[\begin{array}{l} X : \text{Continuous} \\ \quad X : S \to (a, b) \\ \quad E[X] = \int_{-\infty}^{\infty} x \cdot f(x)\, dx \end{array} \right]$$

Let's recall, CDF for continuous random variable. Let X be a continuous random variable with PDF $f(x)$, Then

$$F(y) = \int_{-\infty}^{\infty} f(x)\, dx \qquad \text{Assuming } f(x) \text{ to be continuous.}$$

Then, differentiate the function w.r.t y

$$\frac{d}{dy} F(y) = \frac{d}{dy}\left(\int_{-\infty}^{\infty} f(x)\, dx\right) = f(y)$$

[By Fundamental theorem of Calculus]

Let X be a uniform random variable over $(0,1)$

Let $Y = X^2$

Clearly Y takes the values in $(0,1)$,

Then the PDF of Y,

$$F_Y(y) = P(Y \leq y) \qquad [0 \leq y \leq 1]$$
$$= P(X^2 \leq y)$$
$$= P(X \leq \sqrt{y})$$

$$\Rightarrow F_Y(y) = F_X(\sqrt{y})$$

$$\frac{d}{dy} F_Y(y) = \frac{d}{dx} F_X(\sqrt{y}) = \frac{d}{dx} \int_{-\infty}^{y^{1/2}} f(x)\, dx$$

6/1

Let X be a uniform random variable over $(0,1)$

Consider X^n, $n \in \mathbb{N}$

Clearly X^n takes the values in $(0,1)$

Claim: To find PDF of X^n

Sol: Let $Y = X^n$

$$F_Y(x) = P(Y \leq x) \qquad \forall x \in (0,1)$$
$$= P(X^n \leq x)$$
$$= P(X \leq \sqrt[n]{x})$$

$$\Rightarrow F_Y(x) = F_X(\sqrt[n]{x})$$

$$\frac{d}{dx} F_Y(x) = \frac{d}{dx} F_X(\sqrt[n]{x})$$

$$= \frac{1}{n} x^{1/n - 1}, \quad 0 \leq x \leq 1$$

X is uniform in $(0,1)$

$$F_X(x) = \int_{-\infty}^{x} f(t)\, dt$$
$$= \int_{-\infty}^{x} 1 \cdot dt = \int_{-\infty}^{x} dt$$

$$\left(f(t) = \frac{1}{\beta - \alpha}, \text{ in } (\alpha, \beta) \quad \text{\& Here, } \beta - \alpha = 1 - 0 = 1 \right)$$

Thus, the PDF of X^n is

$$f(x) = \begin{cases} \frac{1}{n} x^{\frac{1}{n} - 1}, & 0 \leq x \leq 1 \\ 0, & \text{otherwise} \end{cases}$$

For $n=2$, the PDF of $y = X^2$ is $\frac{1}{2}x^{\frac{1}{2}-1} = \frac{1}{2\sqrt{x}}$

- Let X be a uniform random variable on (α, β). Consider $Y = X^n$ $(n \in \mathbb{N})$
  Find the PDF of X^n

X is a uniform random variable in (α, β)

Take $Y = X^2$

$\Rightarrow$ Y is a continuous random variable on (α^2, β^2)

$$F_Y(y) = P(Y \leq y)$$
$$= P(X^2 \leq x^2)$$

$$\begin{tikzcd}
(\alpha, \beta) \arrow[r, "\frac{x-\alpha}{\beta-\alpha}"] \arrow[d, Rightarrow] & (0,1) \\
(\alpha^2, \beta^2) \arrow[r, "\frac{x-\alpha^2}{\beta^2-\alpha^2}"] & (0,1)
\end{tikzcd}$$

Let $Y = X^n$, here X is uniform on (α, β)

$$F_Y(x) = P(Y \leq x)$$
$$= P(X^n \leq x)$$
$$= P(X \leq \sqrt[n]{x})$$

$$\frac{d}{dx} F_Y(x) = \frac{d}{dx} F_X(\sqrt[n]{x})$$

$$= F_X(x^{1/n}) \cdot \frac{1}{n} x^{1/n - 1}$$

$$\boxed{\begin{array}{l}\text{The PDF of } Y = X^n \text{ is} \\ \qquad F_X(x^{1/n}) \cdot \frac{1}{n} x^{1/n - 1}\end{array}}$$

$$E[X^2] = \int_{-\infty}^{\infty} x \cdot F_X(x^{1/2}) \cdot \frac{1}{2} \cdot x^{-1/2}$$

$$= \frac{1}{2} \int_{\alpha^2}^{\beta^2} x \cdot \frac{1}{\beta - \alpha} \cdot \frac{1}{\sqrt{x}}$$

$$= \frac{1}{2(\beta - \alpha)} \int_{\alpha^2}^{\beta^2} \sqrt{x}\, dx$$

$$= \frac{1}{2(\beta-\alpha)} \cdot \left(\frac{x^{3/2}}{3/2}\right]_{\alpha^2}^{\beta^2}$$

$$= \frac{1}{2(\beta-\alpha)} \cdot \frac{2}{3} \cdot \left(\beta^{\frac{3\cdot 2}{2}} - \alpha^{\frac{3\cdot 2}{2}}\right)$$

$$= \frac{1}{3(\beta-\alpha)} \cdot (\beta^3 - \alpha^3)$$

$$= \frac{1}{3(\beta-\alpha)} (\beta^2 + \alpha\beta + \alpha^2)(\beta - \alpha)$$

$$\therefore E[x] = \frac{\alpha^2 + \alpha\beta + \beta^2}{3}$$

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

$$= \frac{\alpha^2 + \alpha\beta + \beta^2}{3} - \left(\frac{\alpha^2 + \alpha\beta + \beta^2}{3}\right)^2$$

$$= \frac{(\alpha-\beta)^2}{12}$$

11/2

→ Exponential random variable:

A continuous random variable X is said to be exponential with parameter λ if its PDF is the following

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

Expectation:

Graph:

$$E[X] = \int_{-\infty}^{\infty} x \cdot f(x)\, dx$$

$$= \int_{-\infty}^{0} x \cdot 0 \cdot dx + \int_{0}^{\infty} x \cdot \lambda e^{-\lambda x}\, dx$$

$$= \lambda \cdot \left[\int_{-\infty}^{0} x \cdot 0\, dx + \int_{0}^{\infty} x \cdot e^{-\lambda x}\, dx\right]$$

$$= \lambda \left(x \cdot \frac{-1}{\lambda} e^{-\lambda x} - \frac{1}{\lambda^2} \cdot e^{-\lambda x}\right]_0^{\infty}$$

$$= \left(-x \cdot e^{-\lambda x} - \frac{1}{\lambda} e^{-\lambda x}\right]_0^{\infty}$$

D	I
$+\ x$	$e^{-\lambda x}$
$-\ 1$	$\frac{-1}{\lambda} e^{-\lambda x}$
$+\ 0$	$\frac{1}{\lambda^2} e^{-\lambda x}$

$$= \lim_{x\to\infty} \left(-x \cdot e^{-\lambda x} - \frac{1}{\lambda} e^{-\lambda x}\right) - \lim_{x \to 0} \left(-x \cdot e^{-\lambda x} - \frac{1}{\lambda} e^{-\lambda x}\right)$$

$$= -\lim_{x\to\infty} \left(x \cdot e^{-\lambda x}\right) + \lim_{x\to 0} \left(x e^{-\lambda x} + \frac{1}{\lambda} e^{-\lambda x}\right)$$

$$= 0 + \frac{1}{\lambda}$$

$$\boxed{E[x] = \frac{1}{\lambda}}$$

$$E[x^2] = 2/\lambda^2$$

<u>H.W</u> Prove that $Var[X] = 1/\lambda^2$, Given $E[x] = 2/\lambda^2$

$$E[X^2] = \int_{-\infty}^{\infty} x^2 \cdot f(x)\, dx$$

$$= \int_{-\infty}^{0} 0 \cdot dx + \int_{0}^{\infty} x^2 \cdot \lambda \cdot e^{-\lambda x} dx$$

$$= \lambda \cdot \int_{0}^{\infty} x^2 \cdot e^{-\lambda x} dx$$

$$= \lambda \left[\frac{-x^2}{\lambda} \cdot e^{-\lambda x} - \frac{2x}{\lambda^2} \cdot e^{-\lambda x} - \frac{2}{\lambda^3} e^{-\lambda x} \right]_0^{\infty}$$

$$= \left(x^2 \cdot e^{-\lambda x} + \frac{2x}{\lambda} e^{-\lambda x} + \frac{2}{\lambda^2} e^{-\lambda x} \right]_{\infty}^{0}$$

$$= 2/\lambda^2 - \lim_{x \to \infty} \left[\underbrace{(e^{-\lambda x})}_{\text{Rate of growth is more}} \left(x^2 + 2x/\lambda + 2/\lambda^2 \right) \right]$$

D	I
$+ \; x^2$	$e^{-\lambda x}$
$- \; 2x$	$\frac{-1}{\lambda} e^{-\lambda x}$
$+ \; 2$	$\frac{1}{\lambda^2} e^{-\lambda x}$
$- \; 0$	$\frac{-1}{\lambda^3} e^{-\lambda x}$

$$\boxed{\therefore E[X^2] = 2/\lambda^2}$$

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

$$= 2/\lambda^2 - 1/\lambda^2$$

$$= 1/\lambda^2$$

$$\boxed{\therefore Var[X] = 1/\lambda^2}$$

- Example :

Suppose that the duration of a phone call in minutes is an exponential random variable with parameter $\lambda = \frac{1}{2}$. If someone answers immediately ahead of you at a public telephone booth, find the probability that you will have to wait.

(a) More than 10 minutes

Sol:

Let X be the length of the call made by the person in the booth.

$$P(X > 10) = \int_{10}^{\infty} \lambda \cdot e^{-\lambda x}\, dx$$

$$= \lambda \cdot \int_{10}^{\infty} e^{-\lambda x}\, dx$$

$$= \lambda \cdot \left(\frac{-1}{\lambda} \cdot e^{-\lambda x}\right]_{10}^{\infty}$$

$$= e^{-\lambda x}\Big]_{\infty}^{10}$$

$$= e^{-10\lambda}$$

$$\boxed{P(X > 10) = e^{-5}} \approx 0.365$$

(b) between 10 and 20 minutes

$$P(10 < X < 20) = \int_{10}^{20} \lambda \cdot e^{-\lambda x}\, dx$$

$$= \lambda \cdot \left(\frac{-1}{\lambda} e^{-\lambda x}\right]_{10}^{20}$$

$$= \left(e^{-\lambda x}\right]_{20}^{10}$$

$$= e^{-10\lambda} - e^{-20\lambda}$$

$$= e^{-5} - e^{-10}$$

$$= e^{-10}(e^{5} - 1)$$

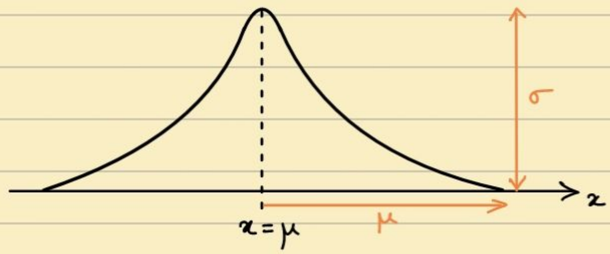
$$\boxed{P(10 < X < 20) \approx 0.233}$$

→ Normal Random Variables :

A random variable X is said to be normal distribution, if its PDF is

$$f(x) = \frac{1}{\sqrt{2\pi}\cdot\sigma}\cdot e^{\frac{-(x-\mu)^2}{2\cdot\sigma^2}} \qquad [-\infty < x < \infty]$$

Here, μ, σ^2 are the parameters of X.



As f is an even function,

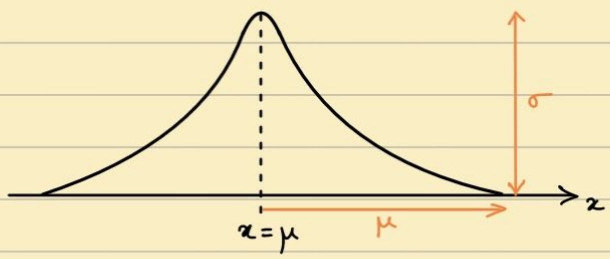
$f(-x) = f(x)$,

The Graph is symmetric about the line $x = \mu$

18/2

- X is normal if its PDF is the following

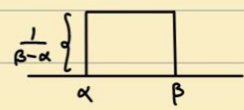
$$f(x) = \frac{1}{\sqrt{2\pi}\cdot\sigma}\, e^{-(x-\mu)^2/2\sigma^2} \qquad -\infty < x < \infty$$



For PDF :

$$\int_{-\infty}^{\infty} f(x)\,dx = 1$$

Uniform → $f(x) = \frac{1}{\beta-\alpha}$, $\alpha < x < \beta$



Exponential → $f(x) = \lambda . e^{-\lambda x}$

Check :

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\cdot\sigma}\cdot e^{-(x-\mu)^2/2\sigma^2}\,dx$$

$$= \frac{1}{\sqrt{2\pi}\cdot\sigma}\int_{-\infty}^{\infty} e^{-(x-\mu)^2/2\sigma^2}\,dx$$

Substitute $y = (x-\mu)/\sigma$

$$= \frac{1}{\sqrt{2\pi}\cdot\sigma} \cdot \int_{-\infty}^{\infty} e^{-y^2/2} \cdot \sigma \, dy$$

$$= \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{\infty} e^{-y^2/2} \, dy$$

$$\frac{y}{\sqrt{2}} = t$$

$$\Rightarrow dy = \sqrt{2}\, dt$$

$$\frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{\infty} e^{-t^2} \cdot dt \cdot \sqrt{2}$$

$$= \frac{\sqrt{2}}{\sqrt{2\pi}} \cdot \frac{\sqrt{\pi}}{2} \Big(\mathrm{erf}(t)\Big]_{-\infty}^{\infty}$$

$$= \frac{1}{2} \cdot \Big(\mathrm{erf}(t)\Big]_{-\infty}^{\infty}$$

$$= \frac{1}{2}(1-(-1)) = 1$$

Let $I = \int_{-\infty}^{\infty} e^{-y^2/2} \, dy$

$$\Rightarrow I^2 = \int_{-\infty}^{\infty} e^{-y^2/2} dy \cdot \int_{-\infty}^{\infty} e^{-x^2/2} dx$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{(-x^2+y^2)/2} \, dy \cdot dx$$

$$\left[\begin{array}{l} x = r\cos\theta \Rightarrow dx = -r\sin\theta \cdot d\theta \\ y = r\sin\theta \Rightarrow dy = r\cos\theta \cdot d\theta \\ \\ \quad dy \cdot dx = -r^2 \sin\theta\cos\theta \, d\theta \cdot d\theta \end{array}\right]$$

$$= \int_0^{\infty}\int_0^{2\pi} e^{-r^2/2} \, r\,dr \cdot d\theta$$

$dy \cdot dx = r \cdot dr\, d\theta \longrightarrow ?$

$$= \int_0^{\infty} r \cdot e^{-r^2/2} \cdot \int_0^{2\pi} d\theta \cdot dr$$

$$= 2\pi \cdot \int_0^{\infty} r \cdot e^{-r^2/2} \, dr$$

$$= 2\pi \cdot \int_0^{\infty} e^{-r^2/2} \, d(r^2/2)$$

$$= 2\pi \Big((-1) \cdot e^{-r^2/2}\Big]_0^{\infty}$$

$$\Rightarrow I^2 = 2\pi$$

$$\Rightarrow \boxed{I = \sqrt{2\pi}}$$

As $\text{PDF} = \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{\infty} e^{-y^2/2} \, dy$

$$\Rightarrow \text{PDF} = \frac{1}{\sqrt{2\pi}} \cdot \sqrt{2\pi}$$

$$= 1$$

$$\therefore \int_{-\infty}^{\infty} f(x)\, dx = 1$$

Next we find $E[X]$ and $Var[X]$

Mid-Sem : Upto Normal distribution

19/2 :

$$E[X] = \int_{-\infty}^{\infty} x \cdot f(x)\, dx$$

$$= \int_{-\infty}^{\infty} x \cdot \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot e^{-(x-\mu)^2/2\sigma^2} \cdot dx$$

$$= \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot \int_{-\infty}^{\infty} x \cdot e^{-(x-\mu)^2/2\sigma^2}\, dx$$

$$= \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot \int_{-\infty}^{\infty} [(x-\mu)+\mu] \cdot e^{-(x-\mu)^2/2\sigma^2}\, dx$$

$$= \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot \int_{-\infty}^{\infty} (x-\mu) \cdot e^{-(x-\mu)^2/2\sigma^2}\, dx + \underbrace{\frac{\mu}{\sqrt{2\pi} \cdot \sigma} \cdot \int_{-\infty}^{\infty} e^{-(x-\mu)^2/2\sigma^2}\, dx}_{\text{Density Function}}$$

$$y = x - \mu \Rightarrow dy = dx$$

$$= \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot \int_{-\infty}^{\infty} y \cdot e^{-y^2/2\sigma^2}\, dy + \mu$$

$$= \mu$$

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

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi} \cdot \sigma} (x-\mu)^2\, e^{-(x-\mu)^2/2\sigma^2}\, dx$$

Substitute $y = \frac{x-\mu}{\sigma}$

$$= \frac{1}{\sqrt{2\pi}\cdot\sigma} \cdot \int_{-\infty}^{\infty} y^2 \sigma^2 . e^{-y^2/2} \sigma . dy$$

$$= \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y^2 . e^{-y^2/2} dy$$

$$= \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} -y . d(e^{-y^2/2})$$

$$= \frac{\sigma^2}{\sqrt{2\pi}} \left\{ \left[\left(-y . e^{-y^2/2}\right]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} -e^{-y^2/2} dy \right\}$$

$$= \frac{\sigma^2}{\sqrt{2\pi}} \cdot \left\{ 0 + \int_{-\infty}^{\infty} e^{-y^2/2} dy \right\}$$

$$= \frac{\sigma^2}{\sqrt{2\pi}} \cdot \sqrt{2\pi} = \sigma^2$$

$\longrightarrow$ Let X be a normal distribution with parameters μ and σ^2, i.e. the PDF is

$$f(x) = \frac{1}{\sqrt{2\pi}\cdot\sigma} \cdot e^{-(x-\mu)^2/2\sigma^2} \quad , -\infty < x < \infty$$

Let's define

$$Y = \alpha \underline{x} + \beta \, , \quad \alpha, \beta \in \mathbb{R}$$

Why Y is normal ?

