

Statistical Methods

**By. Reem
Nazzaal**



**Dr. Yousef
Mubarak**



Introduction to statistics

26.7.2022

* **Statistics** : collect, analyze, present and interpret the data.

* **Mathematics vs Statistics** :

Maths → The values are exact.

Statistics → May have many possible answers.

→ engineering work :

① In the laboratory work.

② In research.

* **Constant** has fixed value. } → The difference between variable & constant.

* **Variables** has any value

└─ Quantitative variable (measured numerically) < @ Discrete.
└─ Qualitative " (can't be a numerical value) < @ Continuous.

@ Discrete → (countable) [integer]

@ Continuous → any number.

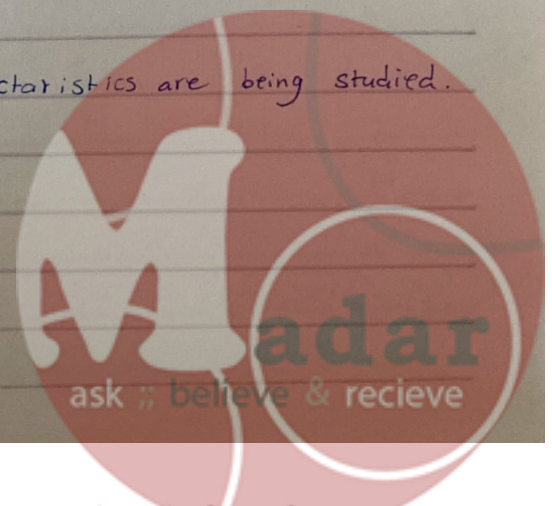
* **Types of statistics** :-

① **Descriptive** statistics → tables, graphs, summary. (چند و)

② **Inferential** statistics → consists of methods that use sample ~~that~~ results to help make decision about population

+ **population** : consists of all elements whose characteristics are being studied.

* **sample** : portion of the population.



* **Sample set** :- The set of all possible outcomes of a random experiment.

* **Sampling** :- ① random sample.

② stratified sample. (تقسيم عشوائي)

③ systematic sample. (تقسيم منتظم)

④ cluster sample. : we use it when we have large population.

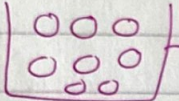
* **Survey** :- The collection of information from the elements of a population.

* **probability** :- measures of the likelihood of an event to occur.

* **experiment** :- outcomes and sample space.



* Sample space (Ω): all possible outcomes including ^{them} in Ω
 ↳ and you have to ... Ω outcome

* example:  population.

→ We want to see how many balls are defective and how many balls are good.

$\Omega = \left\{ \begin{array}{l} \text{defective} \\ \text{good} \end{array} \right\}$; $\Omega = \{(D, D), (D, G), (G, D), (G, G)\}$

* We can describe sample space by using two methods (1) Venn Diagram (2) Tree Diagram.
 Venn Diagram ↳ we have to include all options in one block

* Simple & compound events:-

(*) Simple event: an event that includes one and only one of the final outcomes for an experiment.

(*) Compound event: is a collection of more than one outcome for an experiment.

* Calculating probability:-

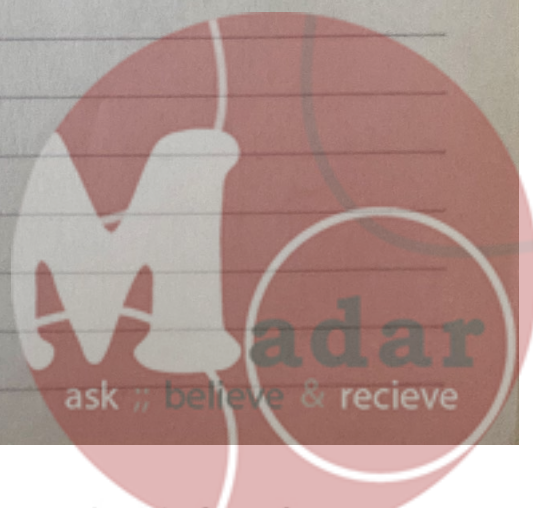
→ lies between 0 & 1

• impossible event M , $P(M) = 0$

• for a sure event C , $P(C) = 1$

The sum of the probabilities of all simple event for an experiment is always 1.

$$\sum_{i=1}^n P(E_i) = 1$$



* Three Types of probabilities :

① Classical probability : Two or more outcomes that have the same probability of occurrence .

$$P(E^c) = \frac{1}{\text{Total num of outcomes}}$$

$$P(A) = \frac{\text{num of outcomes for favorable to A}}{\text{Total num of outcomes}} \quad (\text{ingeneral})$$

example :- Tossing a coin $\begin{matrix} H \\ \swarrow \\ T \end{matrix}$

$$P(H) = 1/2$$

$$P(T) = \frac{1}{2}$$

$$\begin{aligned} \sum P(E_i) &= P(T) + P(H) \\ &= \underline{\underline{1}} \end{aligned}$$



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

$$P(1) = 1/6$$

$$P(4) = 1/6$$

$$P(2) = 1/6$$

$$P(5) = 1/6$$

$$P(3) = 1/6$$

$$P(6) = 1/6$$

$$\rightarrow P = \sum_{i=1}^6 P(\Sigma_i) = 1/6 + 1/6 + 1/6 + 1/6 + 1/6 + 1/6 = 1$$

* Find the probability of obtaining an ~~event~~ even num in one roll of the die.

$$\{2, 4, 6\} ; P(\text{even num}) = \frac{3}{6} = \frac{1}{2}$$

② Relative frequency concept of probability : an experiment repeated n times and an event $\{A\}$ is observed f times, according to the relative frequency concept.

$$P(A) = \frac{f}{n}$$

example 1) Suppose that you have to select 10 items from 500 and we need randomly to select these 10 cars and we need to find how many cars are considered as ~~defect~~ lemon.
* What will be the probability the next car manufactured will be lemon?

$$f = 10, n = 500$$

$$P(\text{next car is lemon}) = \frac{10}{500} = \frac{1}{50} = 0.02$$



③ Subjective probability : (personal). يعتمد على الدراء الشخصية.

* Counting rule : إذا بعد اي تجربة لعدد من المخرجات لكل خطوة معينة المنزى
عدد المخرجات لكل خطوة مع بعضى.

Step ①

{H, T}

2

$$2 \times 2 \times 2 = 8$$

Step ②

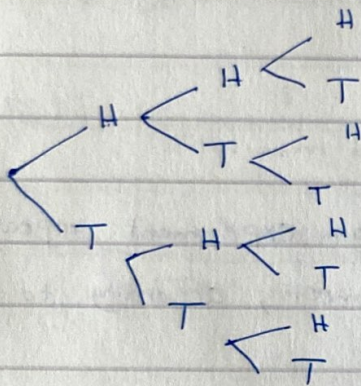
{H, T}

2

Step ③

{H, T}

2



out comes ..

HHH

THH

HHT

THT

HTH

TTH

HTT

TTT

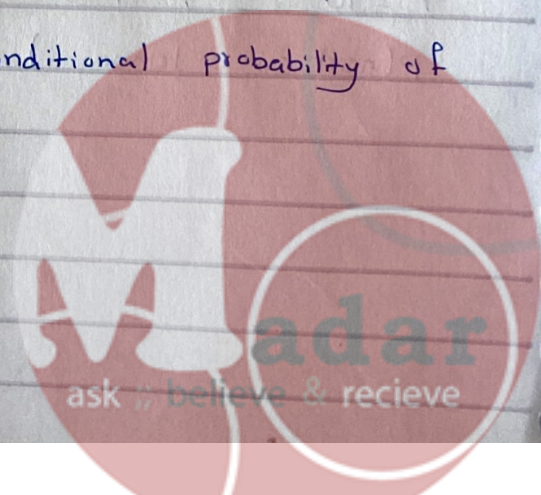
(8)

* Marginal and Conditional probability :

→ simple → does not consider the effect of any other event.

* Conditional probability : is the probability that an event will occur given that another event has already occurred.

→ If A and B are two events, then the conditional probability of A given B is written as P(A|B).



Example: an automobile manufactory provides speakles acupide with succeded option.

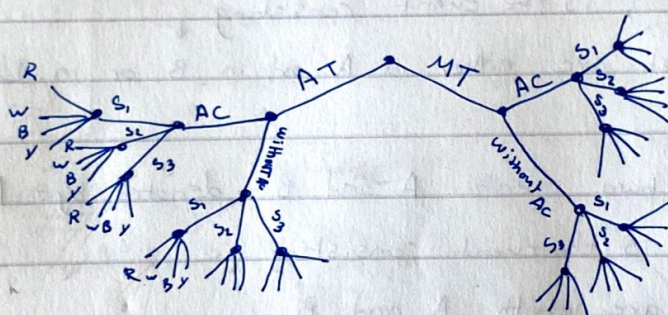
Car :-

- With or without automatic transmtion. 2
- with or without AC 2
- With one of three choices of a stereo system. 3
- with one of four exterial colors ; Red, white, Blue, yellow. 4

What is the num of outcomes?

Using counting rule = $2 \times 2 \times 3 \times 4 = 48$

If we represent this easly using tree diagram



48

$$P(AT, AC, S_1, R) = \frac{1}{48}$$

$$P(AT) = \frac{24}{48} = \frac{1}{2}$$

ترجع للسؤال | Consider an extension of the AT in which author vehicle option is the enternal color and suppose that the enternal cdors are ; Red Black, Blue, Brown but if the exterial color is red , only a black or red enternal is allowed

$$Q = 2 \times 2 \times 3 \times 4 \times 4 = 192 \longrightarrow$$

جدو ذ فيور

ask // believe & recieve

If the external color is white, any color is allowed internal
 " " " " " " Blue, only Black, Red or Blue
 " " " " " " yellow, a brown external can be used.

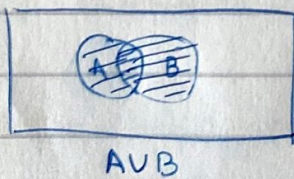
Total outcomes $\rightarrow S = \{120\}$

$$P(AT, Ac, \$, \text{white}, \text{Red}) = \frac{1}{120}$$

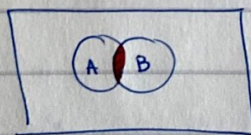
*** Some relations:-**

[1] The union of two events A and B, denoted by $A \cup B$ and read as A or B, is the event consisting of all outcomes that are either in A or in B or in both events

[2] The intersection of two events A and B denoted by $A \cap B$ and read as A and B, is the event consisting of all outcomes that are in both A and B

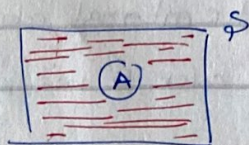


$A \cup B$



$A \cap B$

[3] The complement of an event A, denoted by A' (complement of A) is the set of all outcomes in S that are not contained in A.



$$A \cup A' = S$$

example :- $S = \{AA, AN, NA, NN\}$; IF $E_1 = \{AA, AN, NA\}$ and

$E_2 = \{AN, NN, NA\}$ then $E_1 \cup E_2$ & $E_1 \cap E_2$?

$$E_1 \cup E_2 = \{AA, AN, NA, NN\}$$

$$E_1 \cap E_2 = \{AN, NA\}$$

$$E_1' = \{NN\}$$

$$E_2' = \{AA\}$$

$$E_1 \cup E_1' = \{AA, AN, NA, NN\} = S$$

$$E_2 \cup E_2' = S$$

$$E_1 \cap E_1' = \emptyset \text{ null}$$

example :- PC sheets (49)
 scratch resistance
 shock resistance

shock \ scratch	high	Low
high	40	4
low	2	3

A = high shock resistance

B = scratch =

+ what is the num of sample $A \cap B$?

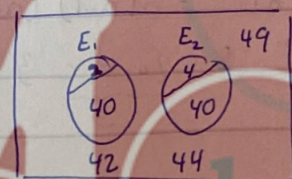
$$\boxed{40}$$

what is the num of sample A' ? $49 - 40 = 9$

what is the num of sample $A \cup B$? 46

* useful results :-

$$(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$$



$$40 + 2 + 4 = 46$$

ask :: believe & recieve

→ permutation & combinations :-

31-8

permutation: is an arrangement of an object in a definite order

combinations: is a selection of an objects without regard to order

outcomes with arrangement = {abc, acb, bac, bca, cab, cba}

$$p(cba) = 1/6$$

The num of permutation (arrangements) of n different elements is $n!$

$$n! = n(n-1)(n-2) \dots \times 1, 0! = 1$$

$$S = \{a, b, c\} \quad 3!$$

3

$$3 \times 2 \times 1 = \underline{6}$$



31-8

* If we need r elements from total of n different element by the permutation technique (arrangement) then P_n^r where n is the element in sample space or r is the number of objects to be withdrawn.

$n-(r-1)$ ← smallest

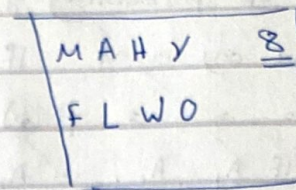
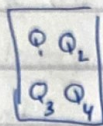
$n-0$ $n-1$ $n-2$ —
1st 2nd 3rd

$$n-r+1 = \frac{n!}{(n-r)!}$$



example:

تأليف من الكتب وكتاب واحد
من قبل يصنع سؤال واحد



*Arrangement is very important,
so we talk about permutation.

exam

(r)

Sample of TA

(n)

Sample space

$${}_8P_4 = \frac{n!}{(n-r)!} = \frac{8!}{(8-4)!} = \frac{8!}{4!}$$

$${}_8P_4 = \frac{8 \times 7 \times 6 \times 5 \times 4!}{4!} = 1680$$

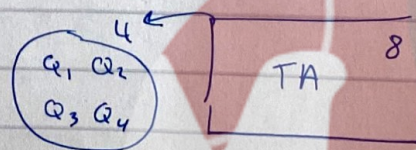
*Combination If we need to select r from a sample space from n elements without arrangement.