Sol: Assume $\alpha > 0$

$$F_Y(a) = P(Y \leq a) \quad , a \in \mathbb{R}$$

$$= P(\alpha x + \beta \leq a)$$

$$= P(\alpha x \leq a - \beta)$$

$$= P\left(X \leq \frac{a-\beta}{\alpha}\right)$$

$$= P_X\left(\frac{a-\beta}{\alpha}\right)$$

Differentiate the function w.r.t a,

$$f_Y(a) = f_X\left(\frac{a-\beta}{\alpha}\right) \cdot \frac{1}{\alpha}$$

$$= \frac{1}{\alpha} \cdot \frac{1}{\sqrt{2\pi}\cdot\sigma} \cdot e^{-\left[\left(\frac{a-\beta}{\alpha}\right) - \mu\right]^2 / 2\sigma^2}$$

$$= \frac{1}{\sqrt{2\pi}\cdot \alpha\sigma} \cdot e^{-(a-\beta-\alpha\mu)^2/2\alpha^2\sigma^2}$$

$$\left[\begin{array}{l} \text{Here, } \sigma' = \alpha\sigma \\ \qquad \mu' = \beta + \alpha\mu \end{array} \right]$$

$$= \frac{1}{\sqrt{2\pi}\cdot\sigma'} \cdot e^{-(a-\mu')/2\sigma'^2}$$

Y is normal with parameters $\beta + \alpha\mu$ and $\alpha\sigma$.

<u>20/2</u>

If X is normal, then $\alpha X + \beta$ is also normal.

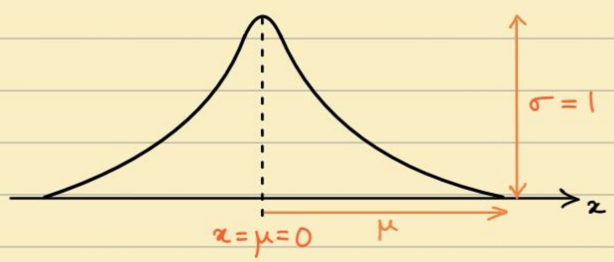
$$X \rightsquigarrow (\mu, \sigma)$$
$$\alpha X + \beta \rightsquigarrow (\alpha\mu + \beta, \alpha\sigma)$$

[H·W : $\alpha < 0$]

Let $Z = \frac{X-\mu}{\sigma} = \frac{X}{\sigma} + \left(\frac{-\mu}{\sigma}\right)$

Z is also normal with parameters $\left(\frac{1}{\sigma}(\mu) + \left(\frac{-\mu}{\sigma}\right), \frac{1}{\sigma}(\sigma)\right)$

$$\equiv (0, 1)$$



Hence, Z is also called Standard Normal Variable.

- <u>Properties</u>:

(1) The PDF of Z is

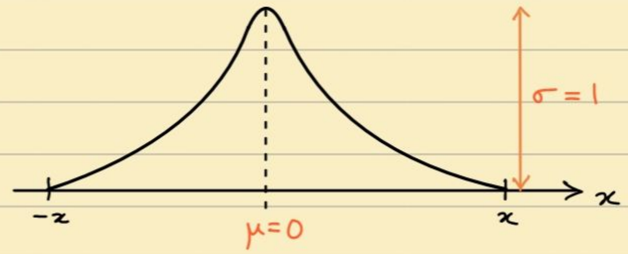
$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

(2) The CDF of Z is denoted by

$$\Phi(x) = F_Z(x)$$

(3) For any non-negative $x \in \mathbb{R}$,

$$\boxed{\Phi(-x) = 1 - \Phi(x)}$$



$$P(Z \le -x) = P(Z \ge x)$$
$$\Phi(-x) = 1 - P(Z \le x)$$
$$= 1 - \Phi(x)$$

Observation:

Let X be a normal random variable with (μ, σ)

$$F_X(a) = P(X \le a)$$

$$= P\left(\frac{X-\mu}{\sigma} \le \frac{a-\mu}{\sigma}\right)$$

$$= P\left(Z \le \frac{a-\mu}{\sigma}\right)$$

$$= \Phi\left(\frac{a-\mu}{\sigma}\right)$$

Ex. If X is normal with parameters $\mu = 3$ and $\sigma^2 = 9$. Find

(a) $P(2 < X < 5)$

$$= P\left(\frac{2-3}{3} < \frac{X-3}{3} < \frac{5-3}{3}\right)$$

$$= P\left(\frac{-1}{3} < \frac{X-3}{3} < \frac{2}{3}\right)$$

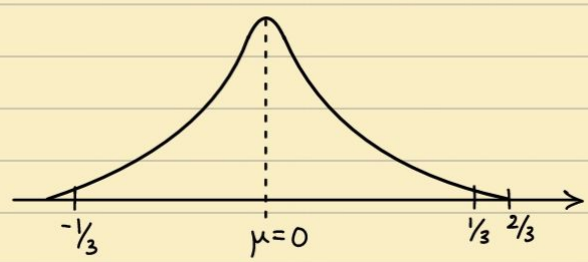
$$= P\left(\frac{-1}{3} < Z < \frac{2}{3}\right)$$

$$= P\left(Z < \frac{2}{3}\right) - P\left(Z \le \frac{-1}{3}\right) \longrightarrow \Phi(2/3) - \Phi(-1/3)$$

$$= P(Z < 0.66) - P(Z \le -0.33) \qquad = \Phi(2/3) - (1 - \Phi(1/3))$$

$$= 0.7454 - 0.3707$$

$$= 0.3747$$



- <u>Theorem</u>:

Let X be a binomial distribution with parameters (n, p), Then,

$$P\left(a < \frac{X - np}{\sqrt{np(1-p)}} \leq b\right) = \Phi(b) - \Phi(a) \qquad \left[\text{When } n \longrightarrow \text{large}\right]$$

Problem Set - 3

1. A poker hand means a set of five cards selected at random from usual deck of playing cards.

   (a) Find the probability that it is a Royal Flush - means that it consists of ten, jack, queen, king, ace of one suit.

   (b) Find the probability that it is four of a kind - means that there are four cards of equal face value.

   (c) Find the probability that it is a full house - means that it consists of one pair and one triple of cards with equal face values.

   (d) Find the probability that it is a straight - means that it consists of five cards in a sequence regardless of suit.

   (e) Find the probability that it consists of three cards of equal face value and two other cards but not a full house.

   (f) Find the probability that it consists of two distinct pairs and another card but does not fall into previous categories.

   (g) Find the probability that it consists a pair and three other cards but does not fall into previous categories.

Total hands = $\binom{52}{5}$

(a) Royal Flush $\longrightarrow$ A, Q, K, J, 10 (same suit)

Total suits : 4

Choosing any one suit : $\binom{4}{1}$

$$\boxed{P(\text{Royal flush}) = \frac{\binom{4}{1}}{\binom{52}{5}}}$$

(b) Four of a kind $\longrightarrow$ 4 cards of equal face value

Total cards in a suit : 13

Only $\binom{4}{4}$ ways to select 4 cards from 4 suits having a same face value.

Such combination can be done for 13 cards in total.

i.e. $13 \times \binom{4}{4}$

Now, 4 cards are taken away : $52 - 4 = 48$.

5^{th} card can be chosen from these 48 cards.

i.e. $\binom{48}{1}$

$$\boxed{\therefore \text{ Result : } \frac{13 \times \binom{4}{4} \times \binom{48}{1}}{\binom{52}{5}}}$$

(c) Full House

First 3 cards of equal face value :

Choose any value $\longrightarrow \binom{13}{1}$

Choose 3 suits out of 4 $\longrightarrow \binom{4}{3}$

(for those 3 cards)

$\therefore$ P (3 cards of equal face value) $= \binom{13}{1} \times \binom{4}{3}$

Next, 2 cards of equal face value :

(different from previous)

$\because$ 1 value is chosen from 13 $\Rightarrow$ 12 left

Choose any value $\longrightarrow \binom{12}{1}$

Choose 2 suit out of 4 $\longrightarrow \binom{4}{2}$

(for those 3 cards)

$\therefore$ P (2 cards of equal face value) $= \binom{12}{1} \times \binom{4}{2}$

$$\boxed{P(\text{Full House}) = \frac{\binom{13}{1} \times \binom{4}{3} \times \binom{12}{1} \times \binom{4}{2}}{\binom{52}{5}}}$$

(d) Straight

Sequence $\longrightarrow$ Total $-$ Same suit cards

Total possible sequences :

$$\left.\begin{array}{c} A-2-3-4-5 \\ 2-3-4-5-6 \\ 3-4-5-6-7 \\ \vdots \\ 9-10-J-Q-K \\ 10-J-Q-K-A \end{array}\right\} \underline{\underline{10}}$$

$\therefore$ Way to choose straight $= 10 \times \binom{4}{1} \times \binom{4}{1} \times \binom{4}{1} \times \binom{4}{1} \times \binom{4}{1} - 10 \times \binom{4}{1}$

$= (10 \times 4^5) - 40$

$$\boxed{\therefore P(\text{straight}) = \frac{(10 \times 4^5) - 40}{\binom{52}{5}}}$$

(e) Three cards having equal face value + 2 other (not full house)

Choose one value : $\binom{13}{1}$

Choose 3 suits out of 4 : $\binom{4}{3}$

Now, Choose 2 values from the remaining 12 values : $\binom{12}{2}$

Each card can be any of 4 suits : $\binom{4}{1} \times \binom{4}{1}$

$\therefore$ Way to choose : $\binom{13}{1} \times \binom{4}{3} \times \binom{12}{2} \times \binom{4}{1} \times \binom{4}{1}$

$$\boxed{\therefore P(\text{result}) = \frac{\binom{13}{1} \times \binom{4}{3} \times \binom{12}{2} \times \binom{4}{1} \times \binom{4}{1}}{\binom{52}{5}}}$$

(f) $\underbrace{\text{Two distinct pairs}}$ + 1 totally different card

Choose 2 values out of 13 : $\binom{13}{2}$

Choose 2 suits for each pair : $\binom{4}{2} \times \binom{4}{2}$

Now, choose the last card from rest 11 values : $\binom{11}{1}$

This last card can be of any 4 suits : $\binom{4}{1}$

Way to choose : $\binom{13}{2} \times \binom{4}{2} \times \binom{4}{2} \times \binom{11}{1} \times \binom{4}{1}$

$$\boxed{P(\text{result}) : \frac{\binom{13}{2} \times \binom{4}{2} \times \binom{4}{2} \times \binom{11}{1} \times \binom{4}{1}}{\binom{52}{5}}}$$

(g) $\underbrace{\text{1 pair}}$ + 3 other cards, different from the pair

Choose 1 value out of 13 : $\binom{13}{1}$

Choose 2 suits for the pair : $\binom{4}{2}$

Now, choose 3 other values from the remaining 12 values : $\binom{12}{3}$

Each card can be from any suit : $\binom{4}{1} \times \binom{4}{1} \times \binom{4}{1}$

Way to choose : $\binom{13}{1} \times \binom{4}{2} \times \binom{12}{3} \times \binom{4}{1} \times \binom{4}{1} \times \binom{4}{1}$

$$\boxed{\therefore P(\text{result}) = \frac{\binom{13}{1} \times \binom{4}{2} \times \binom{12}{3} \times \binom{4}{1} \times \binom{4}{1} \times \binom{4}{1}}{\binom{52}{5}}}$$

2. A bridge distribution means a distribution of the usual deck of playing cards among four persons to be called N, E, S, W, each getting 13 cards.

   (a) Show that the probability p of W receiving exactly k aces is same as the probability that an arbitrary hand of 13 cards contains excatly k aces.

   (b) What is the probability that N and S together get k aces? Here $k = 0, 1, 2, 3, 4$.

   (c) Find the probability that N, S, E, W get a, b, c, d spades respectively

N $\rightarrow$ 13 cards
E $\rightarrow$ 13 cards
S $\rightarrow$ 13 cards
W $\rightarrow$ 13 cards

(a) The probability of getting exactly k aces in a 13-card hand:

$$P(k) = \frac{\binom{4}{k}\overbrace{\binom{48}{13-k}}^{\text{Choosing remaining non-ace cards}}}{\underbrace{\binom{52}{13}}_{\rightarrow \text{ 13 cards out of 52}}}$$

Since, the desk is shuffled randomly, every player's hand follows the same distribution as any randomly chosen 13-card hand

Hence, the probability that W recieves exactly k aces is the same.

(b) N and S together recieve 26 cards from the 52-card deck. The number of aces in their combined hand:

$$\boxed{P(k) = \frac{\binom{4}{k}\binom{48}{26-k}}{\binom{52}{26}}}$$

$\leftarrow$ Probability that N and S together recieve exactly k aces.

(c) 13 spades in the deck.

Each player gets 13 cards

i.e. Partitioning 13 spades among 4 players.

$$P(a,b,c,d) = \frac{\binom{13}{a}\binom{13-a}{b}\binom{13-a-b}{c}\binom{13-a-b-c}{d}}{\binom{52}{13,13,13,13}}$$

$$[a+b+c+d = 13]$$

3. I have n sticks. Each is broken into two pieces - one long and one short piece. These $2n$ pieces are paired at random two form n sticks.

(a) What is the probability that they are joined to form original sticks?

(b) Find the probability that all long parts are paired with short parts.

(a) The total no. of ways to pair up 2n pieces into n pairs :

$$\frac{(2n)!}{n! \cdot 2^n} \longrightarrow n \text{ pairs from } 2n \text{ objects}$$

(Order within each pair does not matter)

To form original sticks, each long piece must be paired with its original short piece, ∴ Only 1 way

$$P = \frac{1}{\left(\frac{(2n)!}{2^n \cdot n!}\right)}$$

$$\Rightarrow \boxed{P = \frac{n! \cdot 2^n}{(2n)!}}$$

(b) n! ways of pairing long pieces with the short pieces.

Total no. of ways to randomly pair 2n pieces :

$$\frac{(2n)!}{n! \cdot 2^n}$$

$$\therefore \quad P = \frac{n!}{\left(\frac{(2n)!}{n! \cdot 2^n}\right)} = \frac{(n!)^2 \cdot 2^n}{(2n)!}$$

4. In how many ways can two rooks of different colours be put on a chess board so that they can take each other?

First rook $\longrightarrow$ 64 choices

Second rook $\rightarrow$ 7 (Horizontal) + 7 (Vertical)

= 14 choices

Total ways = 64×14

$\boxed{\therefore \text{Total ways} = 896}$

5. Show that it is more probable to get at least one ace with four dice than at least one double ace in 24 throws of two dice. This is apparently known as de Mere's paradox. There is, of course, no paradox – it just 1 so happens that Chevalier de Mere thought that both probabilities are equal.

Instance 1 : The probability of rolling atleast a 1 in 4 rolls of a fair die

Instance 2 : The probability of rolling atleast 2 1s in 24 rolls of 2 fair dice

As, P(atleast one success) = 1 − P(No success)

(i) P(No ace in one roll) $= \frac{5}{6}$

$\Rightarrow$ P(No ace in 4 rolls) $= \left(\frac{5}{6}\right)^4$

$$\boxed{\therefore P(\text{atleast one ace}) = 1 - \left(\frac{5}{6}\right)^4 \approx 0.518}$$

(ii) P(Double ace in one throw) $= 1/6 \times 1/6 = 1/36$

$\Rightarrow$ P(No double ace in one throw) $= 35/36$

$\Rightarrow$ P(No double ace in 24 throws) $= (35/36)^{24}$

$$\boxed{\therefore P(\text{atleast one double ace}) = 1 - \left(\frac{35}{36}\right)^{24} \approx 0.492}$$

As $0.518 > 0.492$,

$$\boxed{P(\text{Atleast one ace in 4 rolls}) > P(\text{Atleast one double ace in 24 throws})}$$

6. If n balls are placed at random into n cells, find the probability that exactly one cell remains empty.

Total no. of Possible Arrangements $= n^n$

No. of favourable Arrangements :

To have exactly one empty cell, we distribute n balls into $n-1$ cells, since one cell is empty.

$\therefore$ $n-1$ cells must contain atleast two balls.

- 'n' ways to choose which cell will be empty.
- Distributing n balls into $n-1$ cells : (Such that no cell is empty)

$$(n-1)^n - \binom{n-1}{1}(n-2)^n + \binom{n-2}{2}(n-3)^n - \ldots\ldots + (-1)^{n-2}\cdot\binom{n-1}{n-2}(1)^n$$

[Principle of Inclusion - Exclusion]

Now,

$$P = \frac{n \times \left[(n-1)^n - \binom{n-1}{1}(n-2)^n + \binom{n-2}{2}(n-3)^n - \ldots\ldots + (-1)^{n-2}\cdot\binom{n-1}{n-2}(1)^n\right]}{n^n}$$

$$\boxed{\therefore P = \frac{(n-1)^n - \binom{n-1}{1}(n-2)^n + \binom{n-2}{2}(n-3)^n - \ldots\ldots + (-1)^{n-2}\cdot\binom{n-1}{n-2}(1)^n}{n^{n-1}}}$$

For large n, i.e. $n \longrightarrow \Rightarrow P \rightarrow 1/e$

i.e. Poisson distribution with parameter $\lambda = 1$

7. A man is given n keys, in a random order, of which only one fits the door. He tries the keys, one after the other, to open the door. This procedure may require $1, 2, 3 \cdots, n$ trials. Show that each of these n has probability $\frac{1}{n}$.

Since the keys are arranged in random order, each of the n keys is equally likely to be in any of the n positions.

The probability that the correct key is in the first position : $1/n$
(only one trial is required)

The probability that the correct key is in the second position : $1/n$
(first key was incorrect, second key is correct)

& Generally,

The probability that the correct key is in the k^{th} position : $1/n$
(first '$k-1$' keys were incorrect and the k^{th} key is correct)

$\because$ Keys are randomly arranged, each key has an equal chance of appearing in any of the n positions.

$\boxed{\therefore \text{Probability of needing exactly } k \text{ trials} = 1/n \quad \forall k = 1, 2, 3, \ldots, n}$

8. A box contains 90 good and 10 defective items. If 10 items are selected then what is the probability that none of them is defective?

Given :

$P(\text{Good}) = 90/100 = 0.9$

$P(\text{Defective}) = 10/100 = 0.1$

10 items were selected. → (n)

Total Items (N) = 100

No. of defective items (m) = 10

Desired no. of defective items (i) = 0

This follows $\underbrace{\text{hypergeometric distribution}}$ → Probability of k successes in n draws from a finite population without replacement

$$P(X=i) = \frac{\binom{m}{i}\binom{n-m}{n-i}}{\binom{N}{n}}$$ ← The probability of drawing k defective items in n draws.

$$\Rightarrow P(X=0) = \frac{\binom{10}{0} \times \binom{90}{10}}{\underbrace{\binom{100}{10}}}$$

$\binom{10}{0}$: 1 way to choose 0 defective items

$\binom{90}{10}$: No. of ways to choose 10 good items.

→ Choosing 10 items from total 100 items.

$$\Rightarrow \boxed{P(X=0) = \frac{\binom{90}{10}}{\binom{100}{10}}}$$

9. If n men among whom are A and B, stand in a row what is the probability that there are exactly r men between A and B? What if they stand in a ring (what does this mean?) and the clock-wise direction is used for counting the number between A and B?

Total men : n

Total no. of ways to arrange n men : $n!$

To have exactly r men b/w A and B :

$A \longrightarrow$ position i

$B \longrightarrow$ position $i+r+1$

& $1 \le i \le n-r-1$ ($\because$ A & B should fit within n positions)

$\therefore$ A $\longrightarrow$ '$n-r$' possible starting positions

The remaining $n-2$ men can be arranged in $(n-2)!$ ways

A & B can exchange positions $\Rightarrow$ Multiply result by 2

$\therefore$ Favourable outcomes $= 2 \times (n-2)! \times (n-r)$

$$\therefore P = \frac{2(n-2)! \cdot (n-r)}{n!}$$

$$\Rightarrow \boxed{P = \frac{2(n-r)}{n(n-1)}}$$

When they stand in a ring, Circular arrangement fixes one person's position, i.e. $n! \longrightarrow (n-1)!$ total distinct arrangements

If A position is fixed as reference, B must be placed $r+1$ positions away from A in any direction.

Hence, giving only 2 possible solutions for B.

The remaining $n-2$ men can be arranged in $(n-2)!$ ways

$\therefore$ No. of favourable outcomes $= 2 \times (n-2)!$

$$\therefore P = \frac{2 \times (n-2)!}{(n-1)!} = \frac{2}{n-1}$$

$$\boxed{\therefore P = \frac{2}{n-1}}$$

If only clockwise direction is used for counting,
B gets to have only 1 position.

$\therefore$ No. of favourable outcomes $= (n-2)!$

$$P = \frac{(n-2)!}{(n-1)!} \Rightarrow \boxed{P = \frac{1}{n-1}}$$

10. What is the probability that two throws with three dice each will show the same configuration, if the dice are distinguishable? What if the dice are not distinguishable?

(i) Total unique configurations $= 6 \times 6 \times 6 = 216$ per throw

For two throws,

the probability that the second throw matches the first in exact configuration $= 1/216$

$\because$ Regardless of the outcome of the first throw, there is only one specific configuration out of 216 possible that will match it.

$$\boxed{\therefore\ P = 1/216}$$

(ii) For indistinguishable dice,

$(1,2,3) \equiv (3,2,1)$

No. of unique outcomes $= \binom{6+3-1}{3} = \binom{8}{3} = 56$

$\left[\binom{n+k-1}{k-1} \text{ for } x_1 + x_2 + x_3 + \dots x_k = n \right]$

$\because$ 2nd throw must match the first,

There are 56 favourable outcomes out of $56 \times 56 = 3136$ pairs

$$\therefore P = \frac{56}{3136} = \frac{1}{56}$$

$$\boxed{\therefore P = 1/56}$$

Problem Set - 4

1. Consider rolling a fair six-sided die, so that $S = \{1, 2, 3, 4, 5, 6\}$. Let $X(s) = s$, and $Y(s) = s^3 + 2$. Let $Z = XY$. Compute $Z(s)$ for all $s \in S$.

Given:

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

$X(s) = s$

$Y(s) = s^3 + 2$

$Z = XY$

$Z = (s)(s^3 + 2)$

$\Rightarrow \boxed{Z(s) = s^4 + 2s}$

s	Z(s)
1	3
2	20
3	87
4	264
5	635
6	1308

2. Suppose a university is composed of 55% female students and 45% male students. A student is selected to complete a questionnaire. There are 25 questions on the questionnaire administered to a male student and 30 questions on the questionnaire administered to a female student. If X denotes the number of questions answered by a randomly selected student, then compute $P(X = x)$ for every real number x.

Given:

55% female $\longrightarrow$ 30 questions

45% male $\longrightarrow$ 25 questions

$X \longrightarrow$ no. of questions answered a randomly selected student

$P(X = 30) = 0.55$

$P(X = 25) = 0.45$

Otherwise, $P(X = x) = 0$

$$\boxed{P(X = x) = \begin{cases} 0.45, & x = 25 \\ 0.55, & x = 30 \\ 0, & \text{otherwise} \end{cases}}$$

3. Consider flipping a fair coin. Let $Z = 1$ if the coin is heads, and $Z = 3$ if the coin is tails. Let $W = Z^2 + Z$.

(a) What is the probability function of Z?

(b) What is the probability function of W?

Given:

Heads $\Longrightarrow Z = 1$

Tails $\Longrightarrow Z = 3$

$W = Z^2 + Z$

(a) $P(Z=1) = 1/2$

$P(Z=3) = 1/2$

otherwise, $P(Z=x) = 0$

$$\boxed{P(Z=x) = \begin{cases} 1/2, & x = 1 \\ 1/2, & x = 3 \\ 0, & \text{otherwise} \end{cases}}$$

(b) $Z = 1 \Rightarrow W = 1^2 + 1 = 2$

$Z = 3 \Rightarrow W = 3^2 + 3 = 12$

$P(W=2) = P(Z=1) = 1/2$

$P(W=12) = P(Z=3) = 1/2$

$$\boxed{P(W=\omega) = \begin{cases} 0.5, & \omega = 2 \\ 0.5, & \omega = 12 \\ 0, & \text{otherwise} \end{cases}}$$

4. Let X be a Binomial random variable with parameter $(12, p)$. For what value of p is $P(X = 11)$ maximized?

Given:

$n = 12$

$p = p \in [0, 1]$

$$P(X=i) = \binom{n}{i} p^i (1-p)^{n-i}$$

$$\Rightarrow P(X=11) = \binom{12}{11} p^{11} (1-p)$$

$$= 12 \cdot p^{11} \cdot (1-p)$$

$$\frac{d}{dp}\left[P(X=11)\right] = 12 \times \left[11 \times p^{10} - 12 \cdot p^{11}\right] \overset{\text{set}}{=\!=} 0$$

$$\Rightarrow 11 \cdot p^{10} = 12 \cdot p^{11}$$

$$\Rightarrow \boxed{p = \frac{11}{12}} \quad \text{or} \quad p = 0 \ \times$$

5. An urn contains 4 black balls and 5 white balls. After a thorough mixing, a ball is drawn from the urn, its color is noted, and the ball is set aside. The remaining balls are then mixed and a second ball is drawn.

(a) What is the probability distribution of the number of black balls observed?

(b) What is the probability distribution of the number of white balls observed?

Given:

Total no. of balls = 9

Black balls = 4

White balls = 5

$P(W_1) = 5/9$

$P(B_1) = 4/9 \implies$ Black ball was drawn initially.

$P(B_2|B_1) = 3/8$ | $P(W_2|B_1) = 5/8$

$P(B_2|W_1) = 4/8$ | $P(W_2|W_1) = 4/8$

$$\implies P(B_1 \cap B_2) = P(B_1) \cdot P(B_2|B_1) = 4/9 \times 3/8 = 1/6$$

Similarly,

$$P(B_1 \cap W_2) = P(B_1) \cdot P(W_2|B_1) = 4/9 \times 5/8 = 5/18$$

$$P(W_1 \cap B_2) = P(W_1) \cdot P(B_2|W_1) = 5/9 \times 4/8 = 5/18$$

$$P(W_1 \cap W_2) = P(W_1) \cdot P(W_2|W_1) = 5/9 \times 4/8 = 5/18$$

(a) $P(X=0) = P(W_1 \cap W_2) = 5/18$

$P(X=1) = P(B_1 \cap W_2) + P(W_1 \cap B_2) = 5/18 + 5/18 = 5/9$

$P(X=2) = P(B_1 \cap B_2) = 1/6$

$$\boxed{P(X=x) = \begin{cases} 5/18, & x=0 \\ 5/9, & x=1 \\ 1/6, & x=2 \end{cases}}$$

(b) Let Y be the no. of white balls observed. The possible values are 0, 1, 2.

$P(Y=0) = P(B_1 \cap B_2) = 1/6$

$P(Y=1) = P(B_1 \cap W_2) + P(B_2 \cap W_1) = 5/18 + 5/18 = 5/9$

$P(Y=2) = P(W_1 \cap W_2) = 5/18$

$$\boxed{P(Y=y) = \begin{cases} 1/6, & y=0 \\ 5/9, & y=1 \\ 5/18, & y=2 \end{cases}}$$

6. Suppose an urn contains 1000 balls — one of these is black, and the other 999 are white. Suppose that 100 balls are randomly drawn from the urn with replacement. Use the appropriate Poisson distribution to approximate the probability that five black balls are observed.

Given:

$N = 1000$

No. of black balls $= 1$

No. of white balls $= 999$

Probability of drawing a black ball in one draw:

$$p = 1/1000 = 0.001$$

Let X be the no. of black balls observed in 100 draws.

Each draw $\longrightarrow$ Independent & follows Bernoulli style,

$$X \sim \text{Binomial}(\underbrace{n = 100}_{\text{large}}, \underbrace{p = 0.001}_{\text{very small}})$$

$$\left.\begin{aligned} \lambda &= np \\ &= 100 \times 0.001 \\ &= 0.1 \end{aligned}\right\} \text{Following Poisson distribution}$$

For $i = 5$:

$$P(X = i) = \frac{e^{-\lambda} \cdot \lambda^i}{i!}$$

$$\Rightarrow P(X = 5) = \frac{e^{-0.1} \cdot (0.1)^5}{5!}$$

$$\boxed{\therefore P(X = 5) \approx 7.54 \times 10^{-8}}$$

7. Suppose that there is a loop in a computer program and that the test to exit the loop depends on the value of a random variable X. The program exits the loop whenever $X \in A$, and this occurs with probability 1/3. If the loop is executed at least once, what is the probability that the loop is executed five times before exiting?