${}_nC_r$ or $\binom{n}{r}$ is given by

$${}_nC_r = \binom{n}{r} = \frac{n!}{r!(n-r)!} = \frac{{}_nP_r}{r!}$$

$${}_nC_r = \frac{{}_nP_r}{r!} \quad 4 \text{ teaching assistant from th } 8TA$$

$${}_8C_4 = \frac{{}_8P_4}{4!} = \frac{1680}{4!} = \frac{1680}{4 \times 3 \times 2 \times 1} = \frac{1680}{24} = 70$$

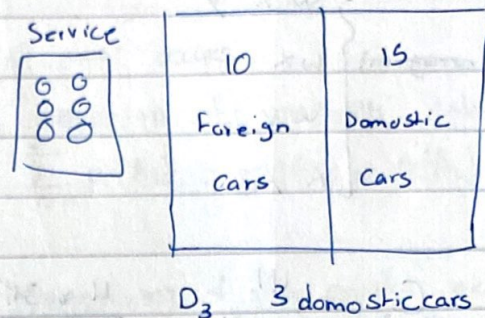


ask // believe & recieve

example :

6 to be selected randomly

② $P(3 \text{ one domestic} \mid 3 \text{ are foreign})$



$$P(D_3) = \frac{N(D_3)}{N} \rightarrow \frac{{}^{15}C_3 * {}^{10}C_3}{{}^{25}C_6}$$

$$N = {}^{25}C_6 = \frac{25!}{6! \cdot 19!} = 177100$$

$${}^{15}C_3 = \frac{15!}{3! \cdot 12!} = 455$$

$${}^{10}C_3 = \frac{10!}{3! \cdot 7!} = 120$$

$${}^{15}C_3 * {}^{10}C_3 = 455 * 120 = 54600$$

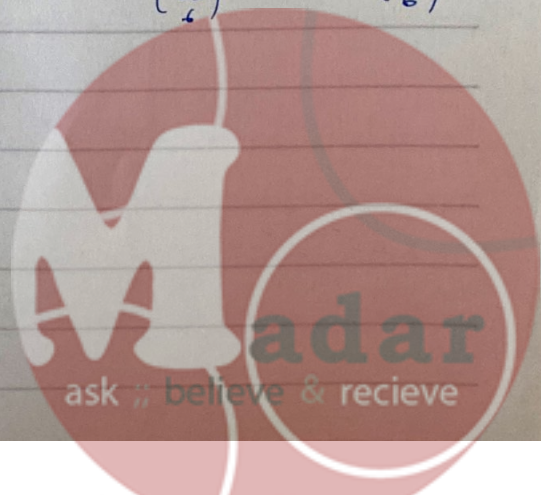
$$P(D_3) = \frac{54600}{177100} = 0.3083$$

$P(\text{at least 3 domestic cars will be selected})$

$$P(D_3) \geq 3 = P(D_3) + P(D_4) + P(D_5) + P(D_6)$$

$$P(D \geq 3) = \frac{54600}{{}^{25}C_6} + \frac{61425}{{}^{25}C_6} + \frac{30030}{{}^{25}C_6} + \frac{5005}{{}^{25}C_6}$$

$$P(D \geq 3) = 0.8530$$



example

outcome

no arrangement

- w x y
- w x z
- w y z
- x y z

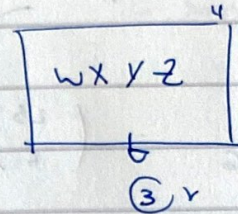
outcomes with arrangement.

wxy, wyx, xwy, xyw, ywx, yxw
wxz, wxz, xwz, xzw, zwx, zwx
wyz, wyz, zwy, zyw, yzw, ywz
xyz, xyz, zxy, zyx, yxz, yzx

$${}^4C_3 = \frac{4!}{3! \cdot 1!} = \frac{4 \times 3!}{3!} = 4$$

$${}^4P_3 = \frac{4!}{1!} = \frac{4 \times 3 \times 2 \times 1}{1} = 24$$

$$\frac{24}{4} = 6$$



example

If ${}_nP_r = 110$ & ${}_nC_r = 55$

$n = ?$

$$r = ? \rightarrow \frac{{}_nP_r}{r!} = {}_nC_r$$

$$r! = \frac{{}_nP_r}{{}_nC_r} = \frac{110}{55} = 2 \rightarrow 2! = 2 \times 1 = 2$$

$r = 2$

$${}_nP_r = {}_nP_2 = 110 = \frac{n!}{(n-2)!} = \frac{n(n-1)(n-2)!}{(n-2)!}$$

$$n^2 - n = 110 \rightarrow n(n-1) = 110$$

$n = 11$



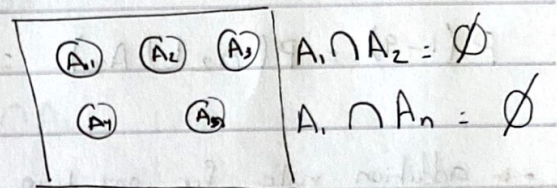
* Axioms of probability :

Axiom ① For any event if $p(A) \geq 0$

Axiom ② $p(S) = 1$, event S must occur.

Axiom ③ If $A_1, A_2, \dots, A_n$ is a finite collection of mutually exclusive event, when $p(A_1 \cup A_2 \cup A_3 \dots A_n) = \sum_{i=1}^n p(A_i) = p(A_1) + p(A_2) + \dots + p(A_n)$

* **exclusive event** $\rightarrow$ There is no intersection between any of the events exists in experiment.

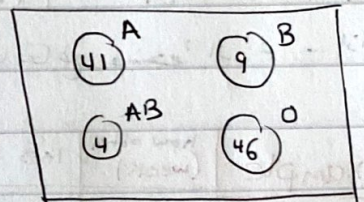


example | 41% type A 9% type B 4% type AB

46% type O

$$p(A \cup B \cup AB) = p(A) + p(B) + p(AB) \\ = 41 + 9 + 4 = 54\%$$

$$p'(A \cup B \cup AB) = p(O) \rightarrow 1 - 0.54 = 0.46$$



A, B, AB, O are mutually exclusive.

* **Addition rule for elementary rule** : The probability of an event A_i is equal to the sum of the probability of the elementary events forming A_i ;

$$A_i = (a_1, a_2, a_3) \text{ then } p(A_i) = p(a_1) + p(a_2) + p(a_3)$$

<< This rule is not valid for nonmutually exclusive events >>



example	how often	frequency	probability
	10 times or more	2100	0.42 A_1
	3 - 9	1500	0.3 A_2
	1 - 2	1200	0.24 A_3
	less than once	200	0.04 A_4
		5000	$\sum_{i=1}^4 P(A_i) = P(A_1) + P(A_2) + P(A_3) + P(A_4)$ = 1.0

$$P(1-9) = P(A_2 \cup A_3) = P(A_2) + P(A_3) = 0.3 + 0.24 = 0.54$$

→ addition rule for any two events ($A_1 + A_2$)

$$P(A_1 \text{ or } A_2) = P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$$

إذا كان في تقاطع بنحسبه مرة واحدة بس . إذا ما كان في تقاطع ونحسبه اثنى - إذا ما كان في تقاطع

example	how often (week)	less than 30 A_5	30 - 50 A_6	> 50 A_7	العمر
$A_1 \geq 10$		200	100	100	400
A_2 3-9		600	900	400	1900
A_3 1-2		400	600	500	1500
$A_4 < 1$		700	500	0	1200
		1900	2100	1000	5000

	A_5	A_6	A_7	
A_1	0.04	0.02	0.02	0.08
A_2	0.12	0.18	0.08	0.38
A_3	0.08	0.12	0.10	0.30
A_4	0.14	0.10	0	0.24
	$\Sigma = 0.38$	$\Sigma = 0.42$	0.20	$\Sigma P(A_5 + A_6 + A_7) = 1.0$

$$\Sigma P(A_1 + A_2 + A_3 + A_4) = 1.0$$

→ Suppose we need to know the probability of those who have their food out 10 or more per week:

$$P(A_1) = P(A_5) + P(A_6) + P(A_7) = 0.08$$

→ The probability of that the person is under 30:

هون بنشوند الاحتمالات لانه يكون اقل من 30 هي الشروط الـ 5 يعني انه يكون حقه شرط انه 30

$$P(A_5) = P(A_1/A_5) + P(A_2/A_5) + P(A_3/A_5) + P(A_4/A_5) = 0.38$$

$$\begin{aligned} \rightarrow P(A_4 \cup A_6) &= P(A_4) + P(A_6) - P(A_4 \cap A_6) \\ &= 0.24 + 0.42 - 0.1 \\ &= 0.56 \end{aligned}$$

$$\rightarrow P(A_4 \text{ or } A_7) \rightarrow P(A_4) + P(A_7) = 0.44$$



* Conditional probabilities : If an event (B) has occurred this will affect the probability of let's say event (A) $\neq 1$. (تأثير حدوث حدث على احتمال حدوث حدث آخر)

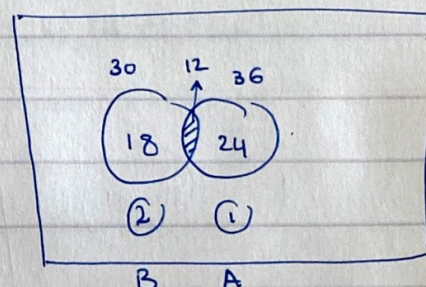
We will use the $P(A/B)$ to represent the conditional probability of A given that event B has occurred.

→ The conditional probability of an event (A) given a event (B) denoted as $P(A/B) = \frac{P(A \cap B)}{P(B)}$

$P(B) \neq \text{zero}$; A must intersect with B.

example | Suppose that there is a comet in 266 air samples, two rare molecules has to be detected, let A denote the event that the sample contains rare molecule (1) and let B denoted the event that consist all air samples in which rare molecule (2) is present.

molecule 1 \ molecule 2	NO	yes
NO	212	24
yes	18	12
		30
		36

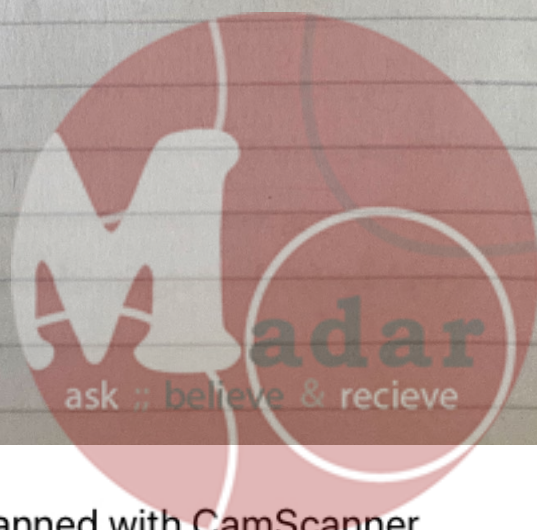


$P(A)$, $P(B)$, $P(A/B)$, $P(B/A)$?

$$P(A) = \frac{36}{266} \quad P(B) = \frac{30}{266}$$

$$P(A/B) = \frac{P(A \cap B)}{P(B)} = \frac{12/266}{30/266} = 0.402$$

$$P(B/A) = \frac{P(B \cap A)}{P(A)} = \frac{12/266}{36/266} = 0.33$$



num of trips \ sex	E_4 Female	E_5 male	
E_1 1 or 2	450	500	950
E_2 3-10	300	800	1100
E_3 over 10	100	350	450
	850	1650	
	2500		

$$P(E_2/E_4)$$

E_1

Traveller is already known

E_2

to be female and we need

E_3

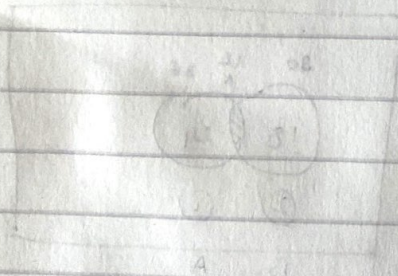
to determine the probability

that she will travel between 3-10 times.

$$P(E_2/E_4) = \frac{P(E_2 \cap E_4)}{P(E_4)} = \frac{300/2500}{850/2500} = 0.353$$

Example
do it your self

Suppose that the department receives 25 applications for a nomination that are ~~exist~~ ^{exist} for a opportunity. 10 are males, 15 are females. 17 hold Master's degree only and 8 have ~~high~~ high diploma. Suppose we are asked to find the probability that a female will be selected if it's known that the department will make it randomly knowing that the ~~person~~ person is having high diploma.



12	5
3	3

12	5
3	3

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{3/25}{5/25} = \frac{3}{5} = 0.6$$

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{3/25}{12/25} = \frac{3}{12} = 0.25$$

$$P(A) = \frac{12}{25} = 0.48$$

$$P(B) = \frac{5}{25} = 0.2$$

$$P(A \cap B) = \frac{3}{25} = 0.12$$



• independent events:

If $P(A/B) = P(A)$, the eq reduces to $P(A \cap B) = P(B) \cdot P(A)$ and we say that the events A & B are independent.

$$P(A/B) = \frac{P(A \cap B)}{P(B)} = P(A)$$

* إذا ما بائز علال probability لا يكون independent (A)

$$P(A/B) = P(A)$$

$$P(B/A) = P(B)$$

$$P(A \cap B) = P(A) \cdot P(B)$$

→ Can be applied on independent events

example | $P(\text{For any probability}) = \frac{1}{12}$; $S = \{(1,H), (1,T), (2,H), (2,T), (3,H), (3,T), (4,H), (4,T), (5,H), (5,T), (6,H), (6,T)\}$

A : the die shows one or two.

B : the coin shows head.

$A \cap B$: the die shows one or two and the coin shows head

$$P(A) = 4/12$$

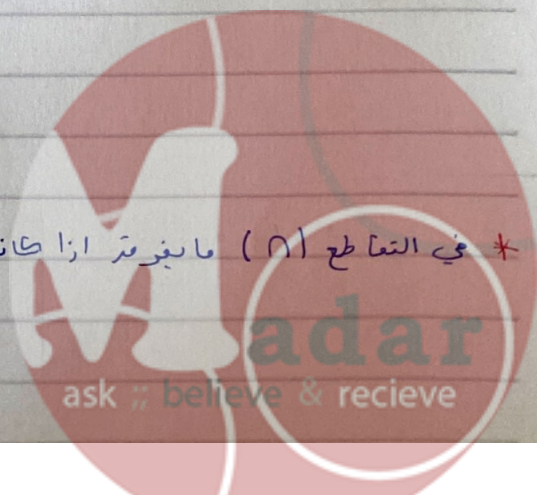
$$P(B) = 6/12$$

$$P(A \cap B) = \frac{2}{12}$$

$$P(A/B) = \frac{P(A \cap B)}{P(B)} = \frac{1/6}{1/2}$$

$$P(B/A) = \frac{P(B \cap A)}{P(A)} = \frac{1/6}{1/3}$$

* في التقاطع (A) ما يغيره اذا كانت $A \cap B$ أو $B \cap A$ لانهم متساوية (فمن بولنا).



$$P(A \cap B) = P(A) \cdot P(B)$$

$$1/6 = 1/6 \cdot (1) \quad (\checkmark)$$

example | consider the event A (water sample contains toxic level of lead) the probability is 0.32 and A_2 (water sample contains toxic level of Hg) is 0.16. the probability to have both exist is 0.1. suppose we're asked what is the probability that a randomly selected sample will contain toxic levels of Hg? what is the probability to have Hg knowing that it has lead?