$X \longrightarrow$ Geometric Random variable

$P(\text{exit}) = 1/3 \longrightarrow$ success

$P(\text{continue}) = 1 - 1/3 = 2/3 \longrightarrow$ Failure

$$P(X = n) = (1 - p)^{n-1} \cdot p$$

$$\Rightarrow P(X = 5) = (1 - 1/3)^{5-1} \cdot (1/3)$$

$$= \left(\frac{2}{3}\right)^4 \left(\frac{1}{3}\right) = \frac{16}{243} \approx \boxed{0.0658}$$

8. Let W be the uniform random variable in the interval $[1, 4]$. Compute each of the following.

(a) $P(W \geq 5)$
(b) $P(W \geq 2)$
(c) $P(W^2 \leq 9)$
(d) $P(W^2 \leq 2)$

Given :

$W \rightarrow$ Uniform Random variable in $[1,4]$

$$f(x) = \begin{cases} \frac{1}{3}, & 1 \leq x \leq 4 \\ 0, & \text{otherwise} \end{cases}$$

(a) $P(W \geq 5)$

$$= \int_5^\infty f(x)\,dx = 0$$

$$\boxed{\therefore P(W \geq 5) = 0}$$

(b) $P(W \geq 2)$

$$= \int_2^\infty f(x)\,dx$$

$$= \int_2^4 \frac{1}{3}\,dx + \cancelto{0}{\int_4^\infty 0 \cdot dx}$$

$$= \frac{1}{3}\left(x\right]_2^4$$

$$\Rightarrow \boxed{P(W \geq 2) = 2/3}$$

$\frac{1}{n} x^{1/n - 1} \cdot F_X(\sqrt[n]{x})$

$$f(y) = \begin{cases} \frac{1}{3} \cdot \frac{1}{2\sqrt{x}}, & 1 \leq x \leq 16 \\ 0, & \text{otherwise} \end{cases}$$

(c) $P(W^2 \leq 9)$

$$= \int_{-\infty}^9 f(x)\,dx$$

$$= \cancelto{0}{\int_{-\infty}^1 0 \cdot dx} + \int_1^9 \frac{1}{3} \cdot \frac{1}{2\sqrt{x}}\,dx$$

$$= \frac{1}{3} \cdot \left(\sqrt{x}\right]_1^9$$

$$= 1/3\,(3-1) = 2/3$$

$$\boxed{\therefore P(W^2 \leq 9) = 2/3}$$

(d) $P(W^2 \leq 2)$

$$= \int_1^2 \frac{1}{3} \cdot \frac{1}{2\sqrt{x}}\,dx$$

$$= \frac{1}{3}\left(\sqrt{x}\right]_1^2$$

$$= \frac{\sqrt{2}-1}{3} \approx 0.138$$

$$\boxed{\therefore P(W^2 \leq 2) = \frac{\sqrt{2}-1}{3}}$$

(OR)

(c) $P(\underbrace{W^2 \leq 9}) = P(1 \leq W \leq 3)$

$\hookrightarrow W \in [-3, 3]$ & on $[-3, 1] \rightarrow f(x) = 0$

Similarly :

(d) $P(\underbrace{W^2 \leq 2}) = P(1 \leq W \leq \sqrt{2})$

$\hookrightarrow W \in [-\sqrt{2}, \sqrt{2}]$ & on $[-\sqrt{2}, 1] \rightarrow f(x) = 0$

9. Let X be the Exponential with parameter 3. Compute each of the following.

(a) $P(0 < X < 1)$
(b) $P(0 < X < 3)$
(c) $P(0 < X < 5)$
(d) $P(2 < X < 5)$
(e) $P(2 < X < 10)$
(f) $P(X > 2)$

$X \rightarrow$ Exponential , $\lambda = 3$

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

$$\Rightarrow f(x) = \begin{cases} 3e^{-3x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

(a) $P(0 < X < 1)$

$$= \int_0^1 f(x) \cdot dx$$
$$= \int_0^1 3e^{-3x}\, dx$$
$$= \left(e^{-3x}\right]_1^0$$
$$= e^0 - e^{-3}$$
$$= 1 - e^{-3} \approx 0.9502$$
$$\Rightarrow \boxed{P(0 < X < 1) = 1 - e^{-3}}$$

(b) $P(0 < X < 3)$

$$= \int_0^3 f(x) \cdot dx$$
$$= \int_0^3 3e^{-3x}\, dx$$
$$= \left(e^{-3x}\right]_3^0$$
$$= e^0 - e^{-9}$$
$$= 1 - e^{-9} \approx 0.999877$$
$$\Rightarrow \boxed{P(0 < X < 3) = 1 - e^{-9}}$$

(c) $P(0 < X < 5)$

$$= \int_0^5 f(x) \cdot dx$$
$$= \int_0^5 3e^{-3x}\, dx$$
$$= \left(e^{-3x}\right]_5^0$$
$$= e^0 - e^{-15}$$
$$= 1 - e^{-15} \approx 0.9999997$$
$$\Rightarrow \boxed{P(0 < X < 5) = 1 - e^{-15}}$$

(d) $P(2 < X < 5)$

$$= \int_2^5 f(x) \cdot dx$$
$$= \int_2^5 3e^{-3x}\, dx$$
$$= \left(e^{-3x}\right]_5^2$$
$$= e^{-6} - e^{-10}$$
$$= e^{-10}(e^4 - 1) \approx 0.00248$$
$$\Rightarrow \boxed{P(2 < X < 5) = e^{-6} - e^{-10}}$$

(e) $P(2 < X < 10)$

$$= \int_2^{10} f(x) \cdot dx$$

$$= \int_2^{10} 3e^{-3x}\, dx$$

$$= \left(e^{-3x}\right]_{10}^{2}$$

$$= e^{-6} - e^{-30}$$

$$= e^{-30}\left(e^{24} - 1\right) \approx 0.00248$$

$$\Rightarrow \boxed{P(2 < X < 10) = e^{-6} - e^{-30}}$$

(f) $P(X > 2)$

$$= \int_2^{\infty} f(x) \cdot dx$$

$$= \int_2^{\infty} 3e^{-3x}\, dx$$

$$= \left(e^{-3x}\right]_{\infty}^{2}$$

$$= e^{-6} - e^{-\infty}$$

$$= e^{-6} \approx 0.00248$$

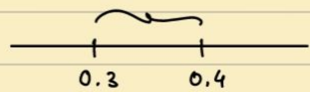
$$\Rightarrow \boxed{P(X > 2) = e^{-6}}$$

10. Suppose X has density f and that $f(x) \geq 2$ for $0.3 < x < 0.4$. Prove that $P(0.3 < X < 0.4) \geq 0.2$.

Given :

$$\underbrace{f(x)}_{} \geq 2 \quad \forall x \in (0.3, 0.4)$$

PDF of X

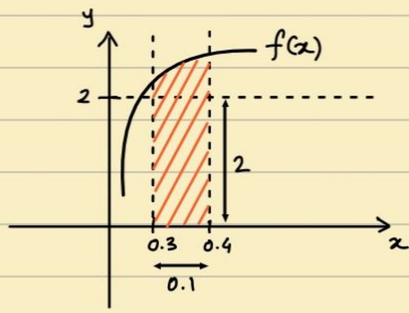


$P(0.3 < X < 0.4)$

$$= P(X \leq 0.4) - P(X \leq 0.3)$$

$$= \int_{-\infty}^{0.4} f(x) \cdot dx - \int_{-\infty}^{0.3} f(x) \cdot dx$$

$$= \int_{0.3}^{0.4} f(x) \cdot dx$$



$$\int_{0.3}^{0.4} f(x)\, dx \geq 2 \times (0.4 - 0.3)$$

$$\geq 2 \times 0.1$$

$$\geq 0.2 \qquad \left[\because f(x) \geq 2,\ x \in (0.3, 0.4)\right]$$

$$\boxed{\therefore\ P(0.3 < X < 0.4) \geq 0.2}$$

Problem Set - 5

1. Establish for which constants c the following functions are densities.

   (a) $f(x) = cx$ on $(0, 1)$ and 0 otherwise.

   (b) $f(x) = cx^n$ on $(0, 1)$ and 0 otherwise, for n a non-negative integer.

   (c) $f(x) = cx^{1/2}$ on $(0, 2)$ and 0 otherwise.

   (d) $f(x) = c \sin x$ on $(0, \pi/2)$ and 0 otherwise.

(a) $$f(x) = \begin{cases} cx, & 0 < x < 1 \\ 0, & \text{otherwise} \end{cases}$$

For $f(x)$ to be PDF,

$$\int_{-\infty}^{\infty} f(x) \cdot dx = 1$$

$$\Rightarrow \cancelto{0}{\int_{-\infty}^{0} f(x) \cdot dx} + \int_{0}^{1} f(x) \cdot dx + \cancelto{0}{\int_{1}^{\infty} f(x) \cdot dx} = 1$$

$$\Rightarrow \int_{0}^{1} cx \cdot dx = 1$$

$$\Rightarrow c\left(\frac{x^2}{2}\right]_0^1 = 1 \Rightarrow c/2 = 1 \Rightarrow \boxed{c = 2}$$

(b) $$f(x) = \begin{cases} cx^n, & 0 < x < 1 \\ 0, & \text{otherwise} \end{cases} \qquad [n \geq 0]$$

For $f(x)$ to be PDF,

$$\int_{-\infty}^{\infty} f(x) \cdot dx = 1$$

$$\Rightarrow \cancelto{0}{\int_{-\infty}^{0} f(x) \cdot dx} + \int_{0}^{1} f(x) \cdot dx + \cancelto{0}{\int_{1}^{\infty} f(x) \cdot dx} = 1$$

$$\Rightarrow \int_{0}^{1} cx^n \cdot dx = 1$$

$$\Rightarrow c\left(\frac{x^{n+1}}{n+1}\right]_0^1 = 1 \qquad [n \geq 0]$$

$$\Rightarrow c\left(\frac{1^{n+1}}{n+1}\right) = 1 \Rightarrow \boxed{c = n+1}$$

(c) $$f(x) = \begin{cases} cx^{1/2}, & 0 < x < 2 \\ 0, & \text{otherwise} \end{cases}$$

For $f(x)$ to be PDF,

$$\int_{-\infty}^{\infty} f(x) \cdot dx = 1$$

$$\Rightarrow \cancelto{0}{\int_{-\infty}^{0} f(x)\cdot dx} + \int_{0}^{2} f(x)\cdot dx + \cancelto{0}{\int_{2}^{\infty} f(x)\cdot dx} = 1$$

$$\Rightarrow \int_{0}^{2} c\sqrt{x}\cdot dx = 1$$

$$\Rightarrow c\left(\frac{x^{3/2}}{3/2}\right]_0^2 = 1$$

$$\Rightarrow c\left(\frac{1^{n+1}}{n+1}\right) = 1 \Rightarrow \boxed{c = n+1}$$

(d) $f(x) = \begin{cases} c\cdot\sin(x), & 0 < x < \pi/2 \\ 0, & \text{otherwise} \end{cases}$

For $f(x)$ to be PDF,

$$\int_{-\infty}^{\infty} f(x)\cdot dx = 1$$

$$\Rightarrow \cancelto{0}{\int_{-\infty}^{0} f(x)\cdot dx} + \int_{0}^{\pi/2} f(x)\cdot dx + \cancelto{0}{\int_{\pi/2}^{\infty} f(x)\cdot dx} = 1$$

$$\Rightarrow \int_{0}^{\pi/2} c\cdot\sin(x)\cdot dx = 1$$

$$\Rightarrow c\left(-\cos(x)\right]_0^{\pi/2} = 1$$

$$\Rightarrow c\left(\cos(0) - \cos(\pi/2)\right) = 1 \Rightarrow \boxed{c = 1}$$

2. Suppose X has density f and Y has density g. Suppose $f(x) > g(x)$ for $1 < x < 2$. Prove that $P(1 < X < 2) > P(1 < Y < 2)$.

PDF of $X \rightarrow f$

$\hookrightarrow \int_{-\infty}^{\infty} f(x) = 1$

PDF of $Y \rightarrow g$

$\hookrightarrow \int_{-\infty}^{\infty} g(y) = 1$

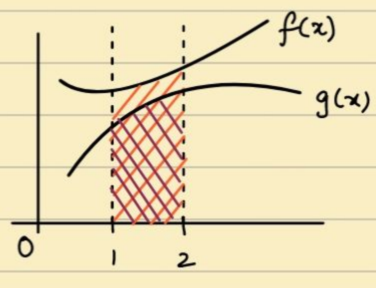
$f(x) > g(x)$ for $1 < x < 2$

$$P(1 < X < 2) = P(X \in (1,2)) = \int_1^2 f(x)\cdot dx$$

$$P(1 < X < 2) = \int_1^2 f(x)\cdot dx$$

$$P(1 < Y < 2) = \int_1^2 g(y)\cdot dy = \int_1^2 g(x)\cdot dx$$

As $f(x) > g(x)$ for $1 < x < 2$



From diagram:

Area under $f(x)$ > Area under $g(x)$

$$\therefore \int_1^2 f(x)\,dx > \int_1^2 g(y)\,dy$$

$$\boxed{\therefore\ P(1<X<2) > P(1<Y<2)}$$

3. Suppose X is a standard normal variable and Y is a normal variable with mean 1 and variance 1. Prove that $P(X < 3) > P(Y < 3)$.

For Y:

$$\left.\begin{array}{l}\mu = 1\\ \sigma^2 = 1\end{array}\right\} P(Y<3) = P\left(\frac{Y-1}{1} < \frac{3-1}{1}\right)$$

$$= P(Y-1 < 2)$$

$$= P(Z < 2)$$

$$= 0.97725$$

From Standard Normal Table

$$\left.\begin{array}{l}\mu = 0\\ \sigma^2 = 1\end{array}\right\} P(X<3) = 0.99865$$

$$P(Y<3) = \int_{-\infty}^{3} f(y)\cdot dy$$

$$= \int_{-\infty}^{3} \frac{1}{\sqrt{2\pi}} \cdot e^{-(x-1)^2/2}\,dx$$

$$= \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{3} e^{-(x-1)^2/2}\cdot dx$$

$$= 0.977249$$

$$f(y) = \frac{1}{\sqrt{2\pi}\cdot\sigma} \cdot e^{-(x-\mu)^2/2\sigma^2}$$

$$= \frac{1}{\sqrt{2\pi}} \cdot e^{-(x-1)^2/2}$$

$$\boxed{\therefore\ P(X>3) > P(Y>3)}$$

4. Consider rolling one fair six-sided die, so that $S = \{1, 2, 3, 4, 5, 6\}$, and $P(s) = 1/6$ for all $s \in S$. Let X be the number showing on the die, so that $X(s) = s$ for $s \in S$. Let $Y = X^2$. Compute the cumulative distribution function $F_Y(y) = P(Y \le y)$, for all $y \in R$.

Given:

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

$$P(s) = 1/6 \quad \forall s \in S$$

$$X(s) = s \quad \forall s \in S$$

$$Y = X^2$$

$$F_Y(y) = P(Y \le y) \quad \forall y \in \mathbb{R}$$

$$\left[\begin{array}{l}\text{As PDF of } Y = X^n \text{ is } F_X(x^{1/n}) \cdot \frac{1}{n} \cdot x^{1/n - 1}, \\ \quad \text{PDF of } Y = X^2 \text{ is } F_X(x^{1/2}) \cdot \frac{1}{2} \cdot x^{1/2 - 1} \\ \qquad \Rightarrow F_Y(y) = \frac{1}{2\sqrt{x}} \cdot F_X(\sqrt{x})\end{array}\right]$$

$$Y = X^2 = \{1, 4, 9, 16, 25, 36\}$$

$$F_Y(y) = P(Y \le y)$$

$\forall y < 1 \Rightarrow F_Y(y) = 0$

$\forall y \in [1, 4) \Rightarrow F_Y(y) = P(Y = 1)$

$= P(X = 1) = 1/6$

$\forall y \in [4, 9) \Rightarrow F_Y(y) = P(Y = 1) + P(Y = 4)$

$= P(X = 1) + P(X = 2)$

$= 1/6 + 1/6 = 2/6 = 1/3$

Similarly,

$\forall y \in [9, 16) \Rightarrow F_Y(y) = P(Y \le 9) = 3/6 = 1/2$

$\forall y \in [16, 25) \Rightarrow F_Y(y) = P(Y \le 16) = 4/6 = 2/3$

$\forall y \in [25, 36) \Rightarrow F_Y(y) = P(Y \le 25) = 5/6$

$\forall y \ge 36 \Rightarrow F_Y(y) = P(Y \le 36) = 6/6 = 1$

$$\boxed{F_Y(y) = \begin{cases} 0, & y < 1 \\ 1/6, & 1 \le y < 4 \\ 1/3, & 4 \le y < 9 \\ 1/2, & 9 \le y < 16 \\ 2/3, & 16 \le y < 25 \\ 5/6, & 25 \le y < 36 \\ 1, & y \ge 36 \end{cases}}$$

5. Let X is a normal random variable with mean +8 and ~~mean~~ Var. 4. Compute each of the following in terms of the cumulative distributive function,

(a) $P(X \ge 5)$
(b) $P(2 \ge X \ge 7)$
(c) $P(X \ge 3)$

Given: $\mu = 8$

$\sigma^2 = 4 \Rightarrow \sigma = 2$

(a) $P(X \geq 5)$

$$= P\left(\frac{X-8}{2} \geq \frac{5-8}{2}\right)$$

$$= P\left(Z \geq \frac{-3}{2}\right)$$

$$= 1 - P(Z < -1.5)$$

$$= 1 - 0.06681$$

$$\boxed{\therefore P(X \geq 5) = 0.93319}$$

(c) $P(X \geq 3)$

$$= P\left(\frac{X-8}{2} \geq \frac{3-8}{2}\right)$$

$$= P\left(Z \geq \frac{-5}{2}\right)$$

$$= 1 - P(Z < -2.5)$$

$$= 1 - 0.00621$$

$$\boxed{\therefore P(X \geq 3) = 0.99379}$$

(b) $P(2 \geq X \geq 7)$

$$= P\left(\frac{2-8}{2} \geq Z \geq \frac{7-8}{2}\right)$$

$$= P(-3 \geq Z \geq -0.5)$$

$$= P(Z \leq -3) + 1 - P(Z < -0.5)$$

$$= 0.00135 + 1 - 0.30854$$

$$= 0.69281$$

$$\boxed{\therefore P(2 \geq X \geq 7) = 0.69281}$$

6. Suppose $F_Y(y) = y^3$ for $0 \leq y < 1/2$, and $F_Y(y) = 1 - y^3$ for $1/2 \leq y \leq 1$. Compute each of the following.