* independent events :

* Two events are independent if, and only if, any one of the following statements is true :

$$(1) P(A|B) = P(A)$$

$$(2) P(B|A) = P(B)$$

$$(3) P(A \cap B) = P(A) P(B)$$

* The multiplication rule :

* For independent events E_1, E_2

$$P(E_1 \cap E_2) = P(E_1) * P(E_2)$$

(2) multiplication rule for two events (E_1, E_2)

$$P(E_2|E_1) = \frac{P(E_2 \cap E_1)}{P(E_1)}$$

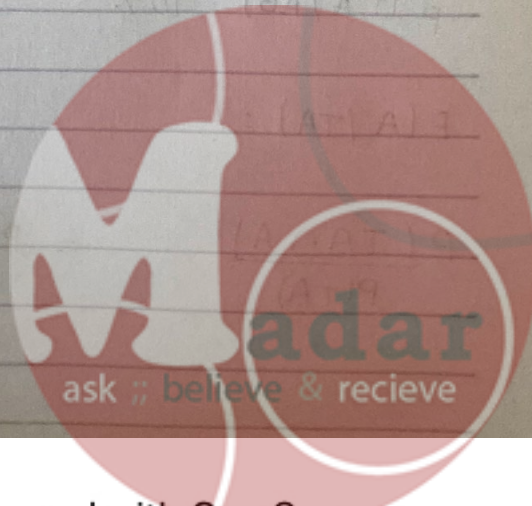
$$P(E_2 \cap E_1) = P(E_1 \cap E_2)$$

$$P(E_2|E_1) = \frac{P(E_1 \cap E_2)}{P(E_1)}$$

$$P(E_1 \cap E_2) = P(E_1) P(E_2|E_1)$$

$$P(E_2 \cap E_1) = P(E_2) P(E_1|E_2)$$

$$P(E_1|E_2) = \frac{P(E_2 \cap E_1)}{P(E_2)}$$



example | With draw 2 cards (without replacement)

if both cards are aces

Whats the probability to select two aces

52
cards

A_1 = ace on the first draw

A_2 = ace on the 2nd draw

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

$$P(A_1) = \frac{4}{52}$$

$$P(A_2) = \frac{3}{51}$$

$$P(A_2 | A_1) = \frac{3}{51} \quad \text{so}$$

example | ~~extended~~

A	B	AB	O
41%	9%	4%	46%

$$P(T.A/O) = 4\%$$

$$P(T.A/A) = 88\%$$

$$P(T.A/B) = 4\%$$

$$P(T.A/AB) = 10\%$$

$$P(A/TA) = ?$$

$$\frac{P(TA \cap A)}{P(TA)}$$

$$= P(A/TA)$$

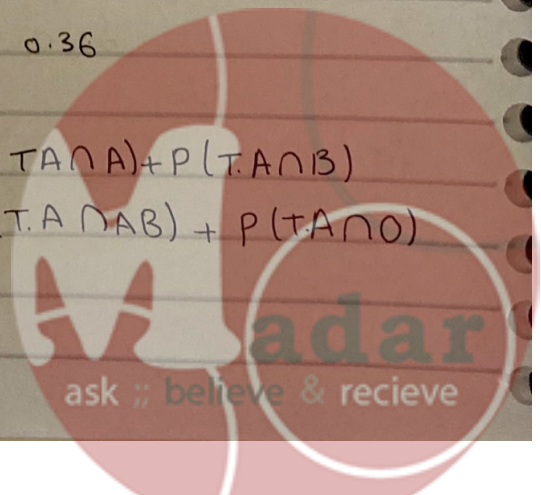
$$P(TA/A) \neq \frac{P(TA \cap A)}{P(A)}$$

$$P(TA \cap A) = P(A) \cdot P(TA/A)$$
$$= 0.41 \times 0.88$$

$$P(TA \cap A) = 0.36$$

$$P(TA) = P(TA \cap A) + P(TA \cap B)$$

$$+ P(TA \cap AB) + P(TA \cap O)$$



$$P(TA) = (0.4 \times 0.88) + (0.09 \times 0.04) + (0.04 \times 0.1) + (0.46 \times 0.04) \\ = 0.378$$

$$P(A/TA) = \frac{P(A \cap TA)}{P(TA)} = \frac{0.36}{0.378} = 0.95$$

* Bay's Theorem 8 Let $A_1, A_2, A_3, \dots, A_n$ be a collection of mutually exclusive events which partition S . Let B be an event such that $P(B) \neq \text{zero}$, Then for any of the events $A_i = i, 1, 2, 3, \dots, n$:

$$P(A_i/B) = \frac{P(A_i \cap B)}{P(B)} = \frac{P(B/A_i) \cdot P(A_i)}{\sum_{i=1}^n P(B/A_i) \cdot P(A_i)}$$

* This is the Bay's theorem 5

example | suppose that we have electric company doing calculators and these calculators manufactured in 3 different locations.

For.

$P(A/D)$? $P(B/D)$? $P(C/D)$?

A	B	C
50%	30%	20%
$P(D/A) = 5\%$	$P(D/B) = 7\%$	$P(D/C) = 12\%$

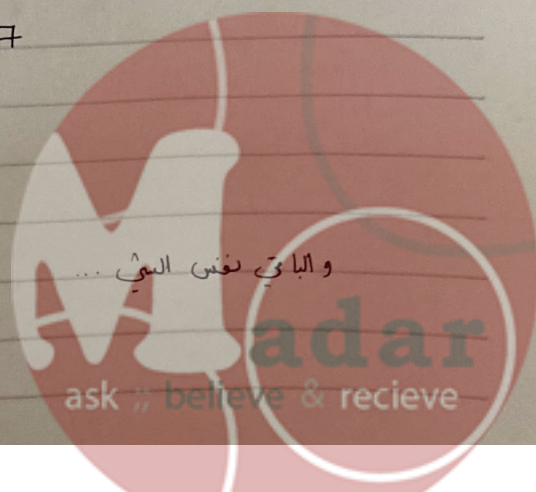
$$P(A/D) = \frac{P(A \cap D)}{P(D)}$$

$$P(A \cap D) = P(A) \cdot P(D/A) = 0.5 \times 0.05 = 0.025$$

$$P(D) = P(A \cap D) + P(B \cap D) + P(C \cap D)$$

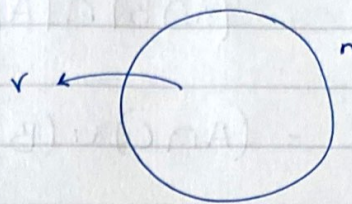
$$= (0.025) + (0.3 \times 0.07) + (0.2 \times 0.12) = 0.07$$

$$\therefore P(A/D) = \frac{0.025}{0.07} = 0.357$$



① permutation of objects that one n of all different

$${}_n P_r = \frac{n!}{(n-r)!}$$

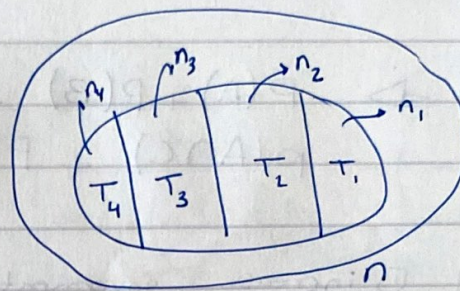


$${}_n C_r = \frac{n!}{r!(n-r)!} = \frac{{}_n P_r}{r!}$$

$${}_n P_{n_1, n_2, n_3, \dots}$$

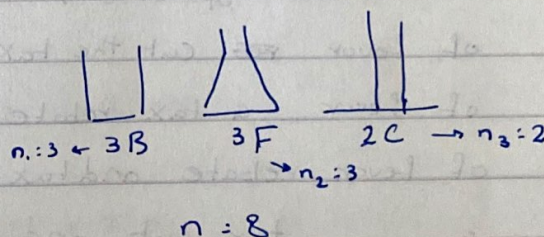
$$= \frac{n!}{n_1! \cdot n_2! \cdot n_3! \cdot \dots}$$

$$= \binom{n}{n_1} \binom{n-n_1}{n_2} \binom{n-n_1-n_2}{n_3} \dots$$



example

ڪيئن سڪن اڻڄاتل ڀرتي؟

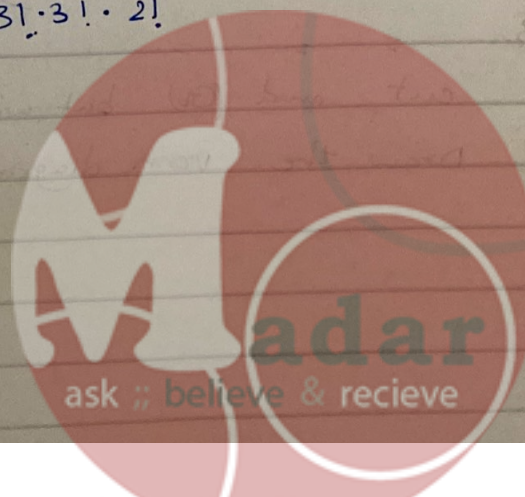


$${}_8 P_{332} = \frac{8!}{3! \cdot 3! \cdot 2!}$$

$$= 560$$

$${}_8 P_{332} = \binom{8}{3} \binom{5}{3} \binom{2}{2}$$

$$\frac{8!}{3! \cdot 3! \cdot 2!} \times \frac{5!}{3! \cdot 2!} \times \frac{2!}{2! \cdot 0!} = \frac{8!}{3! \cdot 3! \cdot 2!} = 560$$



* When we have three events A, B, C

$$1) A \cap B \cap C = (A \cap B) \cap (A \cap C)$$

$$2) (A \cup B) \cap C = (A \cap C) \cup (B \cap C)$$

$$3) P(A \cup B \cup C) = ? \rightarrow P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

بالأخير يتبقى $\rightarrow P(A) + P(B) + P(C) - P(B \cap C) - P(A \cap B) - P(A \cap C) + P(A \cap B \cap C)$

Example 1) increase governmental spending (G)

2) cut the tax (C)

3) ~~tax~~ rebate the tax (R)

50% of favor option (G)

45% of favor ~~and~~ cut the tax (C)

30% of favor a tax rebate (R)

10% of favor rebate and tax cut

25% " " tax cut and increase governmental spending

20% " " = rebate " " " "

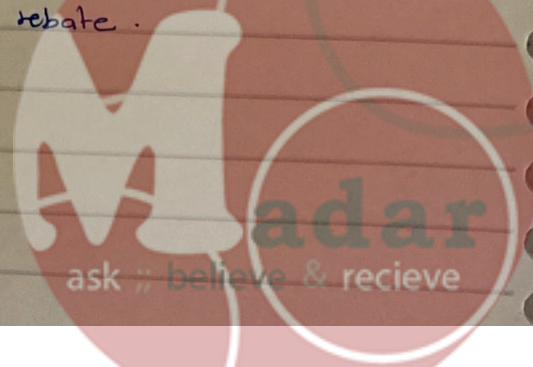
1% " " = " " = " " = " " = and cut the tax

① What is the probability that the candidate favor exactly one of the options?

② what is the probability that none of the options will be favored?

③ " " " " = the candidate will favor a tax cut and (G) but will oppose a tax rebate.

④ Draw the venn diagram.

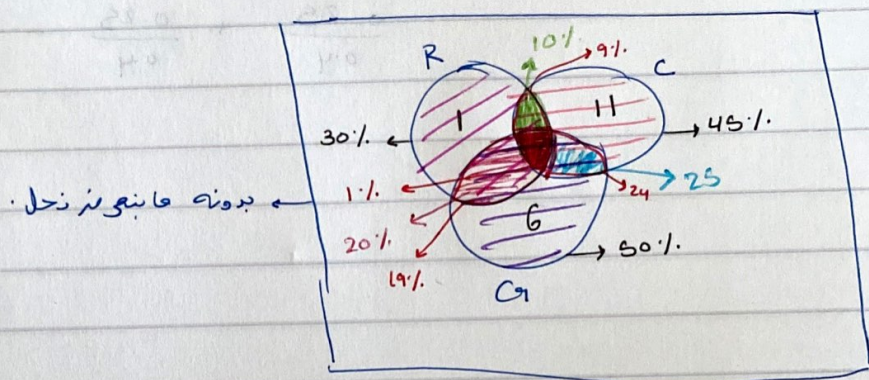


A) $p(\text{exactly 1 option}) = 1\% + 11\% + 6\% = 18\% = 0.18$

B) $p(G \cup R \cup C) = p(G) + p(R) + p(C) - p(G \cap R) - p(G \cap C) - p(R \cap C) + p(G \cap R \cap C)$
 $= 0.5 + 0.3 + 0.45 - 0.20 - 0.25 - 0.1 + 0.01 = 0.71$

$p'(G \cup R \cup C) = 1 - 0.71 = 0.29$

C) $p((C \cap G) \cap R') = 0.24$



* For events A, B and C when we have conditions

① $P((A \cup B) | C) = \frac{P((A \cup B) \cap C)}{P(C)} = P(A|C) + P(B|C) - P((A \cap B) | C)$

② $P(A \cap B \cap C) = p(A) * p(B|A) * p(C|(A \cap B))$

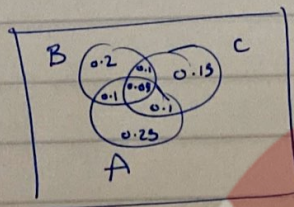
IF independent $P(A \cap B \cap C) = p(A) \cdot p(B) \cdot p(C)$

example

A = married

B = B.S.C

C = citizen



$$(1) P(A) = 0.25 + 0.1 + 0.1 + 0.05 = 0.5$$

$$(2) P(B \cap A) = 0.1 + 0.05 = 0.15$$

$$(3) P(C' \cap A \cap B) = 0.1$$

$$(4) P((A' \cup B') | C) = \frac{P(A' \cap C)}{P(C)} + \frac{P(B' \cap C)}{P(C)} - \frac{P(A' \cap B' \cap C)}{P(C)}$$

$$\frac{0.25}{0.4} + \frac{0.25}{0.4} - \frac{0.15}{0.4} = \underline{0.875}$$



example

arrests	num of people	$P(X=x) = \frac{\text{num}}{\text{total num}}$
0	300	0.6
1	100	0.2
2	60	0.12
3	20	0.04
4	10	0.02
5	5	0.01
6	5	0.01

range (0-1)

600

$\sum = 1.00$

random
Variable (real)

* Discrete probability Density : We need to $\rightarrow$ random variable
 $\rightarrow$ probability

* Two functions are used so

① Density function

② Cumulative distribution function

① We'll use the $f(x)$ for the density function of a discrete density. let X be a discrete random variable then the function f given by $f(x) = P(X=x)$, for x real is called the density function of X . $\rightarrow$ where x is the value in outcome.

② $f(x)$ is a probability, $f(x) \geq 0$ regardless the value of the random variable X .



③ If we sum f over all values of X that occur with non-zero probability, the sum must be 1.

$$f(x) \geq 0, \quad \sum_{\text{all } x} f(x) = 1 \quad \left. \begin{array}{l} \text{must always take place in random} \\ \text{discrete variable} \end{array} \right\}$$

$$p(x=1) = 0.2$$

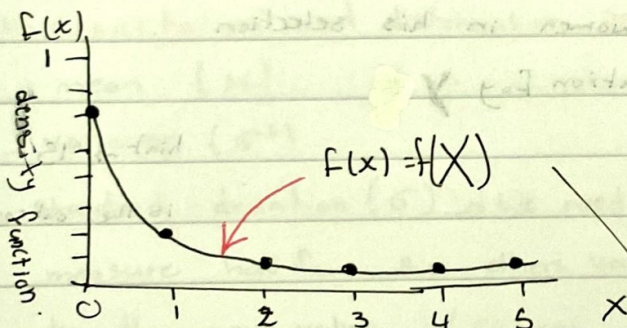
$$p(x=2) = 0.12$$

$$p(x \leq 2) = \sum_{x=0}^2 p(x_i) = p(x=0) + p(x=1) + p(x=2) = 0.92$$

$$p(x \geq 3) = p(x=3) + p(x=4) + p(x=5) + p(x=6) = 0.08$$

* Cumulative distribution 1-

$p(X \leq x)$ We have to accumulate up to that point
 accumulation the probability



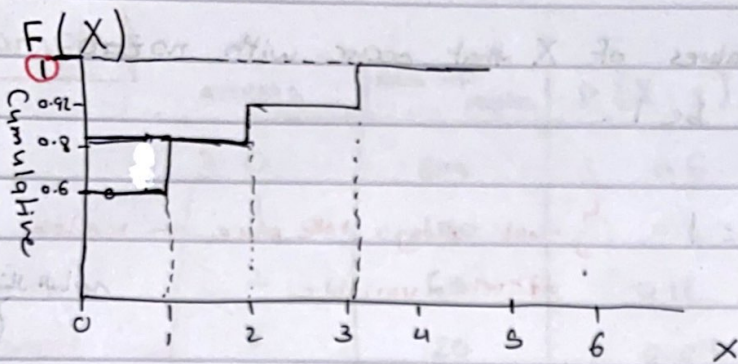
random variable

This represents the density function by this graphical representation.

$$p(X = x)$$

arrests	$f(x)$	Cumulative
0	0.6	0.6
1	0.2	$p(x \leq 1) = 0.8$
2	0.12	$p(x \leq 2) = 0.92$
3	0.04	$p(x \leq 3) = 0.96$
4	0.02	$p(x \leq 4) = 0.98$
5	0.01	$p(x \leq 5) = 0.99$
6	0.01	$p(x \leq 6) = 1.0$
$\sum f(x) = 1$		





random variable.

* This is cumulative distribution which is represent direct calculation for probability to let the variable less than or equal to (X)
 $P(X \leq x)$

example

We have 6 engineers 3 are male & 3 are female and they are working in a factor. The manager want to select 2 engineers to make a special job. He decided to select the 2 engineers randomly. let y = num of women in his selection.

find the probability distribution for y .

hint → 15

is the outcomes



Example 1 3 times tossing coin :

outcomes : { HHH, HHT, HTH, THH, TTH, THT, HTT, TTT }

let X denote the total number of heads obtained in the three tosses. Then X is a random variable whose possible values are 0, 1, 2, 3.

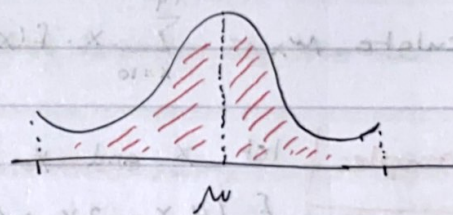
X	Total num of outcomes in sample space	$P(X=x)$	$P(X \leq x)$
0	1	$1/8$	$1/8$
1	3	$3/8$	$4/8$
2	3	$3/8$	$7/8$
3	1	$1/8$	$8/8$
	8	$\Sigma = 1$	

Expectation and distribution parameters :

mean (μ)

variance (σ^2)

standard deviation (σ) → We need to measure how far the others value compared to the mean value.



Mean (expected value) $E(X)$ (central value of the random variable)

let X be a discrete random variable with set of possible values D and probability mass density function $f(x)$. The expected value or the mean value of X , denoted by $E(X)$ or μ_x is $E(X) = \mu_x = \sum_{x \in D} x \cdot f(x)$.

* Rules of expectation:

let X and Y be random variable and let C is any real num,
then:

$$(1) E(c) = C \rightarrow E(c) = \sum_x c \cdot f(x) = c \sum f(x)$$

$$(2) E(cX) = c E(X)$$

$$(3) E(X+Y) = E(X) + E(Y)$$

$$(4) E(X^2) = \sum x^2 \cdot f(x)$$

example

X	10	11	12	13	14	15	16	17	18	19
f(x)	0.3	0.05	0.05	0.15	0.1	0.1	0	0.05	0.2	0

$$\text{Calculate } \mu_x = \sum_{x=10}^{19} x \cdot f(x) = (10 \times 0.3) + (11 \times 0.05) + (12 \times 0.05) + \dots = 13.45$$

example

let X and Y with $E(X) = 7$, $E(Y) = -5$, find the
 $E(4X - 2Y + 6)$

$$\begin{aligned} E(4X - 2Y + 6) &= E(4X) - E(2Y) + E(6) = 4E(X) - 2E(Y) + 6 \\ &= 28 + 10 + 6 = 44 \end{aligned}$$

* let X be a random variable has the following distribution

X	18	20	22
f(x)	0.2	0.6	0.2

$$E(X) = (18 \times 0.2) + (20 \times 0.6) + (22 \times 0.2) = 20$$

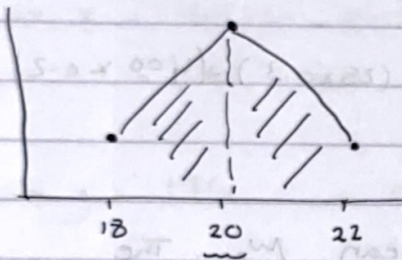


let y be a random variable

y	0	15	25	40
$f(y)$	0.2	0.3	0.3	0.2

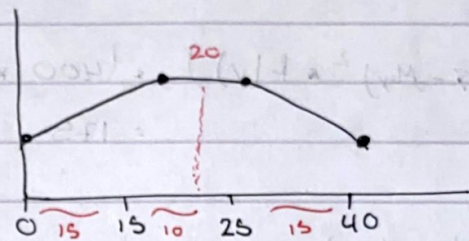
case 2

$$E(y) = (0 \times 0.2) + (15 \times 0.3) + (25 \times 0.3) + (40 \times 0.2) = 20$$



case 1

well distribution



case 2

no well distribution

mean value μ_x Scattering

Standard deviation σ and Variance σ^2 μ_y μ_x

in case 1 $\rightarrow X - \mu_x = 18 - 20 = -2$
 $22 - 20 = +2$

but

in case 2 $\rightarrow y - \mu_y = 0 - 20 = -20$
 $15 - 20 = -5$
 $25 - 20 = +5$
 $40 - 20 = +20$

$$\rightarrow (X - \mu_x)^2 = (X - 20)^2 = (-2)^2 + (2)^2 = 8$$

$$\rightarrow (y - \mu_y)^2 = (y - 20)^2 = (-20)^2 + (-5)^2 + (5)^2 + (20)^2 = 850$$



	22	20	22
$(x - \mu)^2$	4	0	4
$f(x)$	0.2	0.6	0.2

$$\sum_{i=1}^n ((x - \mu)^2 \cdot f(x)) = (4 \times 0.2) + (0 \times 0.6) + (4 \times 0.2)$$

$$= 1.6$$

$(y - \mu_y)^2$	400	25	25	400
$f(y)$	0.2	0.3	0.3	0.2

$$\sum_{i=1}^n ((y - \mu_y)^2 \cdot f(y)) = (400 \times 0.2) + (25 \times 0.3) + (25 \times 0.3) + (400 \times 0.2)$$

$$= 175$$

* let X be a random variable with mean μ_x . The variance (σ^2) of X , denoted by $\text{var } X$ or σ^2 , is given

$$\text{var}(X) = \sigma^2 = E[(X - \mu)^2] = \sum_x (x - \mu)^2 \cdot f(x)$$

$$\sigma^2 = E(x^2 - 2\mu x + \mu^2) = E(x^2) - 2\mu E(x) + \mu^2$$

$$\sigma^2 = E(x^2) - (E(x))^2$$

↓
give the same answer

$$E(x^2) = \sum x^2 \cdot f(x) - [E(x)]^2$$



expectation $E(x)$

$$E(x) = \sum_{\text{all } x} x_i \cdot f(x_i)$$

$$\begin{aligned} \text{Var } x &= \sigma^2 = E((x - \mu)^2) = \sum (x - \mu)^2 \cdot f(x) \\ &= E(x^2) - (E(x))^2 = \sum x^2 \cdot f(x) - (\sum x \cdot f(x))^2 \end{aligned}$$

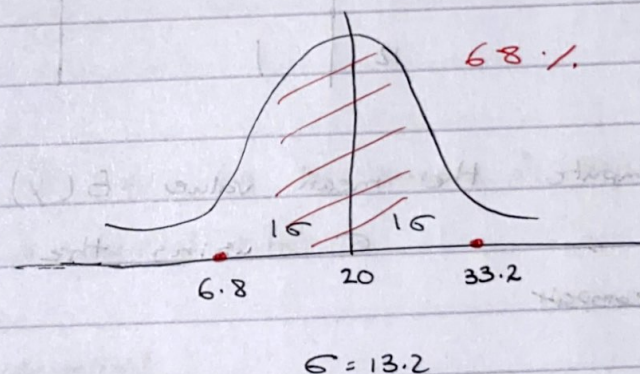
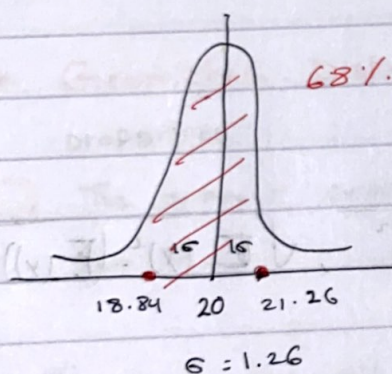
$$\begin{aligned} E(x^2) &= (18^2 \times 0.2) + (20^2 \times 0.6) + (22^2 \times 0.2) \\ &= 401.6 \end{aligned}$$

$$\begin{aligned} \text{Var } x &= 401.6 - 20^2 = 401.6 - 400 \\ &= 1.6 \quad \Rightarrow \sigma = \sqrt{\text{Var}(x)} = \sqrt{1.6} \end{aligned}$$

x	$f(x)$
18	0.2
20	0.6
22	0.2

$$E(x) = 20$$

$$\text{Var } y = 175 \quad \Rightarrow \sigma = \sqrt{\text{Var}(y)} = \sqrt{175}$$



example

The random variable y is the sum of the dice when a pair of fair dice is roll

y	out comes	$P(Y=y)$
2	1	$(1/6) \cdot (1/6)$
3	2	$(1/6) \cdot (1/6)$
4		$(1/6) \cdot (1/6) + (1/6) \cdot (1/6)$
5		$(1/6) \cdot (1/6) + (1/6) \cdot (1/6)$
6		$(1/6) \cdot (1/6) + (1/6) \cdot (1/6) + (1/6) \cdot (1/6)$
7		$(1/6) \cdot (1/6) + (1/6) \cdot (1/6) + (1/6) \cdot (1/6) + (1/6) \cdot (1/6)$
8		$(1/6) \cdot (1/6) + (1/6) \cdot (1/6) + (1/6) \cdot (1/6)$
9		$(1/6) \cdot (1/6) + (1/6) \cdot (1/6) + (1/6) \cdot (1/6)$
10		$(1/6) \cdot (1/6) + (1/6) \cdot (1/6)$
11		$(1/6) \cdot (1/6) + (1/6) \cdot (1/6)$
12	1	$(1/6) \cdot (1/6)$

① Compute the mean value $E(y)$

② using the $\sqrt{\sum (y - \mu_y)^2}$, $\sqrt{E(y^2) - (E(y))^2}$

③ compare



★ Rules For Variance

let X and Y be random variables and C any real number.

Then :

- ① $\text{Var}(C) = 0$
- ② $\text{Var}(C \cdot X) = C^2 \cdot \text{Var}(X)$
- ③ If X & Y are independent then $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$

example let X, Y be random variables

$$\sigma_x^2 = 9, \sigma_y^2 = 3 \text{ then}$$

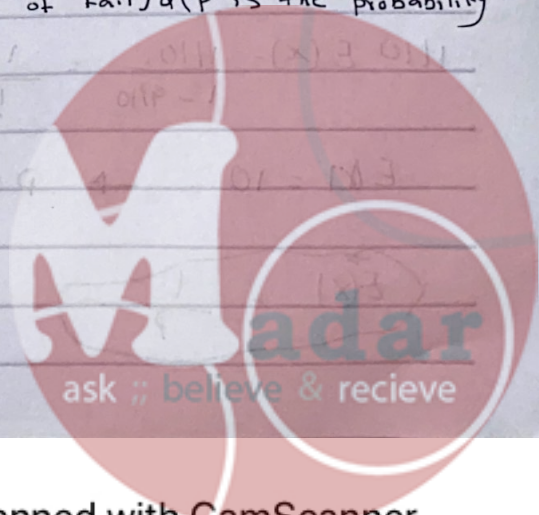
$$\begin{aligned} \text{Var}(4X - 2Y + 6) &= \text{Var}(4X) + \text{Var}(-2Y) + \text{Var}(6) \\ &= 16 \text{Var}(X) + 4 \text{Var}(Y) + 0 \\ &= (16 \times 9) + (4 \times 3) = \underline{156} \end{aligned}$$

★ Geometric Distribution:

properties :

- ① The experiment consists of series of trials and the outcome of each trial can be either a success or a fail
- ② The trials are all identical & independent
- ③ The random variable X denotes the number of trials needed to obtain the first success

★ If we say that $q = 1 - p$ (q is the probability of fail) & (p is the probability of success) then $\rightarrow P(X = y) = q^{y-1} \cdot p$



example | $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$

we need to select from this random space one at a time in such a way that at each stage in the selection process the integer chosen is just as likely to be one digit as any other.