(a) $P(1/3 < Y < 3/4)$

(b) $P(Y = 1/3)$

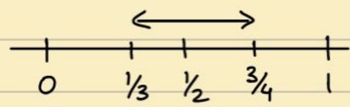
(c) $P(Y = 1/2)$

Given:

$$F_Y(y) = \begin{cases} y^3, & 0 \leq y < 1/2 \\ 1 - y^3, & 1/2 \leq y \leq 1 \end{cases}$$

(a) $P(1/3 < Y < 3/4)$

$$= P(1/3 < Y < 1/2) + P(1/2 \leq Y < 3/4)$$



As, $P(a < Y < b) = F_Y(b) - F_Y(a)$

$$= P(1/3 < Y < 1/2) + P(1/2 \leq Y < 3/4)$$

$$= \cancel{F_Y(1/2)} - F_Y(1/3) + F_Y(3/4) - \cancel{F_Y(1/2)}$$

$$= -\left(\frac{1}{3}\right)^3 + \left[1-\left(\frac{3}{4}\right)^3\right]$$

$$= \frac{935}{1728} = \boxed{0.54109}$$

(b) $\boxed{P(Y = 1/3) = 0}$

$$\left[\because \int_{1/3}^{1/3} f(y)\,dy = 0\right]$$

For continuous distributions,

$P(Y = y) = 0$ for any individual point.

(c) $\boxed{P(Y = 1/2) = 0}$

Similar to (b)

7. Let X is a Exponential random variable with parameter 3. Compute the function F_X.

Given : $\lambda = 3$

PDF: $f(x) = \begin{cases} 3e^{-3x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$

$$F_X(x) = P(X \leq x) = \int_{-\infty}^{x} f(x) \cdot dx$$

If $x < 0 \Rightarrow F_X(x) = 0$

If $x \geq 0$:

$$F_X(x) = P(X \leq x) = \cancel{\int_{-\infty}^{0} f(x) \cdot dx} + \int_{0}^{x} f(x) \cdot dx$$

0

$$= \int_0^x 3e^{-3x}\,dx$$

$$= 3 \cdot \left(\frac{-1}{3} \cdot e^{-3x}\right]_0^x$$

$$= -\left(e^{-3x} - e^0\right)$$

$$= 1 - e^{-3x}$$

$$\boxed{\therefore\ F_X(x) = \begin{cases} 1 - e^{-3x}, & x \geq 0 \\ 0, & x < 0 \end{cases}}$$

8. Let X be a Bernoulli random variable with parameter $1/3$, and let $Y = 4X - 2$. Compute the joint cumulative density function $F_{X,Y}$.

Given:

$p = 1/3$ for X - Bernoulli Random variable

$Y = 4X - 2$

$P(X = 1) = 1/3$

$P(X = 0) = 1 - 1/3 = 2/3$

If $X = 0 \Rightarrow Y = 4(0) - 2$

$\Rightarrow Y = -2$

If $X = 1 \Rightarrow Y = 4(1) - 2$

$\Rightarrow Y = 2$

$\therefore$ Y takes value in $-2, 2$.

Joint CDF : $F_{X,Y}(x,y) = P(X \le x, Y \le y)$

Case (i) $x < 0$, $y < -2$

$F_{X,Y}(x,y) = 0$

$\because$ X is always 0 or 1 & Y is always -2 or 2.

(ii) $0 \le x < 1$, $y < -2$

$F_{X,Y}(x,y) = 0$

$\because$ Y is always -2 or 2

(iii) $x < 0$, $y \ge -2$

$F_{X,Y}(x,y) = 0$

$\because$ X is always 0 or 1.

(iv) $0 \le x < 1$, $-2 \le y < 2$

$P(X = 0, Y = -2) = P(X = 0)$

$= 2/3$

$\Rightarrow F_{X,Y}(x,y) = 2/3$

(v) $x \ge 1$, $y < -2$

$F_{X,Y}(x,y) = 0$

$\because$ Y is always -2 or 2

(vi) $x \ge 1$, $-2 \le y < 2$

$P(X = 0, Y = -2) = P(X = 0)$

$= 2/3$

$\Rightarrow F_{X,Y}(x,y) = 2/3$

(vii) $0 \le x < 1$, $y \ge 2$

$P(X = 0, Y = -2) = P(X = 0)$

$= 2/3$

$\Rightarrow F_{X,Y}(x,y) = 2/3$

(viii) $x \ge 1$, $y \ge 2$

$F_{X,Y}(x,y) = P(X = 0, Y = -2) + P(X = 1, Y = 2)$

$= 2/3 + 1/3$

$= 1$

$$F_{X,Y}(x,y) = \begin{cases} 0, & x<0 \ \& \ y<-2 \\ 2/3, & 0\le x<1,\ -2\le y<2 \\ 2/3, & x\ge 1,\ y<2 \\ 2/3, & 0\le x<1,\ y\ge 2 \\ 1, & x\ge 1,\ y\ge 2 \end{cases}$$

9. Let X be a Bernoulli random variable with parameter $1/4$, and let $Y = 7X$. Compute the joint cumulative density function $F_{X,Y}$.

Given :

$p = 1/4$

$X \rightarrow$ Bernoulli Random Variable

$Y = 7X$

$P(X=1) = 1/4$

$P(X=0) = 1 - 1/4 = 3/4$

If $X=0 \Rightarrow Y = 7(0) \Rightarrow Y=0$

If $X=1 \Rightarrow Y = 7(1) \Rightarrow Y=7$

$\therefore$ Y takes value in 0, 7

Joint CDF : $F_{X,Y}(x,y) = P(X\le x, Y\le y)$

Case (i) : $x<0$, $y<0$

$F_{X,Y}(x,y) = 0$

$\because$ X and Y are non-negative

(ii) $0\le x<1$, $y<7$

$F_{X,Y}(x,y) = P(X=0, Y=0)$
$= P(X=0)$
$= 3/4$

(iii) $x\ge 1$, $y<7$

$F_{X,Y}(x,y) = P(Y=0)$
$= P(X=0)$
$= 3/4$

[X can be 0 or 1]

(iv) $x<1$, $y\ge 7$

$F_{X,Y}(x,y) = P(X=0)$
$= 3/4$

[Y can be 0 or 1]

(v) $x\ge 1$, $y\ge 7$

$F_{X,Y}(x,y) = P(X=0) + P(X=1)$
$= 3/4 + 1/4 = 1$

[X can be 0 or 1 & Y can be 0 or 7]

$$F_{X,Y}(x,y) = \begin{cases} 0, & x<0 \text{ or } y<0 \\ 3/4, & 0 \le x < 1 \text{ and } y < 7 \\ 3/4, & x \ge 1 \text{ and } y < 7 \\ 3/4, & x < 1 \text{ and } y \ge 7 \\ 1, & x \ge 1 \text{ and } y \ge 7 \end{cases}$$

10. Let X be the number of heads when flipping three fair coins. Let $Y = 1$ if $X \geq 1$, with $Y = 0$ if $X = 0$. Find the density function of Y.

Possible Values of X : 0, 1, 2, 3

$X \longrightarrow$ Binomial distribution

The probability of obtaining exactly k heads in 3 flips of a fair coin :

$$P(X=k) = \binom{3}{k}\left(\frac{1}{2}\right)^k\left(\frac{1}{2}\right)^{3-k}$$

$$= \binom{3}{k}\left(\frac{1}{2}\right)^3$$

$$\therefore P(X=k) = \frac{\binom{3}{k}}{8}$$

$$P(X=0) = \frac{\binom{3}{0}}{8} = \frac{1}{8} \qquad P(X=2) = \frac{\binom{3}{2}}{8} = \frac{3}{8}$$

$$P(X=1) = \frac{\binom{3}{1}}{8} = \frac{3}{8} \qquad P(X=3) = \frac{\binom{3}{3}}{8} = \frac{1}{8}$$

As, $Y = \begin{cases} 1, & X \ge 1 \\ 0, & X = 0 \end{cases}$

@ $Y=0 \Rightarrow P(Y=0) = P(X=0) = 1/8$

@ $Y=1 \Rightarrow P(Y=1) = P(X=1) + P(X=2) + P(X=3)$

$= 3/8 + 3/8 + 1/8$

$= 7/8$

$$\boxed{\therefore P(Y=y) = \begin{cases} 1/8, & y=0 \\ 7/8, & y=1 \end{cases}}$$

$\longleftarrow$ Indicates P(no head occuring)

$\longleftarrow$ Indicates P(atleast one head occurs)

11. Let X be the number showing on a fair six-sided die, so that $P(X = x) = 1/6$ for $x = 1, 2, 3, 4, 5$, and 6. Find the density function for Y.

X ?

<u>Given:</u>

$P(X=x) = 1/6 \quad \forall x \in \{1,2,3,4,5,6\}$

$$\boxed{f(x) = \begin{cases} 1/6, & x \in \{1,2,3,4,5,6\} \\ 0, & \text{otherwise} \end{cases}}$$

$F_X \longrightarrow X$

$Y = \alpha X + \beta = ?$

$F_X(x) = F = P(X \leq x)$

$F_Y(x) = P(Y \leq x)$

$= P(\alpha X + \beta \leq x)$

$= P(\alpha X \leq x - \beta)$

$= P\left(X \leq \frac{x-\beta}{\alpha}\right)$

$= F_X\left(\frac{x-\beta}{\alpha}\right)$

$\Rightarrow F_Y(x) = F_X\left(\frac{x-\beta}{\alpha}\right)$

$P \longrightarrow p_1, \ 1-p_1$

$F \longrightarrow p_2, \ 1-p_2$

$p_1 = \frac{x_1}{n}$

$\Rightarrow x_1 = np_1$

$x_2 = np_2$

$X = x_1 \,|\, x_2$

$= np_1 + np_2$

$= n\underbrace{(p_1 + p_2)}_{p'}$

$R \longrightarrow$ fred getting correct

$K \longrightarrow$ fred knowing answer

$p(K) = p$

$p(K|R) = \frac{P(K \cap R)}{P(R)}$

$= \frac{p}{P(R)} \in [0,1]$

$\geq p$ ②

$\rightarrow \underbrace{\text{know} \longrightarrow \text{correct}}$

$\rightarrow$

$P(R|K) = 1 = \frac{P(R \cap K)}{P(K)}$

$\Rightarrow 1 = \frac{P(R \cap K)}{p}$

$\Rightarrow P(R \cap K) = p$

$P(R|\bar{K}) = \frac{1}{n}$

$P(R) = \underbrace{p}\left(1 - \frac{1}{n}\right) + \frac{1}{n} \times 1$ know

$\frac{1}{n} = \frac{P(R \cap \bar{K})}{P(\bar{K})} = \frac{P(R \cap \bar{K})}{1-p}$

$\underbrace{P(R \cap \bar{K})} = \frac{1-p}{n}$

$p(K|R) = \left(\frac{p}{p + \frac{1-p}{n}}\right)$

$\underbrace{\text{knowledge}}_{p} + \underbrace{\text{luck}} = P(R)$

$p + P(R \cap \bar{K}) = P(R)$

$p + \frac{1-p}{n} = P(R)$

27/2

→ **The De-Moivre Laplace Limit theorem**:

Let X be a binomial distribution with parameters (n, p), Then [if n is sufficiently large], for any $a < b$

$$P\left(a < \frac{X - np}{\sqrt{np(1-p)}} < b\right) = \Phi(b) - \Phi(a)$$

Ex. Let X be the no. of times that a fair coin is flipped 40 times, lands heads. Find the value of $P(X = 20)$.

Sol:

$X \longrightarrow$ Binomial distribution (clearly)

From

i.e. $P(X=i) = \binom{n}{i} p^i (1-p)^{n-i}$ PnS-18

Alternatively:

$P(19.5 < X < 20.5)$ [Approximation $\longrightarrow$ Assumption] (19.9, 20.1)

$$= P\left(\frac{19.5 - 20}{\sqrt{10}} < \frac{X - 40(\frac{1}{2})}{\sqrt{40(\frac{1}{2})(1-\frac{1}{2})}} < \frac{20.5 - 20}{\sqrt{10}}\right)$$

$$= \Phi\left(\frac{1}{2\sqrt{10}}\right) - \Phi\left(\frac{-1}{2\sqrt{10}}\right)$$

$= \Phi(0.158) - \Phi(-0.158)$ $\qquad \Phi(-0.158) = 1 - \Phi(0.158)$

$= [(0.56356) \times 2] - 1$

$= 0.12712$ $\qquad [\Phi(0.158) \approx \Phi(0.16)]$

Ex. The ideal size of a first-year class at a particular college is 150 students. Knowing from past experiences, On an average, 30% of those accepted for admissions in the college actually attend, uses a policy of approving the applications of 450 students. Compute the probability that more than 150 first-year students attend this college.

Sol:

Let X be the no. of first-year students attending the college.

$P(X > 150) = ?$

$$= P\left(\frac{X - 450 \times (\frac{1}{3})}{\sqrt{450 \cdot (\frac{1}{3})(\frac{2}{3})}} > \frac{150 - 150}{\sqrt{100}}\right)$$

$$= P\left(\frac{X - 150}{10} > 0\right)$$

$= 1 - P(Z < 0)$

$= 1 - 0.5$

$= 0.5$

3/3

→ **Moments :**

- **Observations :**

Let X be a random variable, we say that $c \in \mathbb{R}$ is a median of X if

$P(X \geq c) \geq \frac{1}{2}$ and

$P(X \leq c) \geq \frac{1}{2}$

$$P(X \geq c) + P(X \leq c) = 1 \implies P(X \geq c) = \tfrac{1}{2}, \quad P(X \leq c) = \tfrac{1}{2}$$

But @c ⇒ Jump / Discontinuity : (Multiple Medians)

$$P(X \geq c) \geq \tfrac{1}{2}$$

$$P(X \leq c) \geq \tfrac{1}{2}$$

A $c_1 \in \mathbb{R}$ is called a Mode of X if

$$P(X = c_1) \geq P(X = x) \quad \forall x$$

But if X → continuous ⇒ $P(X=x)$ $\forall x$ ⇒ Take neighbourhood of x as an interval

∴ Suppose f is the PDF of X.

Mode $= \max(f(x))$

Theorem :

Let X be a random variable with mean μ and let m be a median of X.

Then,

(a) The value of c that minimizes the mean squared error is $E(X-c)^2$ is $c = \mu$

(b) The value of c that minimizes the mean absolute error is $E|X-c|$ is $c = m$.

Def^n: Let X be a random variable with mean μ and variance σ^2. For any $n \in \mathbb{N}$, the n^{th} moment of X is $E[X^n]$.

Observe that the 1^{st} Moment is the expected value of X.

$E[X^n] = ?$ $\forall n \geq 2$

5/3

Recall : X

$E[X^n] \longrightarrow n^{th}$ moment of X

1^{st} moment is the mean.

$E[(X-\mu)^2]$ is the variance.

Def^n :

Let X, Y be two discrete random variables.

We say that X and Y are independent if

$$P(X=a, Y=b)$$

$$= P(X=a) \cdot P(X=b) \qquad [a, b \in \mathbb{R}]$$

If X and Y are continuous,

$$P(X \in A, y \in B) = P(X \in A) \cdot P(Y \in B)$$

$$[\forall A, B \subseteq \mathbb{R}]$$

✶ [Central Limit Theorem]

→ Moment generating function :

The moment generating functions of an random variable of X is

$$M(t) = E[e^{tx}]$$

as a function of t, if this is finite on some open interval $(-a, a)$ containing 0.

- Ex. MGF

Let X be a Bernoulli Random variable, Then, the MGF of X is

$$M(t) = E[e^{tx}]$$

Sol : $E[x] = \sum_x x \cdot f(x)$

PnS-15

$$\Rightarrow E[e^{tx}] = \sum_x e^{tx} \cdot P(x)$$

$$= e^{t(0)} \cdot P(0) + e^{t(1)} \cdot P(1)$$

$$= e^0 \cdot (1-p) + e^t (p)$$

$$= 1 - p + e^t \cdot p$$

$$\boxed{\therefore M(t) = (1-p) + e^t (p)}$$

$M(t) = p \cdot e^t + q$, where , $q = 1-p$

Ex. Let $X \sim Unif(a,b)$, Then , the MGF of X is

$$M(t) = E[e^{tx}]$$

$$= \int_a^b e^{tx} \cdot P(x)\, dx$$

$$= \int_a^b e^{tx} \cdot \frac{1}{b-a}\, dx$$

$$= \frac{1}{b-a} \int_a^b e^{tx} \cdot dx$$

$$= \frac{1}{b-a} \left(\frac{1}{t} e^{tx}\right]_a^b$$

$$= \frac{1}{t(b-a)} \cdot (e^{tb} - e^{ta})$$

$$= \frac{e^t(e^b - e^a)}{t(b-a)}$$

$$\Rightarrow \boxed{M(t) = \frac{e^t(e^b - e^a)}{t(b-a)}}$$

→ Theorem :

If two random variables have the same MGF, they must have the same distribution , i.e. same PDFs

No proof , Beyond the course (Too advanced !)

→ Theorem :

If X and Y are independent , then the MGF of $X+Y$ is

$$M_{X+Y}(t) = M_X(t) \cdot M_Y(t)$$

Proof :

$X : S \to \mathbb{R}$

$Y : S \to \mathbb{R}$

$X + Y : S \to \mathbb{R}$

– To be continued later –

6/4

- X and Y are two random variables
  $\Rightarrow g(X, Y)$ is also a random variable

Special $\longrightarrow$ $X+Y$ is also a random variable

$$X : S \longrightarrow \mathbb{R}$$
$$Y : S \longrightarrow \mathbb{R}$$

$$\Rightarrow X+Y(s) = X(s) + Y(s) \quad , \quad s \in S$$

$$S \longrightarrow \mathbb{R} \times \mathbb{R} \xrightarrow{+} \mathbb{R}$$

$$S \xrightarrow{(x,y)} (X(s), Y(s)) \longrightarrow \mathbb{R}$$

$$+ \circ (X, Y) s$$
$$+ \circ (X(s), Y(s)) = X(s) + Y(s)$$

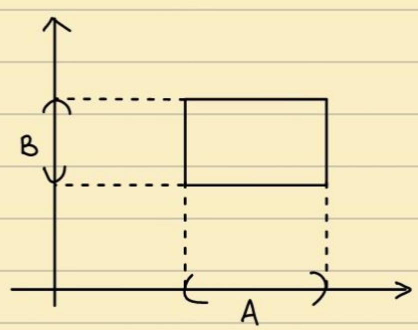
[$f \circ g(x)$ notation]

$$S \xrightarrow{(x,y)} \mathbb{R} \times \mathbb{R} \xrightarrow{g(x,y)} \mathbb{R}$$

Let P_X be the PDF of X and
P_Y be the PDF of Y.
Define a function $P_{X,Y} : \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R}$ is called the Joint Probability distribution function of X and Y if
$P_{X,Y}(x, y) = P(X=x, Y=y) \ \forall x, y$, if x & y are discrete.

If X & Y are continuous, a function $f_{X,Y}(x,y)$ is called the Joint Probability distribution of X and Y if

$$P(X \in A, Y \in B) = \iint_{B\,A} f_{X,Y}(x,y)\, dx \cdot dy \quad , \text{ for any } A \subseteq \mathbb{R} \ \& \ B \subseteq \mathbb{R}$$



Ex. X is uniform distribution on $[1, n]$
Y is uniform distribution on $[1, m]$

$$P(X = i) = P_X(i) = 1/n$$
$$P(Y = j) = P_Y(j) = 1/m$$

- X and Y be two discrete random variable with PDF P_X and P_Y.
  Let $P_{X,Y}$ be the Joint PDF of X, Y

  Fix X as x
  $$P_{X,Y}(1) + P_{X,Y}(2) + P_{X,Y}(3) + \cdots$$
  $$= P_X(x)$$

12/3

- X, Y are discrete,

$$\sum_y P_{X,Y}(x, y) = P_X(x)$$

$$\sum_x P_{X,Y}(x, y) = P_Y(y)$$

Recall, any two random variables X, Y are said to be independent, if
$$P(X \in A, Y \in B)$$
$$= P(X \in A) \cdot P(Y \in B) \quad \text{where } A, B \subseteq \mathbb{R}$$

Special case : X and Y are discrete and independent,
$$P_{X,Y}(x, y) = P_X(x) \cdot P_Y(y)$$

$$P_{X,Y}(x, y) = P(X = x) \cdot P(Y = y)$$
$$= P_X(x) \cdot P_Y(y)$$

- If X and Y are continuous and independent, then
  $$f_{X,Y}(x, y) = f_X(x) \cdot f_Y(y)$$

  Outline of the proof :
  $$F_{X,Y}(x, y) = P(X \le x, Y \le y)$$

  Left Joint CDF of X, Y :
  $$= P(X \le x, Y \le y)$$
  $$= F_X(x) \cdot F_Y(y)$$

Differentiate :

$$f_{X,Y}(x,y) = f_X(x)\cdot f_Y(y)$$

Then, two random variables X and Y are independent iff $f_{X,Y} = f_X \cdot f_Y$

H.W : Outline of Proof (Hint : Double Integral)

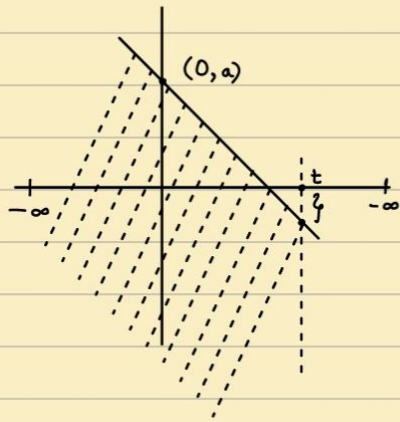
→ Sums of independent random variables :

Suppose that X and Y are independent, continuous random variable with PDF f_X and f_Y. Then, consider $X+Y$.

Over the PDF of $X+Y$, i.e. What is f_{X+Y}

Sol: $F_{X+Y}(a) = P(X+Y \le a)$, $a \in \mathbb{R}$

$$= \iint\limits_{X+Y\le a} f_{X+Y}(x,y)\,dx\,dy$$



$$= \int_{-\infty}^{\infty}\int_{-\infty}^{y-a} f_{X,Y}(x,y)\,dx\cdot dy$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{y-a} f_X(x)\cdot f_Y(y)\,dx\,dy$$

$$= \int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} f_X(x)\,dx\right) f_Y(y)\cdot dy$$

$$= \int_{-\infty}^{\infty} F_X(a-y)\cdot F_Y(y)\cdot dy$$

$$F_{X+Y}(a) = \int_{-\infty}^{\infty} F_X(a-y)\, f_Y(y)\,dy$$

$$\Rightarrow \frac{d}{da}\left(F_{X+Y}(a)\right) = \int_{-\infty}^{\infty}\left(\frac{d}{da}F_X(a-y)\right)\cdot F_Y(y)\,dy$$

H.W

(1) If X and Y be uniform over $(0,1)$. Find the PDF of $X+Y$

(2) If X and Y be two standard normal distribution, Then, Is $X+Y$ also a standard normal distribution.

13/3

$$X \sim \frac{1}{\sqrt{2\pi}\cdot\sigma}\cdot e^{-(x-\mu)^2/2\sigma^2}$$

$$Y \sim \frac{1}{\sqrt{2\pi}\cdot\sigma'}\cdot e^{-(x-\mu')/2\sigma'^2}$$

$$X + Y \sim ?$$

i.e. $X \sim f_X$

$Y \sim f_Y$

$$f_{X+Y} \sim \int_{-\infty}^{\infty} f_X(a-y)\cdot f_Y(y)\, dy = f_{X+Y}(a)$$

Ex. Let $X \sim U(0,1)$

$$f_X = \begin{cases} 1, & \text{if } 0 < x < 1 \\ 0, & \text{otherwise} \end{cases}$$

$Y \sim U(0,1)$

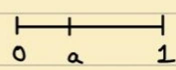
$$f_Y = \begin{cases} 1, & \text{if } 0 < y < 1 \\ 0, & \text{otherwise} \end{cases}$$

$$f_{X+Y}(a) = \int_{-\infty}^{\infty} f_X(a-y)\cdot f_Y(y)\, dy$$

$X + Y$ takes values in $(0, 2)$

$$= \int_{0}^{2} f_X(a-y)\cdot f_Y(y)\, dy$$

$[\text{in } (1,2) \Rightarrow f_Y(y) = 0]$



$$= \int_{0}^{1} f_X(a-y)\cdot dy$$

Case (i): If $0 \le a \le 1$,

$$= \int_{0}^{a} f_X(a-y)\cdot (-dy)$$

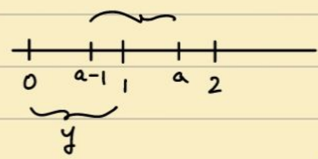
$$= \int_0^a f_Y(a-y)\, dy \qquad [y > a]$$

$$= \int_0^a dy$$

$$= a$$

case -(ii) : If $1 \le a \le 2$

$$= \int_0^1 f_X(a-y)\, dy$$



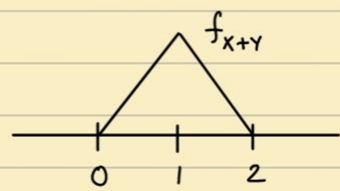
$$= \int_1^{a-1} f_X(a-y) \cdot (-dy)$$

$y \sim 0, 1$
$a-y \sim a, a-1$

$$= \int_{a-1}^{1} f_X(a-y)$$

$$= \int_{a-1}^{1} dy = 2-a$$

$$f_{X+Y}(a) = \begin{cases} a, & 0 \le a \le 1 \\ 2-a, & 1 < a < 2 \\ 0, & a > 2 \end{cases}$$



Not Uniform

Theorem : If $X_1 \sim N(\mu_1, \sigma_1)$
& $X_2 \sim N(\mu_2, \sigma_2)$, Then,

$$X_1 + X_2 \sim N(\mu_1 + \mu_2, \sigma_1 + \sigma_2)$$

Proof : H.W

→ <u>Conditional distributions</u>:

Let X, Y be two continuous random variable with PDFs f_X and f_Y. Then, the conditional density function of X, given $Y = y$, is defined as

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}, \quad \text{if } f_Y(y) > 0$$

If X and Y are discrete,

$$P_{X|Y}(x|y) = \frac{P_{X,Y}(x,y)}{P_Y(y)}, \quad \text{if } P_Y(y) > 0$$

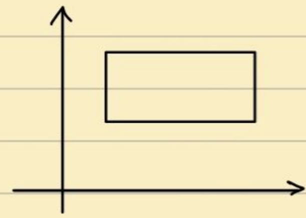
17/3

Theorem: Let X_1 and X_2 be two continuous random variable with Joint PDF f_{X_1,X_2}.

Let $Y_1 = g_1(X_1, X_2)$

$Y_2 = g_2(X_1, X_2)$

$$g_1 : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$$
$$g_2 : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$$



Suppose

(i) The equations $y_1 = g_1(x_1, x_2)$ and $y_2 = g_2(x_1, x_2)$ can be uniquely for x_1, x_2 in terms of y_1, y_2

i.e. $x_1 = h_1(y_1, y_2)$ &

$x_2 = h_2(y_1, y_2)$

(ii) g_1, g_2 have continuous partial derivatives with

$$J(x_1, x_2) = \begin{vmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} \\ \frac{\partial g_2}{\partial x_2} & \frac{\partial g_2}{\partial x_2} \end{vmatrix} \neq 0 \qquad \forall (x_1, x_2)$$

$$f_{Y_1,Y_2}(y_1, y_2) = f_{X_1,X_2}(x_1, x_2) \cdot \frac{1}{|J(x_1,x_2)|}$$

Ex. Let X_1, X_2 be two continuous random variables with joint PDF f_{X_1,X_2}

Define $Y_1 = X_1 + X_2 = g_1(X_1, X_2)$

$Y_2 = X_1 - X_2 = g_2(X_1, X_2)$

Here,

$g_1(X_1, X_2) = x_1 + x_2 = y_1$

$g_2(X_1, X_2) = x_1 - x_2 = y_2$

- The system is consistent.

$$J(x_1, x_2) = \begin{vmatrix} \dfrac{\partial g_1(x_1,x_2)}{\partial x_1} & \dfrac{\partial g_1(x_1,x_2)}{\partial x_2} \\ \dfrac{\partial g_2(x_1,x_2)}{\partial x_1} & \dfrac{\partial g_2(x_1,x_2)}{\partial x_2} \end{vmatrix}$$

$$= \begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix} \neq 0$$

$$f_{Y_1,Y_2}(y_1, y_2) = f_{X_1,X_2}(x_1, x_2) \cdot \frac{1}{(-2)}$$

- Let X_1, X_2 be two random variables & $Y = g(X_1, X_2)$

$$E[Y] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x_1, x_2) \cdot f_{X_1,X_2}(x_1, x_2)\, dx_1\, dx_2$$

Ex. $g(x_1, x_2) = x_1 + x_2$

Then, $Y = x_1 + x_2$

$$E[Y] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x_1 + x_2) \cdot f_{X_1,X_2}(x_1, x_2)\, dx_1\, dx_2$$

$$= \iint x_1 \cdot f_{X_1,X_2}(x_1,x_2)\, dx_1 \cdot dx_2 + \iint x_2 \cdot f_{X_1,X_2}(x_1,x_2)\, dx_1 \cdot dx_2$$

$$= E[X_1] + E[X_2]$$

In General,

$$E[X_1 + X_2 + \dots + X_n] = E[X_1] + E[X_2] + \dots + E[X_n]$$

Ex. If a group of N people throw their hats in the centre of a room. The hats are mixed up & each person selects one. Find the expected number of people select their own hats.

Sol: Let X denote the no. of matches.

$$X = X_1 + X_2 + \dots + X_n$$

where,

$$X_i = \begin{cases} 1, & \text{if the } i^{th} \text{ person gets his/her own hat} \\ 0, & \text{otherwise} \end{cases}$$

$\hookrightarrow$ Bernoulli Random variable

$$E[X] = E[X_1] + E[X_2] + \ldots\ldots + E[X_n]$$
$$= 1/N + 1/N + \ldots\ldots + 1/N$$
$$= 1$$

$\therefore$ Only 1 person is expected to select his/her own hat.

19/3

→ **Covariance:**

Let X and Y be two random variables. Then, the covariance between X and Y, denoted by cov(X, Y) is defined by

$$\mathrm{cov}(X, Y) = E\big[(X - E[X])(Y - E[Y])\big]$$

Remark:

$$\begin{aligned}\mathrm{cov}(X, Y) &= E\big[XY - X\cdot E[Y] - Y\cdot E[X] + E[X]\cdot E[Y]\big] \\ &= E[XY] - E[X\cdot E[Y]] - E[Y\cdot E[X]] + E\big[E[X]\cdot E[Y]\big] \\ &= E[XY] - E[Y]\cdot E[X] - \cancel{E[X]\cdot E[Y]} + \cancel{E[X]\cdot E[Y]} \\ &= E[XY] - E[X]\cdot E[Y]\end{aligned}$$

$$\therefore \mathrm{cov}(X, Y) = E[XY] - E[X]\cdot E[Y]$$

Remark:

If X, Y are independent, cov(X, Y) = 0

$$\begin{aligned}E[XY] &= \iint xy\, f_{X,Y}(x, y)\, dx\, dy \\ &= \int x\, f_X(x)\, dx \cdot \int y \cdot f_Y(y)\, dy \\ &= E[X]\cdot E[Y]\end{aligned}$$

$$\left[E[g(X,Y)] = \iint g(x,y)\, f_{X,Y}(x,y)\, dx \cdot dy \right]$$

$$\therefore E[XY] = E[X]\cdot E[Y]$$
$$\therefore \mathrm{cov}(X,Y) = 0$$

Properties:

(i) $\mathrm{cov}(X, Y) = \mathrm{cov}(Y, X)$

(ii) $\mathrm{cov}(X, X) = \mathrm{var}(X)$

(iii) $\mathrm{cov}(aX, Y) = a\cdot \mathrm{cov}(X, Y)$

(iv) $\mathrm{cov}\left(\sum_{i=1}^{n} X_i, \sum_{j=1}^{n} Y_j\right) = \sum_i \sum_j \mathrm{cov}(X_i, Y_j)$

Proof:

Use the formula

$$E[XY] - E[X]\cdot E[Y] = \mathrm{cov}(X, Y)$$

Remember:

- $E[X+Y] = E[X] + E[Y]$

But:

- $Var(X+Y) = Var(X) + Var(Y) + 2\,cov(X,Y)$

Proof:

$$\begin{aligned} Var(X+Y) &= cov(X+Y, X+Y) \\ &= cov(X,X) + cov(X,Y) + cov(Y,X) + cov(Y,Y) \\ &= Var(X) + 2 \cdot cov(X,Y) + Var(Y) \end{aligned}$$

$$\therefore\ Var(X+Y) = Var(X) + Var(Y) + 2\,cov(X,Y)$$

- Example:

Compute the variance of a binomial distribution with parameters n and p.

Sol:

Let $X_1, X_2, X_3, \ldots\ldots, X_n$ be n independent Bernoulli distributions with parameter p.