Let X the number of trials to obtain the first zero

$$p = \frac{1}{10} \quad q = \frac{9}{10}$$

$$p(x=3) = \frac{9}{10} \times \frac{9}{10} \times \frac{1}{10}$$

$$p(x=6) = \left(\frac{9}{10}\right)^5 \times \frac{1}{10}$$

$$E(X) = \mu = \sum_{x=1}^{\infty} x \cdot f(x) = \sum_{x=1}^{\infty} x \cdot q^{x-1} \cdot p$$

$$= \sum_{x=1}^{\infty} x \cdot \left(\frac{9}{10}\right)^{x-1} \cdot \left(\frac{1}{10}\right) = \frac{1}{10} + \frac{18}{100} + \frac{243}{1000} + \frac{2916}{10000} + \dots \quad (1)$$

→ multiply by (q) → $\frac{9}{10}$

$$\frac{9}{10} E(X) = \frac{9}{100} + \frac{162}{1000} + \frac{2187}{10000} + \frac{26244}{100000} + \dots \quad (2)$$

→ subtract 2 from 1

$$\frac{1}{10} E(X) = \frac{1}{10} + \frac{9}{100} + \frac{81}{1000} + \frac{729}{10000} + \dots$$

No this series is geometric.

$$\frac{1}{10} E(X) = \frac{\frac{1}{10}}{1 - \frac{9}{10}} = \frac{\frac{1}{10}}{\frac{10-9}{10}} = \frac{1}{1}$$

$$E(X) = 10 \quad \rightarrow \quad p = \frac{1}{10}$$

$$E(X) = \frac{1}{p}$$



$$\text{Var} = \frac{q}{p^2} = \frac{1-p}{p^2}$$



* A random variable X has a binomial distribution with parameters (n) and p (probability of success) if its density $f(x)$ is given by $f(x) = \binom{n}{x} p^x (1-p)^{n-x}$

Example | exam contains 5 questions

$$p(1) = p(\text{True answer}) = 1/5 \leftarrow \text{num of question} = 5$$

$$p(1) = 4/5 = 0.8$$

$$p(x=3) = \binom{5}{3} (0.2)^3 (0.8)^2$$

$$= \frac{5!}{3! \cdot 2!} (0.2)^3 (0.8)^2 = 0.05$$

* expectation (E) and var:

The moment generation function for a binomial distribution

$$m_x(t) = (pe^t + q)^n$$

$$E(x) = \mu_x = 1^{\text{st}} \text{ Derivative for the } m_x(t) \Big|_{t=0}$$

$$E(x) = \frac{dm_x(t)}{dt} \Big|_{t=0} = n(pe^t + q)^{n-1} \cdot pe^t \Big|_{t=0}$$

$$E(x) = n(pe^0 + q)^{n-1} \cdot pe^0 = n(p+q)^{n-1} \cdot p$$

$$= n(p + (1-p))^{n-1} \cdot p = np$$

prove that $\text{Var} = npq$ and $\sigma = \sqrt{npq}$

example | 20% of all copies of a particular test fail, let X denote the num among 15 ~~copies~~ randomly selected copies that fail the test

P (the success of the random variable)

$$P(8) = 0.2 \quad P(f) = 0.8$$

find ① $P(X \leq 8) = P(0) + P(1) + P(2) + P(3) + P(4) + \dots + P(8)$

$$P(X=0) = \binom{15}{0} (0.2)^0 (0.8)^{15} = 0.0351843$$

$$P(X=1) = \binom{15}{1} (0.2)^1 (0.8)^{14} = 0.1319413$$

$$P(X=2) = \binom{15}{2} (0.2)^2 (0.8)^{13} = 0.2308974$$

$\vdots$

$$P(X=8) = \binom{15}{8} (0.2)^8 (0.8)^7 = 0.00345476$$

$$\therefore P(X \leq 8) = 0.0351843 + 0.1319413 + \dots$$

Cumulative $\rightarrow$ $= 0.9992144$

② $P(X=8) = \binom{15}{8} (0.2)^8 (0.8)^7 = 0.00345472$

$$P(X \leq 8) - P(X \leq 7)$$

$$0.9992 - 0.9958 = 0.00345$$

③ $P(X \geq 8) = 1 - P(X \leq 7)$

$$= 1 - 0.9958 = 0.00424$$

④ $P(4 \leq X \leq 7) = P(X \leq 3) + P(X \leq 7)$

inclusive $\rightarrow$ $= 0.9958 - 0.6482 = 0.3476$

$$E(X) = np$$

$$= 15 \times 0.2 = 3$$



Hypergeometric Distribution :

example

X is a random variable of the defective containers. What are the possibilities of random variable.

$$\text{outcomes} = \{CC, CN, NC, NN\}$$

$X=0 \quad X=1 \quad X=1 \quad X=2$

non-confusable random variable

$N = 850$	
$N-r$ 800 confirm	50 non-confirm

$$P(X=0) = \frac{800}{850} \times \frac{800}{850} = 0.8858$$

$$P(X=1) = \left(\frac{800}{850} \times \frac{50}{850} \right) + \left(\frac{50}{850} \times \frac{800}{850} \right) = 0.1107$$

$$P(X=2) = \left(\frac{50}{850} \times \frac{50}{850} \right) = 0.00346$$

$n = 2$ random variable

With replacement

$$P(X=0) = \frac{800}{850} \times \frac{799}{849} = 0.8857$$

$$P(X=1) = \left(\frac{800}{850} \times \frac{50}{849} \right) + \left(\frac{50}{850} \times \frac{800}{849} \right) = 0.110857$$

$$P(X=2) = \frac{50}{850} \times \frac{49}{849} = 0.003495$$

without replacement

→ In general, suppose that we have a group of N objects & that r of these objects have a trait of interests.

We are to select n object from the group randomly without replacement let X denote the number of objects objects chosen that have the trait Then:

$$P(X=x) = \frac{\binom{r}{x} \binom{N-r}{n-x}}{\binom{N}{n}}$$



$$\rightarrow N = 850$$

$$r = 50$$

$$n = 2$$

$$p(x=2) = \frac{\binom{50}{2} \binom{800}{0}}{\binom{850}{2}} = \frac{\left(\frac{50!}{2! 48!} \right) \left(\frac{800!}{0! 800!} \right)}{\left(\frac{850!}{2! 848!} \right)} = 0.00339$$

example

let x be the afternoon shift product container.

$$X = 0, 1, 2, 3, 4, 5$$

$$p(x=0) = \frac{8}{20} \times \frac{7}{19} \times \frac{6}{18} \times \frac{5}{17} \times \frac{4}{16}$$

$$= 0.0036119$$

N = 20	
12	8
afternoon	morning
✓	N=8 ✓

$$n = 5$$

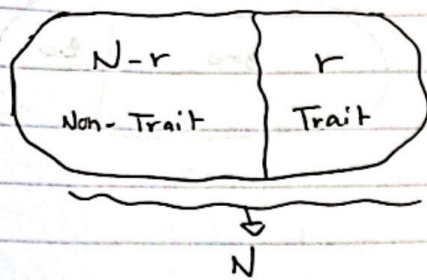
$$p(x=0) = \frac{\binom{r}{x} \binom{N-r}{n-x}}{\binom{N}{n}} = \frac{\binom{12}{0} \binom{8}{5}}{\binom{20}{5}} = \frac{\left(\frac{12!}{0! 12!} \right) \left(\frac{8!}{5! 3!} \right)}{\left(\frac{20!}{5! 15!} \right)}$$

$$p(x=0) = \frac{(1) \left(\frac{8 \times 7 \times 6 \times 5 \times 4}{5! \times 6} \right)}{\left(\frac{20 \times 19 \times 18 \times 17 \times 16 \times 15}{5 \times 4 \times 3 \times 2 \times 1 \times 5!} \right)} = \frac{56}{15504} = 0.0036119$$



16.8.2022

* in hypergeometric distribution:-



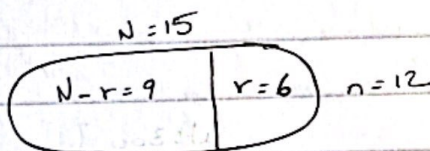
$n \rightarrow$ is the selected item - $\text{بدا نأخذ من العينة}$

$$P(X=x) = \frac{\binom{r}{x} \binom{N-r}{n-x}}{\binom{N}{n}}$$

* If a random variable X has a hypergeometric distribution with parameter N, n, r ; its density is given by:

$$f(x) = \frac{\binom{r}{x} \binom{N-r}{n-x}}{\binom{N}{n}}, \quad \max[0, n-(N-r)] \leq x \leq \min(n, r)$$

For example:-



$$\max \left(0, n - (N-r) \right) \leq x \leq \min(n, r)$$

$$3 \leq x \leq 6$$

② If the num of items sampled (n) is small relative to the number of objects from which the sample is drawn (N), then the binomial distribution can be used to approximate the hypergeometric probability.

if $\frac{n}{N} \leq 0.05$

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



Example 1

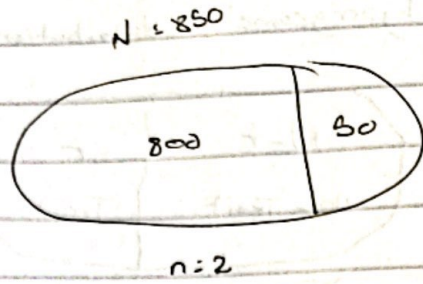
$$X = 0, 1, 2$$

$$P(X=0) = 0.88531$$

$$P(X=1) = 0.11073$$

$$P(X=2) = 0.00396$$

using
Hypergeometric



$$\frac{n}{N} \leq 0.05$$

$$\frac{2}{850} \leq 0.05 \quad \checkmark$$

So we can use binomial

So now using Binomial:-

$$P(X=0) = \binom{2}{0} (0.058)^0 (1 - 0.058)^2$$

$$= 0.8873$$

$$P(X=1) = \binom{2}{1} (0.058)^1 (1 - 0.058)^1$$

$$= 0.109$$

Example 1

$$P(X=4)$$

$$= \frac{\binom{100}{4} \binom{200}{0}}{\binom{300}{4}} = 0.011854$$

$N = 300$

L	AS
100	200
r	$N-r$

$n = 4$

$$\frac{n}{N} = \frac{4}{300} \leq 0.05 \quad (\checkmark)$$

$$(b) P(X \geq 2) = P(X=2) + P(X=3) + P(X=4)$$

$$= \frac{\binom{100}{2} \binom{200}{2}}{\binom{300}{4}} + \frac{\binom{100}{3} \binom{200}{1}}{\binom{300}{4}} + \frac{\binom{100}{4} \binom{200}{0}}{\binom{300}{4}}$$



(C) What is the probability that at least one part in the sample is from the last supplier?

$$P(X \geq 1) = P(X=1) + P(X=2) + P(X=3) + P(X=4)$$

OR

$$P(X \geq 1) = 1 - P(X=0)$$

(D) $E(X) = n \left(\frac{r}{N} \right)$, $Var(X) = n \left(\frac{r}{N} \right) \left(\frac{N-r}{N} \right) \left(\frac{N-n}{N-1} \right)$ ← rules

$$E(X) = 4 \left(\frac{100}{300} \right) = 1.333$$

$$Var = 4 \left(\frac{100}{300} \right) \left(\frac{200}{300} \right) \left(\frac{300-4}{299} \right) = 0.8799$$

* The poisson probability distribution:-

① Consider the number of patients arriving at the emergency ward during a 1 hour interval.

1. occurrence (of the arrival of patients) at the emergency.

2. interval is one hour

3. the total number of patients who may arrive may be 0, 1, 2, 3, 4, 5 ~

4. The patients arrive individually (independent).

* If the following conditions are satisfied to any experiment, then we can apply the poisson's probability distribution :-

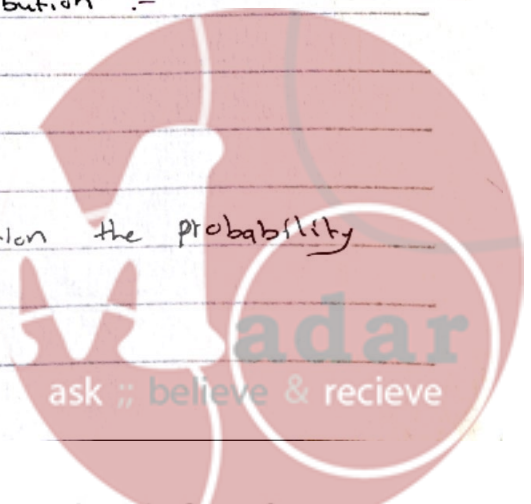
① X is a discrete random variable.

② The occurrence are random.

③ " " " independent

According to the poisson's probability distribution the probability of X occurrence in an interval.

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



→ Where (λ) is the mean number of occurrences in that interval ; and the value of (e) is approximate 2.71828

* The mean number of occurrences in an interval, denoted by (λ) is called the parameter of the poisson's probability distribution or the poisson's parameter.

* The intervals for λ and x must be equal , if they are not , the mean λ should be redefined to made them equal.

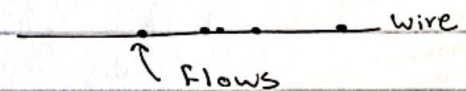
example

$$\lambda = 2.3 / \text{millimeter}$$

$$+ p(x=1) = 2.3$$

$$* p(x=2) = \frac{\lambda^x \cdot e^{-\lambda}}{x!}$$

2.3 millimeter $\times 1$



$$= \frac{(2.3)^2 \cdot e^{-2.3}}{2!} = 0.2651$$

$$* p(x=2 | 3 \text{ millimeter}) = \left(\lambda \times 3 \rightarrow 2.3 \times 3 = 6.9 \right) \text{ interval}$$

$$\rightarrow \frac{(6.9)^2 \cdot e^{-6.9}}{2!} = 0.02399$$



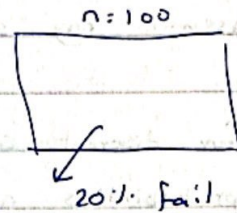
↓ Poisson's Distribution 8

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

Example | 15% students fail

(a) $\lambda = 15$

$$P(X=20) = \frac{(15)^{20} \cdot e^{-15}}{20!} = 0.04181$$



(b) $P(X \leq 9)$ (at most)

$$= P(X=0) + P(X=1) + \dots + P(X=9)$$

$$= 0.0699 \text{ (from the table) } \rightarrow \text{much easier.}$$

(c) $P(10 \leq X \leq 16) = P(X=10) + \dots + P(X=16)$

or $= P(X \leq 16) - P(X \leq 9)$

$$= 0.6641 - 0.0699 = 0.5942$$

(d) at least 20; $P(X \geq 20)$

$$= P(X=20) + P(X=21) + \dots + P(X=100)$$

or simply $= 1 - P(X \leq 19)$

$$1 - 0.8752 = 0.1248$$

using the Binomial

$$P(X=20) = \binom{100}{20} \left(\frac{0.15}{S}\right)^{20} \left(\frac{0.85}{F}\right)^{80} = 0.0402$$

$$n \geq 20$$

$$p \leq 0.15$$

$$n \geq 100$$

$$np \leq 10$$

(X)



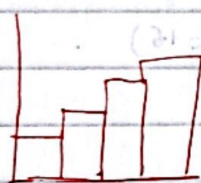
* Continuous Distribution:

a random variable (x) is said to be continuous if its set of possible values is an entire interval of numbers. If for some $A < B$ any number x between A & B and this will include infinity numbers. $\{A \text{ --- } B\}$. The probability that it assume any specific value is zero $p(X=23) = \frac{1}{\infty} = \text{Zero}$.

→ for discrete case $f(x)$ density, which enabled us to compute probabilities associated with the random variable x .

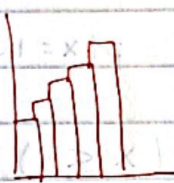
$$f(x) = p(X=x), \quad x \text{ is red.}$$

→ This definition can't be used in the continuous case because $p(X=x)$ is always zero.

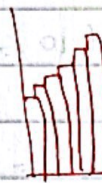


n

less accurat.



n



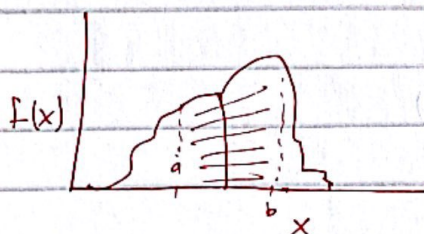
have a huge num. of n

more accurat.

* let x be a continuous random variable. Then the probability distribution or probability density function (of x) is a function $f(x)$ such that for any two numbers a and b with $a < b$.

$$P(a \leq x \leq b) = \int_a^b f(x) dx$$

$$p(X=a) = 0, \quad p(X=b) = 0$$



* That is the probability that x takes on a value in the interval $[a, b]$ is the area under the graph of the density function.

$$p(a \leq x \leq b) = \text{area under the graph}$$

fraction $0 \leq p \leq 1$

→ In the discrete case $\sum_{\text{all } x} f(x) = 1$

→ = continuous $\Rightarrow \int_{-\infty}^{+\infty} f(x) dx = 1$

$-\infty \rightarrow$ اقل رقم

$+\infty \rightarrow$ اكبر رقم

* example | lead concentration range 0.1 to 0.5 gram / liter

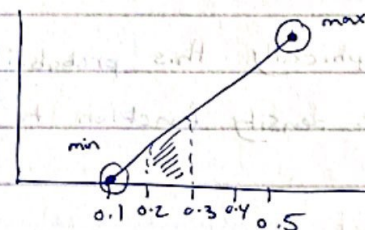
Random 1 liter gasoline

$$p(0.2 \leq x \leq 0.3)$$

$$f(x) = \begin{cases} 12.5x - 1.25 & ; 0.1 \leq x \leq 0.5 \\ 0 & ; \text{else where} \end{cases}$$

$$\int_{0.1}^{0.5} (12.5x - 1.25) dx$$

$$\left. \frac{12.5}{2} x^2 - 1.25x \right|_{0.1}^{0.5} = 1$$



$$p(0.2 \leq x \leq 0.3) = \left. \frac{12.5}{2} x^2 - 1.25x \right|_{0.2}^{0.3} = 0.1875$$



18.8.2022

* Continuous distribution:-

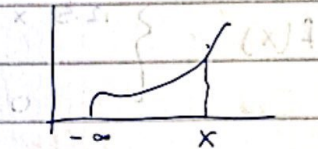
$$\int_{-\infty}^{+\infty} f(x) dx = 1, \quad p(X=x) = 0$$

* Continuous cumulative distribution:-

Let x be a continuous with density (f) The cumulative distribution function for x , denoted by (F) is given by:-

→ $F(x) = P(X \leq x)$, all real value of x . To find $F(x)$ for a specific number of x , we'll integrate the density over all real numbers that are less than or equal to x

$$P(X \leq x) = F(x) = \int_{-\infty}^{+\infty} f(x) dx$$



* Graphically this probability corresponds the area under the graph of the density function to the left of and including the point (x)

example

$$f(x) = 12.5x - 1.25 \quad ; \quad 0.1 \leq x \leq 0.5$$

$$\text{cumulative: } \int_{0.1}^x f(x) dx = \int_{0.1}^x (12.5x - 1.25) dx$$

$$= 6.25x^2 - 1.25x + 0.0625$$

$$P(0.2 \leq x \leq 0.3) = P(x \leq 0.3) - P(x \leq 0.2) \\ = F(0.3) - F(0.2)$$

$$F(0.3) = 6.25(0.3)^2 - 1.25(0.3) + 0.0625 = 0.25$$

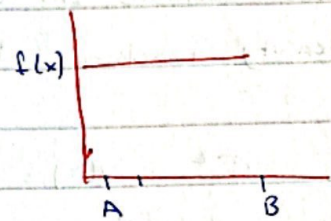


$$F(0.2) = 6.25(0.2)^2 - 1.25(0.2) + 0.0625 = 0.0625$$

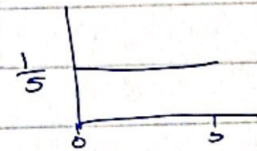
$$P(0.2 \leq x \leq 0.3) = 0.25 - 0.0625 = 0.1875$$

→ A continuous random variable X is said to have a uniform distribution on the interval $[A, B]$, if the probability density function is constant:-

$$f(x) = \begin{cases} \frac{1}{B-A} & , A \leq x \leq B \\ 0 & , \text{other where} \end{cases}$$



example $f(x) = \begin{cases} \frac{1}{5} & , 0 \leq x \leq 5 \\ 0 & \end{cases}$



$$\int_0^5 \frac{1}{5} dx = \frac{1}{5}(5-0) = 1$$

$$P(1 \leq x \leq 3) = F(3) - F(1) = \int_1^3 f(x) dx$$

* Expectation and distribution parameters:-

→ expected value: let X be a continuous random variable, with density function f . let $H(x)$ be a random variable, Then, the expected value of $H(x)$ denoted by $E(H(x))$ is given by:-

$$E(H(x)) = \int_{-\infty}^{\infty} H(x) \cdot f(x) dx$$



example 1 $H(x) = 12.5x - 1.25$; $0.1 \leq x \leq 0.5$

$$\mu = \int_{0.1}^{0.5} x (12.5x - 1.25) dx$$

$$= 0.3667 \text{ g/l}$$

* The variable of a continuous random variable X with probability density function and mean value.

$$\sigma^2 = v(x) = \int_{-\infty}^{\infty} (x - \mu)^2 \cdot f(x) dx$$

$$= E(x - \mu)^2$$

$$= E(x^2) - x(E(x))^2$$

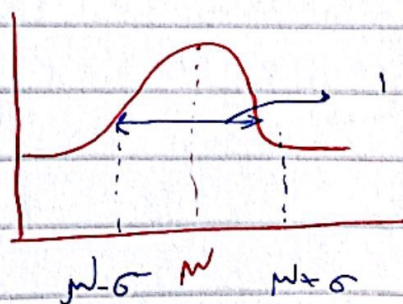
$$\text{ex} \rightarrow \text{var}(x) = 0.0029$$

* Normal distribution: area $\rightarrow$ above mean
area $\rightarrow$ below mean

$\rightarrow$ A continuous random variable (X) is said to have a normal distribution with parameters μ, σ , where

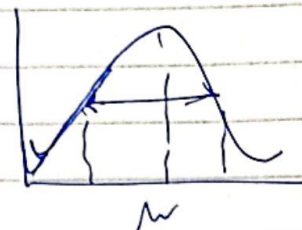
$-\infty \leq \mu \leq \infty$, $0 < \sigma$; If the probability density function of X is

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



$1\sigma \rightarrow 68\%$ from the total area but after 1σ distance

→ 2σ → area will be change
 → Total area = 95.4 %
 2σ



* $P(\mu - \sigma \leq x \leq \mu + \sigma) = 68.27\%$

* $P(\mu - 2\sigma \leq x \leq \mu + 2\sigma) = 95.45\%$

* $P(\mu - 3\sigma \leq x \leq \mu + 3\sigma) = 99.73\%$

* $P(\mu - 6\sigma \leq x \leq \mu + 3\sigma) = 99.9999\%$

* The graph is symmetric, bell-shaped, centered μ with inflection points $\mu \pm \sigma$

Tables → imp

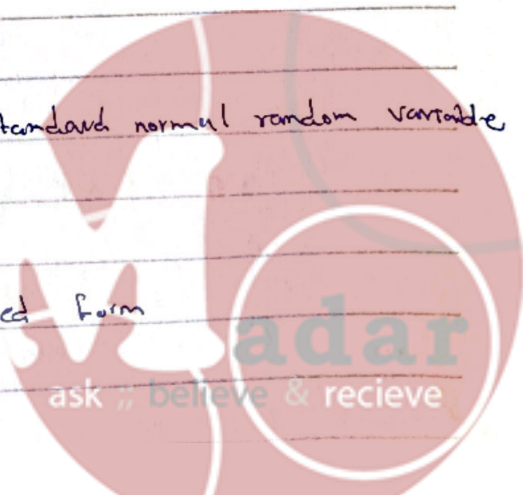
→ In a normal random variable with $\mu=0$ and $\sigma^2=1$ is called standard normal random variable.

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

$$P(z) = \frac{1}{\sqrt{2\pi}} e^{-1/2 * z^2}$$

* The cumulative distribution function for the standard normal random variable is given in Tables.

* The normal density is not integration closed form (use numerical integration)



Normal Distribution :-

 $\mu = 0, \sigma = 1 \rightarrow$ standard normal distribution

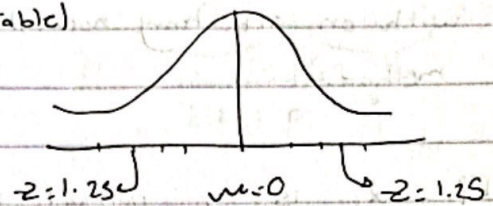
$$f(x) = (1/\sqrt{2\pi}) \cdot e^{(-1/2)x^2} \quad -\infty < x < +\infty$$

$$f(z) = (1/\sqrt{2\pi}) \cdot e^{(-1/2)z^2} \quad -\infty < x < +\infty$$

example (a) $p(z \leq 1.25) = 0.8944$ (Table)

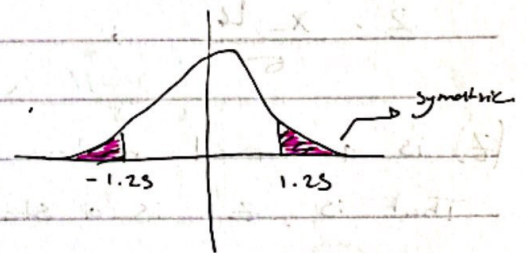
(b) $p(z > 1.25) = 1 - p(z \leq 1.25)$

$$= 1 - 0.8944 = 0.1056$$



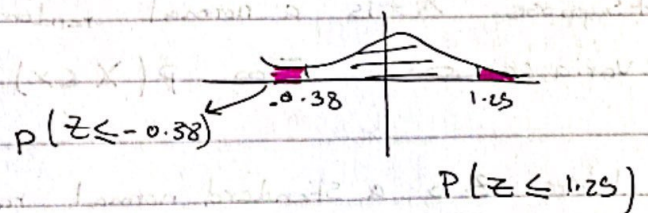
$$p(z \leq -1.25) = 0.1056 \quad (\text{Table})$$

(c) $p(z \leq -1.25) = 0.1056$



(d) $p(-0.38 \leq z \leq 1.25)$

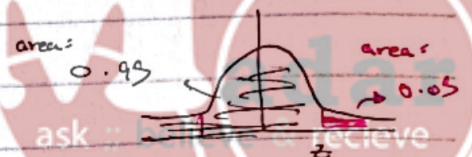
$$= p(z \leq 1.25) - p(z \leq -0.38)$$



$$= 0.8944 - 0.3520$$

$$= 0.5424$$

(e) Find the value z such that $p(z > z) = 0.05$
 $p(z \leq 1.65)$



* If $\mu \neq 0$ or $\sigma \neq 1$, then the normal standard function can't be used directly.

* To use the same approach for an arbitrary normal random variable would require a separate table for every possible pair of μ and σ .

* Fortunately, all normal probability distributions are related algebraically, and appendix can be used to find the probabilities associated with an arbitrary normal random variable using simple transformation method.

If X is a normal random variable with $E(X) = \mu$ and $\text{Var}(X) = \sigma^2$ then the random variable:

$$Z = \frac{X - \mu}{\sigma};$$

(Z) is a normal random variable with $E(Z) = 0$ and $\text{Var}(Z) = 1$.

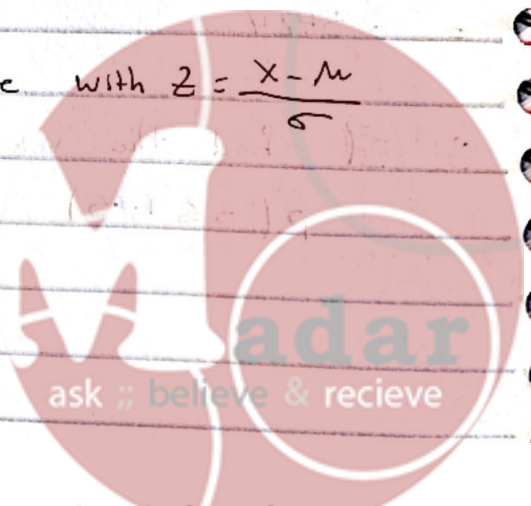
That is, Z , is a standard normal random variable.

* Suppose X is a normal random variable with mean μ and variance σ^2 , Then $P(X \leq x) = P\left(\frac{X - \mu}{\sigma} \leq \frac{x - \mu}{\sigma}\right) = P(Z \leq z)$

Where Z is a standard normal random variable.

$z = \frac{x - \mu}{\sigma}$ is the z -value obtained by standardization X .