Then, $X = X_1 + X_2 + \ldots\ldots + X_n$

$$\begin{aligned} Var(X) &= Var\left(\sum_{i=1}^{n} X_i\right) \\ &= \sum_{i=1}^{n} Var(X_i) + \underbrace{\sum_{i<j}\sum cov(X_i, X_j)}_{\hookrightarrow\ 0\ \because\ \text{Independent}} \\ &= \sum_{i=1}^{n} Var(X_i) \\ &= \sum_{i=1}^{n} p(1-p) \\ &= np(1-p) \end{aligned}$$

Recall:

- $X \rightarrow$ no. of people who select their own hat

Compute $Var(X)$

Sol: $X = \sum_{i=1}^{n} X_i$, where

$$X_i = \begin{cases} 1, & \text{if the } i^{th} \text{ person selects their own hat} \\ 0, & \text{otherwise} \end{cases}$$

$$Var(X) = Var\left(\sum_{i=1}^{n} X_i\right)$$

$$= \sum_{i=1}^{n} Var(X_i) + 2 \sum\sum_{i<j} cov(X_i, X_j)$$

X_i, X_j are independent

$$\Rightarrow P(X_1 = 1, X_j = 1) = P(X_i = 1) \cdot P(X_j = 1)$$

$$X_i X_j = \begin{cases} 1, & \text{if } X_i = 1 \ \& \ X_j = 1 \\ 0, & \text{otherwise} \end{cases}$$

$$\begin{aligned} P(X_i = 1, X_j = 0) &= P(X_i = 1) \cdot P(X_j = 0) \\ &= 1/n \cdot (n-1)/n \\ &= (n-1)/n^2 \end{aligned}$$

$$P(X_i = 1, X_j = 1) \neq 1/n^2$$

20/3

Recall :

$$X = \sum_{i=1}^{n} X_i, \text{ where}$$

$$X_i = \begin{cases} 1, & \text{if the } i^{th} \text{ person selects their own hat} \\ 0, & \text{otherwise} \end{cases}$$

$$\begin{aligned} Var(X) &= Var\left(\sum_{i=1}^{n} X_i\right) \\ &= \sum_{i=1}^{n} Var(X_i) + 2 \sum\sum_{i<j} cov(X_i, X_j) \end{aligned}$$

Recall ,

$$Var(X_i) = \frac{1}{N}\left(1 - \frac{1}{N}\right)$$

$$cov(X_i, X_j) = E[X_i X_j] - E[X_i] \cdot E[X_j]$$

$$X_i X_j = \begin{cases} 1, & \text{if } X_i = 1 \text{ and } X_j = 1 \\ 0, & \text{otherwise} \end{cases}$$

$$\begin{aligned} E[X_i X_j] &= 1 \cdot P(X_i X_j = 1) + 0 \cdot P(X_i X_j = 0) \\ &= P(X_i X_j = 1) \\ &= P(X_i = 1, X_j = 1) \\ &= P(X_i = 1) \cdot P(X_j = 1 \mid X_i = 1) \end{aligned}$$

$$P(A \cap B) = P(A) \cdot P(B|A)$$
$$= 1/n \cdot 1/(n-1)$$
$$\neq E[X_i] \cdot E[X_j]$$

$$cov(X_i, X_j) = E[X_i, X_j] - E[X_i] \cdot E[X_j]$$
$$= 1/n \cdot 1/(n-1) - 1/n^2$$

$$Var(X) = \sum_{i=1}^{n} Var(X_i) + 2 \sum_{i<j}\sum cov(X_i, X_j)$$
$$= \sum_{i=1}^{n} \frac{1}{n}(1 - 1/n) + 2 \sum_{i<j}\sum \left(1/n \cdot 1/(n-1) - 1/n^2\right)$$
$$= n(1/n)(1 - 1/n) +$$
$$=$$

$$X = X_1 + X_2 + X_3 + \ldots\ldots + X_n \Longleftarrow \text{Advantage}$$

- Limit Theorem :

If X is a random variable that takes only non-negative values, then for any $a > 0$,

$$P(X \geq a) \leq \frac{E[X]}{a} \qquad \text{(Markov's Inequality)}$$

Ex. Suppose that it is known that the number of items produced in a factory during a week is a random variable with mean 50.

(a) What can we say about the probability that this weeks production will exceed 75.

$$P(X \geq 76) = \frac{E[X]}{76} = \frac{50}{76}$$
$$\approx 0.658$$

Proof:

For any $a > 0$,

$$\text{let } I = \begin{cases} 1, & \text{if } X \geq a \\ 0, & \text{if } X < a \end{cases}$$

$$E[I] = 1 \cdot P(I=1) + 0 \cdot P(I=0)$$
$$= P(I=1)$$

$$= P(X \geq a)$$

$$I \leq \frac{X}{a} \implies E[I] \leq E\left[\frac{X}{a}\right]$$

$$I(s) \leq \frac{X(s)}{a}$$

$$E[I] \leq \frac{1}{a} E[X]$$

$$\Rightarrow P(X \geq a) \leq \frac{1}{a} \cdot E[X]$$

- Chebyshev's Inequality :

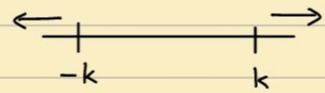
  If X is a random variable with finite mean μ and variance σ^2, then for any $k > 0$,

$$P(|X-\mu| \geq k) \leq \left(\frac{\sigma}{k}\right)^2$$

$$|X - \mu| \geq k$$

$$\text{iff } X-\mu \geq k \quad \text{OR} \quad X - \mu \leq -k$$

$$\implies X \geq \mu + k \quad \text{OR} \quad X \leq \mu - k$$



$$P\Big(X \geq \mu+k) \text{ OR } (X \leq \mu-k)\Big) \leq \left(\frac{\sigma}{k}\right)^2$$

Recall:

(b) If the variance of a week's production is known to equal 25, Then, what can be said about the probability that this week's production will be 40 and 60 ?

Sol: $P(40 \leq X \leq 60)$

$$= P(-10 \leq X - 50 \leq 10)$$
$$= P\left(|X-50| \leq 10\right)$$
$$= 1 - P(|X-50| \geq 10)$$
$$\geq 1 - \frac{25}{10^2}$$
$$\geq 3/4$$

Proof:

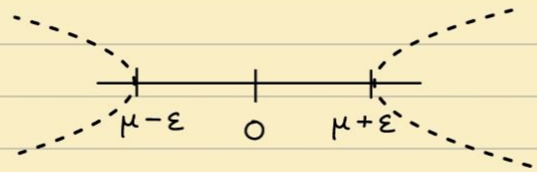
Let $Y = (X-\mu)^2$

$$P(Y \geq k^2) = P\left((X-\mu)^2 \geq k^2\right)$$
$$= P(|X-\mu| \geq k)$$
$$\leq \frac{E[(X-\mu)^2]}{k^2} = \frac{\sigma^2}{k^2}$$

- The weak law of large numbers :

  Let $X_1, X_2, \ldots, X_n$ be a sequence of independent and identically distributed random variables, each having finite mean $E[X] \le \mu$. Then, for any $\varepsilon > 0$.

$$P\left\{ \left| \frac{X_1 + X_2 + \cdots + X_n}{n} - \mu \right| \ge \varepsilon \right\} \longrightarrow 0 \quad \text{As } n \to \infty$$



1/4

- The central Theorem :

  $X_1, X_2, X_3, \ldots$ are independently identically distributed
  with $E[X_i] = \mu_i$
  $Var(X_i) = \sigma_i^2$

  Then, for any $n \ge 0$:

$$P\left\{ \frac{X_1 + X_2 + X_3 + \cdots + X_n - n\mu}{\sigma\sqrt{n}} \le a \right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{-x^2/2}\, dx \qquad \text{as } n \to \infty$$

Q) The no. of students that enroll in a course is a Poisson with mean 100. If the no. of professors ≥ 120 : Two seperate sections
no. of professors < 120 : One seperate section.
Find the probability that the professor will have to teach 2 sections.

Sol: Let X be the no. of students who enroll in a course

$$P(X \ge 120) = \sum_{n=120}^{\infty} P(X = n)$$

$$= \sum_{n=120}^{\infty} \frac{e^{-100} \cdot 100^n}{n!}, \qquad \text{here } \lambda = 100$$

$$= 1 - \sum_{n=0}^{119} \frac{e^{-100} \cdot 100^n}{n!}$$

Approx :

$P(X \ge 120)$ [Use CLT]

$$= P\left(\frac{X - 100}{\sqrt{100}} \ge \frac{120 - 100}{\sqrt{100}} \right)$$

$$= P(Z \ge 2)$$

$= 1 - P(Z < 2)$ $\quad$ [Poisson : $|\mu| = |\sigma^2|$]

$= 0.228$ $\quad$ [$n = 1$]

Ex. Let X_i , $i = 1, 2, 3, \ldots, 10$ be independent random variables each uniformly distributed over $(0, 1)$

Calculate an approximation $P\left(\sum_{i=1}^{10} X_i > 6\right)$

Sol : $P\left(\sum_{i=1}^{10} X_i > 6\right)$

By CLT,

$$P\left(\frac{\sum_{i=1}^{10} X_i - 10 \times 1/2}{\sqrt{10 \times 1/12}} > \frac{6 - 5}{\sqrt{10/12}}\right)$$

$$= P\left(Z > \sqrt{\frac{12}{10}}\right)$$

$$= P\left(Z > \sqrt{1.2}\right)$$

$$= 0.16$$

- CLT for independent random variable :

Let $X_1, X_2, X_3, \ldots\ldots, X_n$ be a sequence of independent random variables having mean $E[X_i] = \mu_i$ and the variance , $Var(X_i) = \sigma_i^2$

If $P(|X_i| < M) = 1$ $\forall$ i, and $\sum_{i=1}^{\infty} \sigma_i^2 = \infty$

[for some M]

Then ,

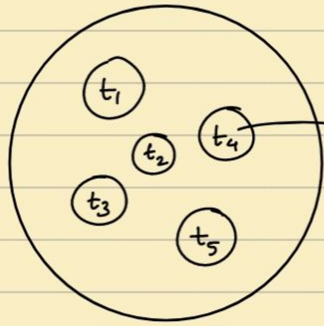
$$P\left(\frac{\sum_{i=1}^{n} (X_i - \mu_i)}{\sqrt{\sum_{i=1}^{n} \sigma_i^2}} \leq a\right) \longrightarrow \phi(a) \quad \text{as } n \to \infty$$

Statistics

→ **Sampling Theory** :

Population : The group of individuals under study is called population or universe.

X



P : parameter – μ

Statistic – $\bar{x}$ (Charecteristic of a Sample)

$$E[X] = \mu$$

- **Types of Sampling** :
  (1) Purposive Sampling – ✗
  (2) Random Sampling – ✓

  We always choose this

- **Unbiased Estimate** :

  A statistic $t = f(x_1, x_2, \ldots, x_n)$,

  a function of the sample values $x_1, x_2, \ldots\ldots, x_n$ is called unbiased estimate of the population parameters.

$$E[t] = \theta$$

– Standard Error : $E[t] - \theta$

- **Standard Error** : (SE)

  $\bar{x}$: Sample mean

$$SE = \sigma/\sqrt{n}$$

n : Sample size

σ^2 : Population variance

s^2 : Sample variance

Sample proportion : p

$$SE = \sqrt{PQ/n}$$

$Q : 1 - P$

Sample Standard Deviation

$$SE = \sqrt{\sigma^2/2n}$$

- <u>Standard Error</u>:

The standard deviation of the sampling distribution of a statistic is called a Standard Error

$$Z = \frac{t - E[t]}{\sqrt{Var(t)}} \sim N(0,1)$$
$$= \frac{t - E[t]}{SE} \sim N(0,1)$$

- <u>Normal Test of Significance</u>:

Procedure :

(1) NULL Hypothesis

Setup the NULL Hypothesis H_0

(2) Alternative Hypothesis

E.g. Negation of the NULL choice

H_1 = Alternative to H_0

(3) Level of Significance

Z_α

3 - 5% : Permissible error range

10% : Too much

(4) Test Statistic

Compute the test statistic

$$Z = \frac{t - E[t]}{SE}$$ under the NULL hypothesis

(5) Conclusion

We compare Z with Z_α.

Ex. A dice is thrown 9000 times and a throw of 3 or 4 is observed 3240 times. Show that the dice cannot be regarded as unbiased.

<u>Sol</u>:

(1) Null Hypothesis

The dice is an unbiased one :

$H_0 = P(3 \text{ or } 4)$

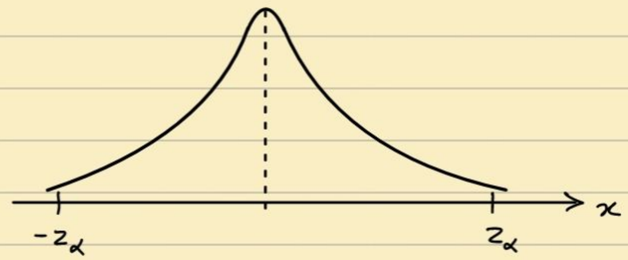
$P = 2/6 = 1/3$

(2) Alternate Hypothesis :

$$H_1 : p \neq 1/3$$

(3) Fix the level of Significance

$$z_\alpha = ?$$



(4) Test for Statistic

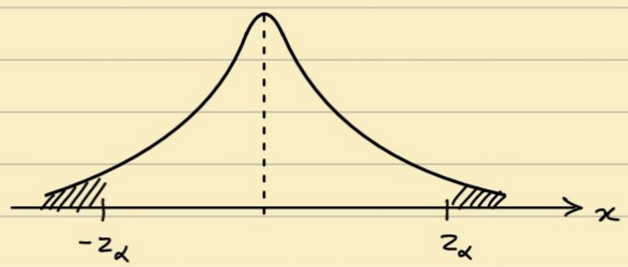
$$9000 \times 1/3 = 3000$$

5/4

→ Critical Values or Significant Values :

A critical value of the test statistic at level of significance z_α is determined by

$$P(|z| > z_\alpha) = \alpha$$



Two - tailed Test	$z_\alpha = 2.58$
Right - tailed Test	$z_\alpha = 2.33$
Left - tailed Test	$z_\alpha = -2.33$

For the 5% level of Significance, z_α

Two - tailed Test	$z_\alpha = 1.96$
Right - tailed Test	$z_\alpha = 1.645$
Left - tailed Test	$z_\alpha = -1.645$

Q) A dice is thrown 9000 times & we get 3 or 4, 3240 times. Show that the dice is biased.

Sol^n: Null Hypothesis $\rightarrow$ $\underbrace{\text{Dice is unbiased}}_{H_0}$

Probability of getting 3 or 4 $\Rightarrow P = 1/3$

Alternate Hypothesis $\longrightarrow$ $\underbrace{\text{Dice is not unbiased}}_{H_1}$

Fix $z_\alpha = 2.58$

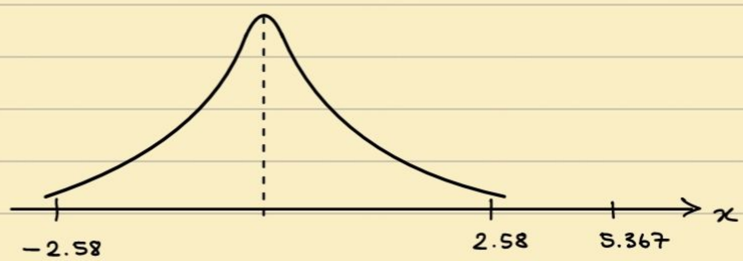
Test statistic :

$$Z = \frac{X - E[X]}{SE[X]}$$

$$\Rightarrow Z = \frac{3240 - (1/3 \times 9000)}{\sqrt{np(1-p)}} \qquad [q = 1 - p]$$

$$= \frac{240}{\sqrt{9000 \times 1/3 \times 2/3}}$$

$$= \frac{240}{\sqrt{2000}} = \frac{24}{\sqrt{20}} = 5.367$$



$\therefore$ We say H_0 is rejected, Since $-2.58 < 5.367 < 2.58$ is false

$\therefore H_1 \longrightarrow$ Accepted

$\therefore$ Dice is biased.