The probability is obtained by entering table with $z = \frac{x - \mu}{\sigma}$



example

$$\mu = 10 \text{ mA}$$

$$\text{Var} = 4 \text{ (mA)}^2 \rightarrow \sigma = 2$$

Let x denote the current in mA

$$P(x > 13)$$

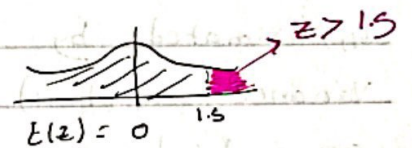
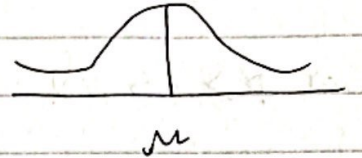
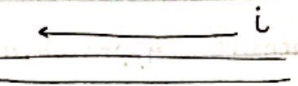
$$P(x > 13) = P\left(\frac{x-10}{2} > \frac{13-10}{2}\right)$$

$$= P(z > 1.5)$$

$$= 1 - P(z \leq 1.5)$$

$$\text{or } = P(z \leq -1.5)$$

$$= 0.0668$$



$$(b) P(9 < x < 11) = P\left(\frac{9-10}{2} < \frac{x-10}{2} < \frac{11-10}{2}\right)$$

$$= P(-0.5 < z < 0.5)$$



$$= \Phi(0.5) - \Phi(-0.5)$$

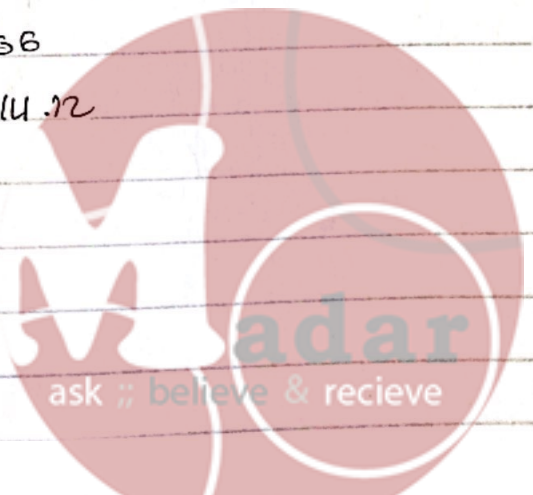
$$0.6915 - 0.3085 = 0.383$$

$$P(9 < x < 11)$$

(c) Find the value of x such that $P(z \leq z) = 0.98$

From the table $\Phi(z) = 0.98$, $z = 2.06$

$$z = \frac{x - \mu}{\sigma} \rightarrow 2.06 = \frac{x - 10}{2} = 14.12$$

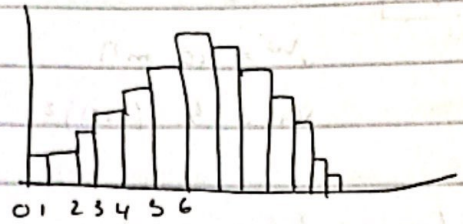


Normal approximation :-

for binomials $P(X=x) = \binom{n}{x} (p)^x (q)^{n-x}$

if $p \leq 0.5$ & $np > 5$

or $p > 0.5$ & $n(1-p) > 5.0$



just the previous part

* let X be a binomial with parameters n and p for large (n) , X is approximated by a normal distribution with $\mu = np$ and Variance $np(1-p)$ OR let X be a binomial random variable based on n trial with success probability p . then IF the binomial probability histogram is not too skewed, X has a probability that can be approximated by normal distribution with $\mu = np$ and $\sigma = \sqrt{npq}$.

$$\left(\frac{0.1 - 11}{3} < \frac{0.1 - X}{3} < \frac{0.1 - 1}{3} \right) q = (11 < X < 1) q (d)$$

$$(0.1 - 11 < 0.1 - X < 0.1 - 1) q =$$



$$P(X=0) = \binom{50}{0} (0.25)^0 (0.75)^{50} = 5.6632 \times 10^{-7}$$

:

$$P(X=10) = \binom{50}{10} (0.25)^{10} (0.75)^{40} = 0.0985184$$

$$P(X \leq 10) = \sum \square = 0.2642$$

$$p < 0.5, np > 5, \mu = np = 12.5, \sigma = 3.06$$

$$P(X \leq 10) = P\left(\frac{X - \mu}{\sigma} \leq \frac{10 + 0.5 - 12.5}{3.06}\right) \neq 0 \quad \neq 1$$

$$= \Phi(-0.65)$$

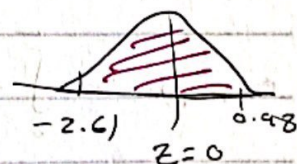
$$= 0.2578 \text{ using Table}$$

$$* P(5 \leq X \leq 15) = \Phi\left(\frac{15 + 0.5 - 12.5}{3.06}\right) - \Phi\left(\frac{5 - 0.5 - 12.5}{3.06}\right)$$

$$= \Phi(0.98) - \Phi(-2.61)$$

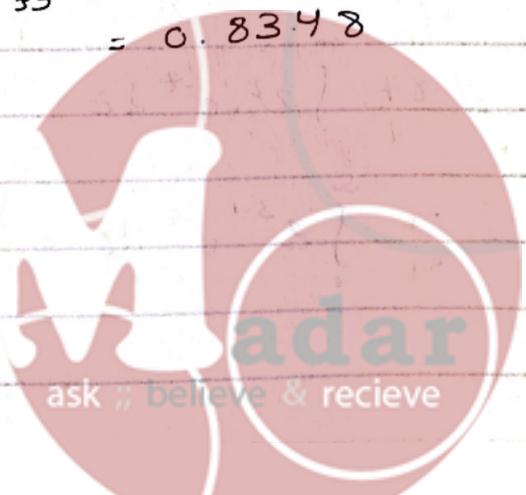
$$= 0.8365 - 0.0045$$

$$\approx 0.832$$



$$\text{using binomial} : \binom{50}{5} (0.25)^5 (0.75)^{45} + \dots + \dots +$$

$$\binom{50}{15} (0.25)^{15} (0.75)^{35} = 0.8348$$



Gamma Distribution:-

① Gamma

② expo

③

denoted by Γ anddefined as $\Gamma(x) =$

$$\int_0^{\infty} z^{x-1} \cdot e^{-z} dz$$

Where $x > 0$

properties:

① $\Gamma(1) = 1$

② for $x > 1$,

$$\Gamma(x) = (x-1) \Gamma(x-1)$$

example $\rightarrow \int_0^{\infty} z^3 \cdot e^{-z} dz = \Gamma(4)$

$$\Gamma(4) = 3 \Gamma(3) = 3 \cdot 2 \Gamma(2)$$

$$= 3 \cdot 2 \cdot 1 \Gamma(1)$$

$$\Gamma(4) = 6$$

example :- $\int_0^{\infty} \frac{1}{34} x^2 \cdot e^{-x/3} dx$

let $z = \frac{x}{3} \rightarrow 3z = x \rightarrow 3dz = dx$

$$\int_0^{\infty} \frac{1}{54} \cdot 9z^2 e^{-z} \cdot 3dz = \int_0^{\infty} \frac{27}{54} z^2 e^{-z}$$

$$= \frac{27}{54} \int_0^{\infty} z^2 \cdot e^{-z} dz$$

$$= \frac{27}{54} \int_0^{\infty} z^{3-1} e^{-z} dz = \frac{27}{54} \Gamma(3) = \frac{27}{54} \cdot 2 \Gamma(2)$$

$$= \frac{27}{54} \cdot 2 \cdot 1 \cdot \Gamma(1) = \frac{54}{54} = 1$$

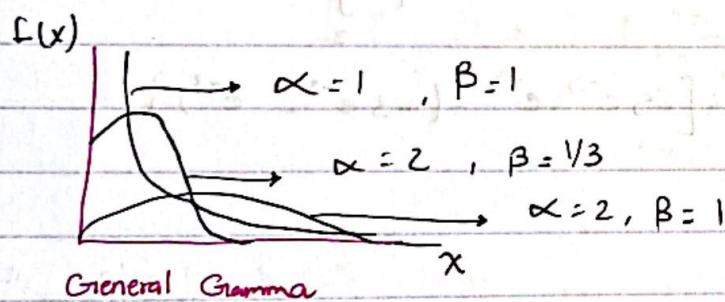


* A continuous random variable x is said to have a gamma distribution if the probability density function of x is defined as

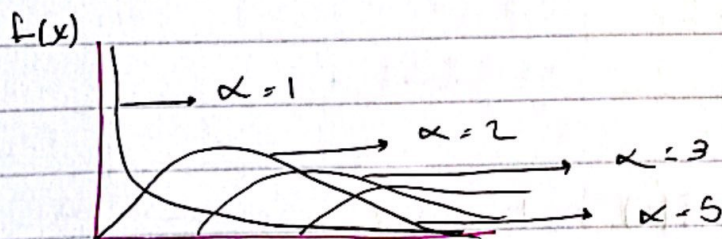
$$f(x) = \frac{1}{\Gamma(\alpha) \beta^\alpha} \cdot x^{\alpha-1} \cdot e^{-x/\beta}, \quad x \geq 0 \quad \left. \vphantom{\frac{1}{\Gamma(\alpha) \beta^\alpha}} \right\} \begin{array}{l} \text{General} \\ \text{Gamma} \end{array}$$

Where the parameters α and β satisfy $\alpha > 0$
 $\beta > 0$

- If $\beta=1 \rightarrow f(x) = \frac{1}{\Gamma(\alpha)} \cdot x^{\alpha-1} \cdot e^{-x} \quad \left. \vphantom{\frac{1}{\Gamma(\alpha)}} \right\} \text{Standard Gamma}$



Vary α and β



$$E(x) = \alpha \beta$$

$$\text{Var}(x) = \alpha \beta^2$$

$$\sigma = \sqrt{\alpha \beta^2} = \beta \sqrt{\alpha}$$



Example reaction time x of a randomly selected individual to a certain stimulus has a standardized Gamma distribution with $\alpha = 2$ second. $\beta = 1$ second.

find $P(3 \leq x \leq 5)$

$$(A) \quad P(3 \leq x \leq 5) = \int_3^5 f(x) dx$$

$$f(x) = \frac{1}{\Gamma(2)\beta} \cdot x^{2-1} \cdot e^{-x} dx$$

$$\begin{aligned} P(3 \leq x \leq 5) &= \int_3^5 \frac{1}{\Gamma(2)} x^{1} \cdot e^{-x} dx = \frac{1}{1} \int_3^5 x e^{-x} dx \\ &= 1 \left[(-x e^{-x} - e^{-x}) \right]_3^5 = [-5e^{-5} - e^{-5} - (-3e^{-3} - e^{-3})] \\ &= 0.1589 \end{aligned}$$

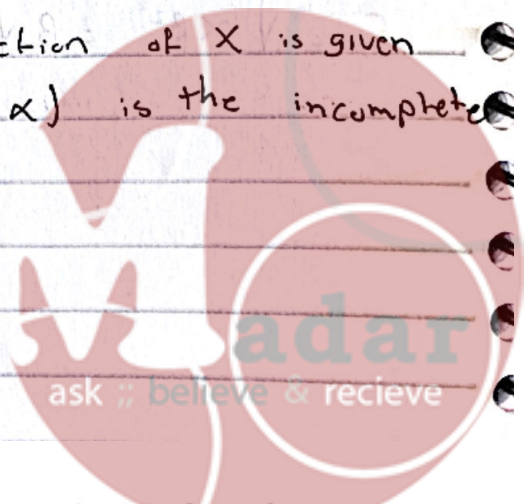
Incomplete Gamma Function (Table)

$$P(3 \leq x \leq 5) = F(5) - F(3)$$

$$\begin{aligned} (B) \quad P(x > 4) &= 1 - P(x \leq 4) = 1 - F(4) = 1 - 0.908 \\ &= 0.092 \end{aligned}$$

How to use the standard table if $\beta \neq 1$

Let x have a gamma distribution with parameter α and $\beta \neq 1$, then for any $x > 0$, the continuous density function of x is given by $P(x \leq x) = F\left(\frac{x}{\beta}; \alpha\right)$ when $F(\cdot; \alpha)$ is the incomplete gamma function



example Survival time X in weeks for mouse male exposed to 240 rads, gamma distribution with $\alpha = 8$ and $\beta = 15 \neq 1$; find:

① $E(X) = \alpha \beta = 8 \times 15 = 120$ weeks

② $\text{Var}(X) = \alpha \beta^2 = 8 \times (15)^2 = 1800$

③ $\sigma = \sqrt{1800} = 42.43$

④ Find the probability that survive between 60 and 120 weeks

$P(60 \leq X \leq 120) = F(120) - F(60) \rightarrow \text{wrong}$

$= P(X \leq 120) - P(X \leq 60)$

$= F\left(\frac{120}{15}; 8\right) - F\left(\frac{60}{15}; 8\right)$

$= F(8; 8) - F(4; 8) \rightarrow \text{from Table}$

$= 0.547 - 0.051 = 0.496$

⑤ at least 30 weeks

$P(X \geq 30) = 1 - P(X < 30) = 1 - P(X \leq 30)$

$= 1 - F\left(\frac{30}{15}; 8\right) \rightarrow 1 - F(2; 8) \rightarrow 0.999$

Applications of Gamma distribution:-

① The exponential distribution

② Chi-Squared

X is said to have an exponential distribution if the density function of X is given by $f(x) = \frac{1}{\beta} \cdot e^{-x/\beta}$ $x > 0, \beta > 0$

and this is gamma function where $\alpha = 1$



* The exponential density function can be easily integrated

$$F(x) = \int_0^x \frac{1}{\beta} e^{-x/\beta} dx = \frac{1}{\beta} [-\beta e^{-x/\beta}]_0^x$$

$$F(x) = 1 - e^{-x/\beta}, \quad x \geq 0$$

$$E(x) = \alpha \beta = \beta$$

$$\text{var}(x) = \alpha \beta^2 = \beta^2$$

example / response time x , has an exponential distribution with expected value as (time) equal to 5 seconds.

$$\textcircled{a} p(x \leq 10) = F(10) = 1 - e^{-x/\beta} = 1 - e^{-10/5}$$

$$p(x \leq 10) = 1 - e^{-2} = 1 - 0.135 = 0.865$$

$$\begin{aligned} \textcircled{b} p(5 \leq x \leq 10) &= F(10) - F(5) \\ &= (1 - e^{-10/5}) - (1 - e^{-5/5}) \\ &= (1 - e^{-2}) - (1 - e^{-1}) = 0.233 \end{aligned}$$



The Chi-squared distribution:

Let ν be a positive integer. Then a random variable X is said to have a chi-squared distribution with parameter ν if the probability distribution function of X is the Gamma density function with:

$$\alpha = \nu/2 \text{ and } \beta = 2$$

The p.d.f. of Chi-squared is:-

$$f(x; \nu) = \begin{cases} \frac{1}{2^{\nu/2} \Gamma(\nu/2)} \cdot x^{(\nu/2)-1} \cdot e^{-x/2}; & x \geq 0 \\ 0; & x < 0 \end{cases}$$

- * The parameter ν is called the number of degree of freedom (d.f.)
- * The symbol χ^2 is often used in place of chi-squared.