Ex. In a sample of 1000 people in Maharashtra, 540 are rice eaters and the rest are wheat eaters. Can we assume that both rice and wheat eaters are equally popular in this at 1% level of Significance

Solⁿ:

Null Hypothesis :

H_0 : Rice and Wheat eaters are equally popular

P : Population Proportion of rice eaters in Maharashtra = $\frac{1}{2}$

Alternative Hypothesis :

$H_1 : P \neq \frac{1}{2}$

Fix level of Significance , $z_{\alpha} = 2.58$

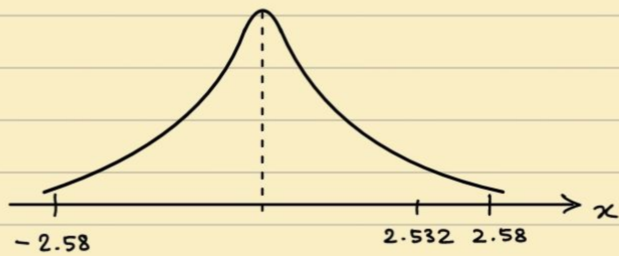
Test Statistic :

$$Z = \frac{X - E[X]}{SE[X]}$$

$$= \frac{540 - 500}{\sqrt{np(1-p)}}$$

$$= \frac{40}{\sqrt{1000 \times \frac{1}{2} \times \frac{1}{2}}}$$

$$= 2.532$$

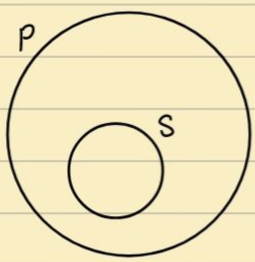


$\because -2.58 < 2.532 < 2.58$

$\therefore H_0 \longrightarrow$ Accepted

$\Rightarrow$ P(rice eaters) = P(wheat eaters) = $\frac{1}{2}$

7/4



Let X be the sample distribution

$X \longrightarrow$ Binomial

$E[X] = nP$ [P : Sample Probability]

$Var[X] = nP(1-P)$

Test Statistic :

$$Z = \frac{X - E[X]}{SE(X)}$$

Remark 1 :

Let X be the no. of units in S with that charecteristic

$$p = X/n$$

$$\text{Proportion of the population} = E[p] = E[X/n] = 1/n \cdot E[X] = 1/n \cdot nP = P$$

$$\begin{aligned} Var(P) &= Var(X/n) \\ &= 1/n^2 \cdot Var(X) \\ &= 1/n^2 \cdot nP \cdot (1-P) \\ &= P(1-P)/n \end{aligned}$$

$$SE(p) = \sqrt{\frac{P(1-P)}{n}}$$

Ex. Twenty people were attacked by a disease and only 18 survived. Will you reject the hypothesis that the survival rate, if attacked by this disease, is 85% in favour of the hypothesis that it is more, at 5% level.

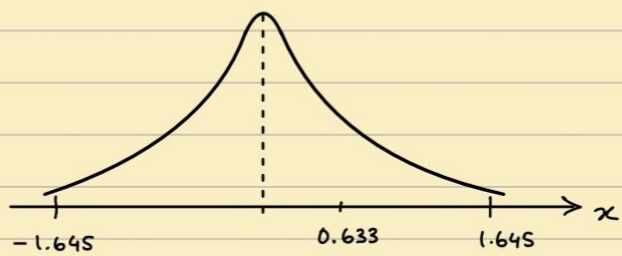
Solⁿ: Null Hypothesis : $H_0 \Longrightarrow P = 0.85$

Alternate Hypothesis : $H_1 \Longrightarrow P > 0.85$

Fix $z_\alpha = 1.645$

Test Statistic :

$$\begin{aligned} Z &= \frac{p - E[P]}{SE(p)} \\ &= \frac{p - P}{\sqrt{\frac{P(1-P)}{n}}} \\ &= \frac{0.90 - 0.85}{\sqrt{\frac{0.85 \times 0.15}{20}}} \\ &= 0.633 \end{aligned}$$



$\& \because -1.645 < 0.633 < 1.645$

$\therefore$ We accept H_0

$\therefore P = 0.85$

Remark 2 :

Probable limits for X are :

$$E[X] \pm 3 \cdot \sqrt{Var(X)}$$

$$= E[X] \pm 3 \cdot SE(X)$$

If $X = p$:

$$= E(p) \pm 3\, SE(X)$$

$$= p \pm 3\sqrt{\frac{p(1-p)}{n}}$$

Ex. A random sample of 500 pineapples are taken from a large consignment and 65 were found to be bad. Show that the SE of the proportion of the bad ones in a sample of a size is 0.015 and deduce that the percentage of bad pineapples in the consignment almost certainly lies between 8.5 and 17.5

Sol^n: $SE(p) = \sqrt{\frac{p(1-p)}{n}}$

$$p = \frac{65}{500} = 0.13 \Rightarrow 1-p = 0.87$$

$$SE = \sqrt{\frac{0.13 \times 0.87}{500}} = 0.015$$

8/4

Absent

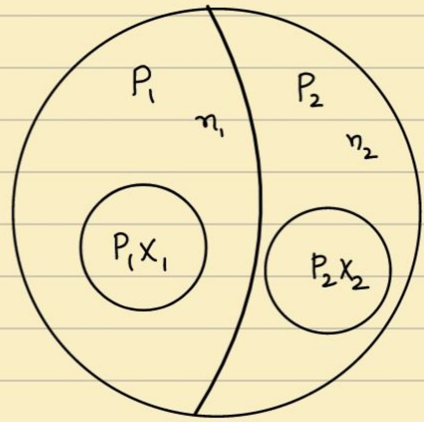
9/4

$$E[P_1] = P_1$$
$$E[P_2] = P_2$$

$H_0 : P_1 = P_2 = P$

$H_1 : P_1 \neq P_2$; $\varphi = 1 - P$

$$P_1 = \frac{X_1}{n_1}$$



$$P_2 = \frac{X_2}{n_2}$$

$$Z = \frac{P_1 - P_2}{\sqrt{\hat{P}\hat{\varphi}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

Take $\hat{p} = \dfrac{X_1 + X_2}{n_1 + n_2}$

$$\Rightarrow E[\hat{p}] = E\left[\frac{X_1 + X_2}{n_1 + n_2}\right]$$
$$= \frac{1}{n_1 + n_2} \cdot E[X_1 + X_2]$$
$$= \frac{1}{n_1 + n_2} \cdot \left(E[X_1] + E[X_2]\right)$$
$$= \frac{1}{n_1 + n_2} \cdot \left(E[n_1 P_1] + E[n_2 P_2]\right)$$

NULL Hypothesis :	Alternate Hypothesis :
$H_0 : P_1 = P_2$	$H_1 : P_1 = P_2$

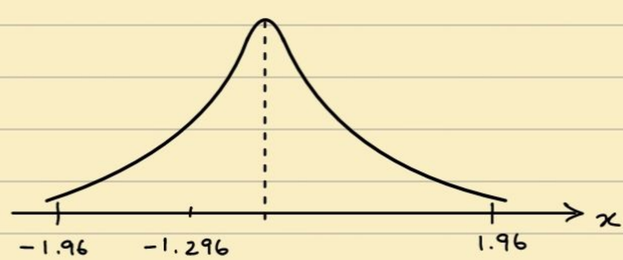
Fix , $Z_\alpha = 1.96$

$$Z = \frac{P_1 - P_2}{\sqrt{\hat{P} \cdot \hat{\varphi}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$\hat{P} = \frac{n_1 P_1 + n_2 P_2}{n_1 + n_2} = \frac{X_1 + X_2}{n_1 + n_2} = \frac{200 + 325}{400 + 600} = 0.525$$

$$\hat{\varphi} = 1 - \hat{P} = 0.475$$

$$Z = \frac{\frac{200}{400} - \frac{325}{600}}{\sqrt{0.525 \times 0.475 \times \left(\frac{1}{400} + \frac{1}{600}\right)}}$$
$$= -1.296$$



$\therefore$ H_0 accepted ($\because -1.96 < -1.296 < 1.96$)

Ex. A company has the head office at Calcutta and branch at Bombay. The personal director wanted to know if the members of the new place would like the introduction of a new place of work and survey was conducted for this purpose. Out of a sample of 500 workers at Calcutta, 62% favoured the new plan.
At Bombay, out of 400 workers, 41% are against the new plan. Is there a significance difference between the two groups towards the new plan

Sol: Here,

$n_1 = 500 \qquad n_2 = 400$

$P_1 = 0.62 \qquad P_2 = 0.59$

Null Hypothesis : $H_0 \Rightarrow P_1 = P_2$

Alternate Hypothesis : $H_1 \Rightarrow P_1 \neq P_2$

For $z_\alpha = 1.96$

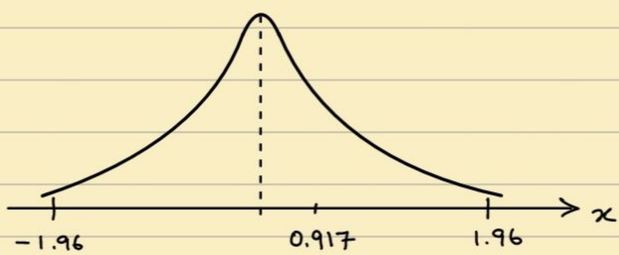
$z = ?$

$X_1 = n_1 P_1 = 500 \times 0.62$ | $X_2 = n_2 P_2 = 400 \times 0.59$

$$Z = \frac{P_1 - P_2}{\sqrt{\hat{P}\hat{Q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$= \frac{0.62 - 0.59}{\sqrt{0.607 \times 0.393 \times \left(\frac{1}{500} + \frac{1}{400}\right)}}$$

$$= 0.917$$



As $-1.96 < 0.917 < 1.96$,

$H_0 \rightarrow$ Accepted

& $H_1 \rightarrow$ Rejected

16/4

Q) In an year, there are 956 births in a town A of which 52.5% were males, while in towns A and B combined, this proportion in a total of 1406 births was 0.496. Is there any significant difference in the proportions of male births in the two towns?

Sol^n:

Null Hypothesis : $H_0 \Rightarrow P_1 = P_2$

Alternate Hypothesis : $H_1 \Rightarrow P_1 \neq P_2$

Fix $z_\alpha = 1.96$ (5%)

$$z = \frac{P_1 - P_2}{\sqrt{\hat{p} \cdot \hat{q} \cdot \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$n_1 = 956$

$n_1 + n_2 = 1406$

$\Rightarrow n_2 = 1406 - n_1$

$= 1406 - 956$

$\therefore n_2 = 450$

P_1 : Proportion of males in the sample town A

$= 0.525$

$$\hat{p} = \frac{X_1 + X_2}{n_1 + n_2} = 0.496$$

$$\Rightarrow \frac{n_1 P_1 + n_2 P_2}{n_1 + n_2} = 0.496$$

$$\Rightarrow \frac{956 \times 0.525 + 450 \times P_2}{1406} = 0.496$$

$\Rightarrow P_2 = 0.434$

$\hat{p} = 0.496$

$\hat{q} = 0.504$

$$z = \frac{0.525 - 0.434}{\sqrt{(0.496)(0.504)\left(\frac{1}{956} + \frac{1}{450}\right)}}$$

$= 3.368$

> 3

H_0 : false

H_1 : Accepted

Ex.

In two large populations, there are 30% and 25% respectively of blue-eyed people. Is this diff. likely to be hidden in samples of 1200 and 900 respectively from the two populations?

Sol^n: Null Hypothesis: $H_0 : p_1 = p_2$ | Alternate Hypothesis: $H_1 : p_1 \neq p_2$

Fix $z_\alpha = 1.96$

$$Z = \frac{P_1 - P_2}{\sqrt{\hat{P}\cdot\hat{Q}\cdot\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$$

$$\hat{P} = \frac{X_1 + X_2}{n_1 + n_2}$$

$X_1 \to n_1 P_1$, $X_2 \to n_2 P_2$

$$= \frac{0.25 - 0.30}{\sqrt{0.27 \times 0.73 \times \left(\frac{1}{1200} + \frac{1}{900}\right)}}$$

$$= -2.25$$

$$\hat{P} = \frac{\frac{30}{100} \times 1200 + \frac{25}{100} \times 900}{2100}$$

$$= \frac{360 + 225}{2100}$$

$$= \frac{585}{2100}$$

$$\Rightarrow \hat{P} = 0.27$$

$$\therefore \hat{Q} = 0.73$$

H_0 : false

H_1 : Accepted

Q) In a random sample of 400 students of university teaching departments, it was found that 300 students failed in the examination. In another random sample of 500 students of the affiliated colleges. The number of failures in the same examination was found to be 300.

Find out whether the proportion of failures in the university teaching dept. is significantly $\geq$ proportion of failures in the university teaching department and affiliated colleges taken together.

17/4

- Sometimes, the sample size, n, is small, we cannot use normal test.
  However, for small samples, we have
  (1) t-test
  (2) F-test
  here, F: Fischer
  (3) Fischer's z-transformation :

- <u>Assumptions for t-test</u>:
  (1) The population is normal.
  (2) The sample is random.
  (3) The standard deviation of the population is known.

Ex. The mean weekly sales of soap bars in departmental stores was 146.3 bars per store. After an advertising campaigning, the mean weekly sales in 22 stores for a typical week increased to 153.7 and showed a standard deviation of 17.2. Was the advertising campaigning successful.

<u>Soln</u> : We try to use t-test

$n = 22$

$\bar{x} = 153.7$

$s = 17.2$

Null Hypothesis : $\mu = 146.3$

Alternate Hypothesis : $\mu > 146.3$

Fix t : $t_0 = 1.72$ (from table)

Test Statistic :

$$t = \frac{\bar{x} - \mu}{\sqrt{s^2/(n-1)}}$$

$$\Rightarrow t = \frac{153.7 - 146.3}{\sqrt{(17.2)^2/21}}$$

$$= 1.97$$

Conclusion : As $1.97 > 1.72$, We reject H_0 & accept H_1.

21/4

→ Example:

A random sample of 10 boys/girls had the following IQ's : 70, 120, 110, 101, 88, 83, 95, 98, 107, 100. Do these data support the assumption of a population mean IQ of 100.

x	$x-\bar{x} = x-97.2$	$(x-\bar{x})^2$
70	-27.2	739.84
120	22.8	519.84
110	12.8	163.84
101	3.8	14.44
88	-9.2	84.64
83	-14.2	201.64
95	2.2	4.84
98	0.8	0.64
107	9.8	96.04
100	2.8	7.84

$\bar{x} = \Sigma x / 10$
$= 972/10$
$= 97.2$

$\Sigma (x-\bar{x})^2 = 1833.60$

Null Hypothesis : H_0
$\mu = 100$

Alternate Hypothesis : H_1
$\mu \neq 100$

Test Statistic :

$$t = \frac{\bar{x} - \mu}{\sqrt{S^2/n}}$$

$$S^2 = \frac{\Sigma (x-\bar{x})^2}{n-1}$$

$$= \frac{1833.60}{9}$$

$$= 203.73$$

$$t = \frac{972 - 100}{\sqrt{\frac{203.73}{10}}} = \frac{2.8}{\sqrt{20.37}}$$

$$= \frac{2.8}{4.514}$$

$$= 0.62$$

As $0.62 < 2.26$

Conclusion : We accept H_0

22/4

Ex.

The height of 10 females of a given locality are found to be 70, 67, 62, 68, 61, 68, 70, 64, 66, 66 inches.

Is it reasonable to believe that the average height is > 64 inches ?

Ex. Below are given the gain in weights (in kgs) of pigs fed on two diets A and B.

Gain in weight :

Diet A : 25, 32, 30, 34, 24, 14, 32, 24, 30, 31, 35, 25

Diet B : 44, 34, 22, 10, 47, 31, 40, 30, 32, 35, 18, 21, 35, 29, 22

Test if the two diets differ significantly as regards their effect on increase in weights.

Solⁿ: Null Hypothesis : H_0

$\mu_x = \mu_y$

There is no significance difference between the mean increase in weights due to diets A and B.

Alternate Hypothesis : H_1

$\mu_x \neq \mu_y$

Test Statistic :

$$t = \frac{\bar{x} - \bar{y}}{\sqrt{s^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

Here, $s^2 = \frac{1}{n_1+n_2-2}\left(\Sigma(x-\bar{x})^2 + \Sigma(y-\bar{y})^2\right)$

$$\left.\begin{aligned}\bar{x} &= 25 \\ \bar{y} &= 30\end{aligned}\right\} \Rightarrow s^2 = \frac{1}{12+15-2}\left(\Sigma(x-\bar{x})^2 + \Sigma(y-\bar{y})^2\right)$$

$$= 71.6$$

$$\Rightarrow t = \frac{28 - 30}{\sqrt{71.6\left(\frac{1}{12} + \frac{1}{15}\right)}} = -0.609$$

As $-2.06 < -0.609 < 2.06$, We accept H_0

Ex.

Samples of two types of electric light bulbs were tested for length of life and the following data were obtained

	Type I	type II
Sample no.s	$n_1 = 8$	$n_2 = 7$
Sample means	$\bar{x} = 1234$ hrs	$\bar{x} = 1036$ hrs
Sample SDs	$s_1 = 36$ hrs	$s_2 = 40$ hrs

Is the difference in the means significant to warrant that type I is superior to type II regarding length of life ?