$$* E(x) = \alpha \beta = \nu/2 \times 2 = \nu$$

$$* \text{var}(x) = \alpha \beta^2 = \nu/2 \times 4 = 2\nu$$

$$* \sigma = \sqrt{2\nu}$$

example |:

$$\nu = 15$$

$$P(\chi^2_{15} \leq 5.33) =$$



Joint Distribution :-

Discrete } univariable
Continuous }

* If two random variable is to be considered simultaneously, we need to study what we call two dimensional or bivariate random variable of both the discrete and the continuous

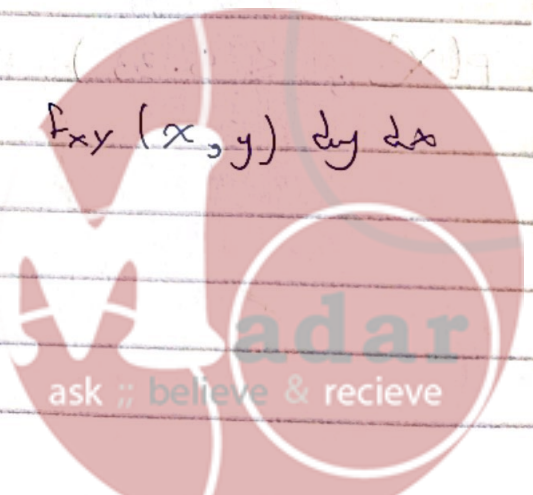
* For a joint densities, let x and y be two discrete random variable defined on the sample space S of an experiment. The joint probability mass function $f(x, y)$ is defined for each pair j (x, y) by $f_{xy}(x, y) = p(x=x_j \text{ and } y=y_j)$

* Properties *

- ① $f_{xy}(x, y) \geq 0$
- ② $\sum_{\text{all } x} \sum_{\text{all } y} f_{xy}(x, y) = 1$

Let x and y be continuous random variables. The ordered pair (x, y) is called a two dimensional continuous random variable. A function f_{xy} such that:

- ① $f_{xy}(x, y) \geq 0 \quad -\infty < x < +\infty \quad ; \quad -\infty < y < +\infty$
- ② $\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f_{xy}(x, y) dy dx = 1$
- ③ $p(a \leq x \leq b \text{ and } c \leq y \leq d) = \int_a^b \int_c^d f_{xy}(x, y) dy dx$



example

$X \backslash Y$	0	1	2	3	$f_{1x}(x)$
0	0.84	0.03	0.02	0.01	0.90
1	0.06	0.01	0.008	0.002	0.08
2	0.01	0.05	0.004	0.001	0.07
$f_{2y}(y)$	0.91	0.09	0.032	0.013	$f(x,y)$

x : defective welds in a car
 y : improperly tightened bolts
 $\downarrow = 1$
 كل وحدة من هي عبارة عن $f(x,y)$
 $\rightarrow = 1$

$$P(x=0 \text{ and } y=3) = 0.01$$

$$\sum_{x=0}^2 \sum_{y=0}^3 f(x,y) = 1 \quad (\text{مجموع الاحتمالات})$$

(A) no errors made by the robot:

$$p(x=0 \text{ and } y=0) = 0.84$$

(B) exactly one error is found:

$$= p(x=1 \text{ and } y=0) + p(x=0 \text{ and } y=1)$$

$$= f_{xy}(1,0) + f_{xy}(0,1)$$

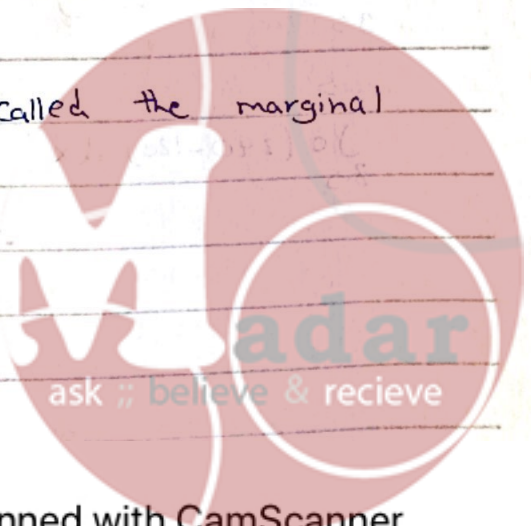
$$= 0.06 + 0.03 = 0.09$$

(C) no improperly tightened bolts:

$$p(y=0) = \sum_{x=0}^2 f_{xy}(x,0) = f_{xy}(0,0) + f_{xy}(1,0) + f_{xy}(2,0)$$

$$= 0.84 + 0.06 + 0.01 = 0.91$$

* The densities for x and y alone are called the marginal densities.



Let (X, Y) be a bivariate discrete random variable with joint density f_{xy} . The marginal density f_x , denoted by f_x , is given by

$$F_x(x) = \sum_{\text{any } y} f_{xy}(x, y)$$

The marginal density for y , denoted by f_y is given by
 $f_y = \sum_{\text{all } x} f_{xy}(x, y)$

* ① $f_{xy}(x, y) = \sum \sum f(x, y) = 1$

② $f_x(x) = 1$

③ $f_y(y) = 1$

ليزمر تطبعة

example | x is the calcium level } 8.5 - 10.5 mg/dL
 y is the cholesterol level } 120 - 240 mg/dL

range (20-29) year

8.5 - 120

8.5 - 240

10.5 - 120

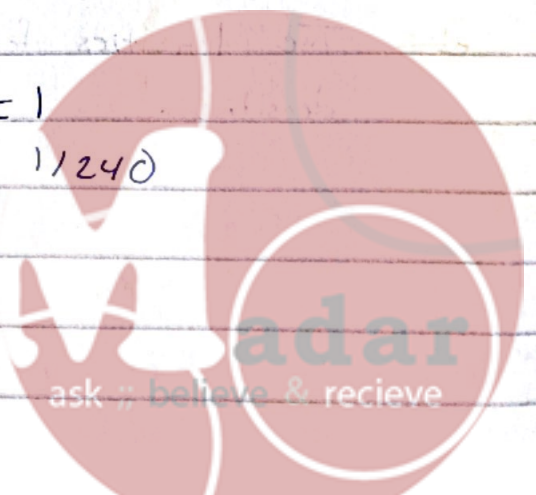
10.5 - 240

$$f_{xy}(x, y) = C$$

$$\int_{8.5}^{10.5} \int_{120}^{240} C \, dy \, dx = 1 \quad \text{from rules}$$

$$\int_{8.5}^{10.5} C(240-120) \, dx = 120 C [10.5-8.5] = 1$$

$$\therefore 240 C = 1 \rightarrow C = 1/240$$



→ find the probability that an individual's calcium level will lie between 9 and 10 mg/dL while the cholesterol level is between 125 and 140 mg/dL



24-8-2022

* Joint Distribution :-

{ Independence }

→ Two events are independent if knowledge that one occurred gives no clue to the likelihood that the other will occur.

* Let A_1 denote the event $X=x$

A_2 " " " " $Y=y$

If X and Y are independent :

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

$$(1) P(X=x \text{ and } Y=y) = P(X=x) \times P(Y=y)$$

or

$$(2) f_{xy}(x,y) = f_x(x) \cdot f_y(y)$$

* Let X and Y be random variable with joint density f_{xy} and the marginal densities f_x and f_y respectively X and Y are independent if and only if:

$$f_{xy}(x,y) = f_x(x) \cdot f_y(y), \text{ for all } x \text{ and } y$$

Example 1 x num of defective welds

y " " " " improperly tightened bolts

$$f_{xy}(x,y) \rightarrow f_{xy}(0,0) = 0.84$$



$$f_{xy}(0,0) \rightarrow f_x(x=0) = 0.9$$

$$f_y(y=0) = 0.91$$

$$f_x(0) \neq f_y(0) \neq f_{xy}(0,0)$$

$$0.9 \times 0.91 \stackrel{?}{=} 0.84$$

$$0.819$$

$$\stackrel{?}{=} 0.84$$

No

, because they are not equal so
 $\therefore$ They are dependent.

example X an individual blood calcium level.

Y = cholesterol "

$$f_{xy}(x,y) = \frac{1}{240}$$

$$f_x(x) = \frac{1}{2}$$

$$f_y(y) = \frac{2}{240}$$

marginal densities

$$f_{xy}(x,y) = \frac{1}{240} \stackrel{?}{=} f_x(x) + f_y(y) = \frac{1}{2} + \frac{2}{240} = \frac{1}{240}$$

$\therefore$ They are independent.

example (X) denote the number of hoses being used on the self-service;

(Y) denote the number of hoses on the full-service



H₁ H₂



H₁ H₂



$x \backslash y$	0	1	2	$F_{X X}$
0	0.1	0.04	0.02	0.16
1	0.08	0.20	0.06	0.34
2	0.06	0.14	0.30	0.5
$F_{Y Y}$	0.24	0.38	0.38	1

(a) What is $p(X=1 \text{ and } Y=1)$: 0.2

(b) compute $p(X \leq 1 \text{ and } Y \leq 1)$

$$p(X=0 \text{ and } Y=0) = 0.1$$

$$p(X=1 \text{ and } Y=1) = 0.2$$

$$p(X=0 \text{ and } Y=1) = 0.04$$

$$p(X=1 \text{ and } Y=0) = 0.08$$

$$0.42$$

(c) Compute $p(X \neq 0 \text{ and } Y \neq 0)$

$$= p(X=1 \text{ and } Y=1) + p(X=2 \text{ and } Y=1) + p(X=2 \text{ and } Y=2) + p(X=1 \text{ and } Y=2)$$

(d) Compute the marginal of X and Y

$$(e) f_{xy}(0,0) = 0.1$$

$$f_X(0) = 0.16, f_Y(0) = 0.24$$

$$f_X(0) \times f_Y(0) = 0.16 \times 0.24$$

$$= 0.0384 \neq 0.1$$

So they're dependent

• independent $\rightarrow$ independent $\rightarrow$ independent $\rightarrow$ independent $\rightarrow$ independent

عشان نتويز اذا dependent اولاد

So they're dependent

* Conditional densities

Assume that (X, Y) is discrete with joint density f_{xy} and marginal densities f_x and f_y . Let A_1 denotes the event that $X=x$ and A_2 denotes the event that $Y=y$

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

$$P(X=x/Y=y) = \frac{P(X=x \text{ and } Y=y)}{P(Y=y)} = \frac{f_{xy}(x, y)}{f_y(y)}$$

In general let (X, Y) be a two-dimensional random variable with joint density f_{xy} and marginal densities f_x and f_y . Then;

① The conditional density for X given $Y=y$ denoted by $f_{x/y}$ is given by:

$$f_{x/y}(x) = \frac{f_{xy}(x, y)}{f_y(y)} \quad ; \quad f_y(y) > 0$$

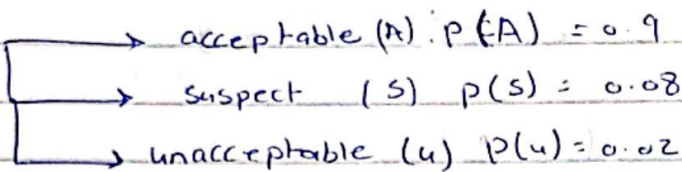
② The conditional density for Y given that $X=x$ denoted by $f_{y/x}(y)$ is given by

$$f_{y/x}(y) = \frac{f_{xy}(x, y)}{f_x(x)} \quad , \quad f_x(x) > 0$$



Example 1

received bit

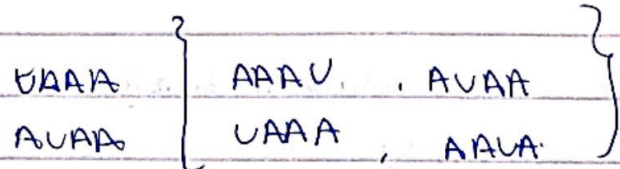


→ let x number of acceptable bits and y number of suspect bits.

→ (independent) rating

① $P(y=0 | x=3)$

② $P(y=1 | x=3)$



$x \backslash y$	0	1	2	3	4
0	⊙	.	.	.	.
1					
2					
3					
4					

UUUU

$$p(u) \times p(u) \times p(u) \times p(u)$$

$$f_{xy}(0,0) = (0.02)^4 = 1.6 \times 10^{-7}$$

pts

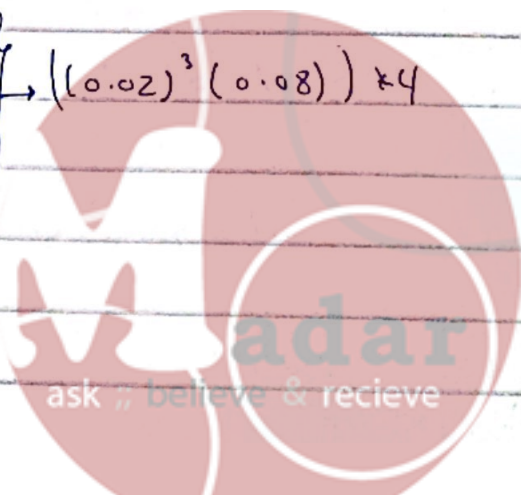
$$p(SUUU) \times 4$$

USUU

UUSU

UUUS

$$\left. \begin{array}{l} USUU \\ UUSU \\ UUUS \end{array} \right\} \rightarrow ((0.02)^3 (0.08)) \times 4$$



28-8-2022

* Conditional Joint probability:

$$P(A_1/A_2) = \frac{P(A_1 \cap A_2)}{P(A_2)} \quad \text{one dimension}$$

$$P(X=x|Y=y) = \frac{P(X=x \text{ and } Y=y)}{P(Y=y)} = \frac{P_{xy}(x, y)}{P_Y(y)}$$

acceptable (A) (0.9)

suspect (S) (0.08)

unacceptable (U) (0.02)

X = acceptable {0, 1, 2, 3, 4}

Y = suspect {0, 1, 2, 3, 4}

Z = unacceptable {0, 1, 2, 3, 4}

X=0 Y=0 Z=4 $\rightarrow P(UUUU) = 0.02^4$

Y=1 Z=3 $\rightarrow P(SUUU) = P(USUU) = P(UUSU) = (0.08)^1 (0.02)^3$

Y=2 Z=2 $\rightarrow P(SSUU) = P(SUSU) = P(UUSS) = (0.08)^2 (0.02)^2$

Y=3 Z=1

Y=4 Z=0

		Y	0	1	2	3	4
X=1	Y	X	.	.	.	.	.
	0	0	.	.	.	.	.
	1	1	.	.	.	.	X
	2	2	.	.	.	X	X
X=2	Y	X	.	.	X	X	X
	0	3	.	.	X	X	X
	2	4	.	X	X	X	X
X=3	Y	X	.	X	X	X	X
	0	1	.	X	X	X	X



X	P
0	$\binom{4}{0} (0.9)^0 (0.1)^4$
1	$\binom{4}{1} (0.9)^1 (0.1)^3$
2	$\binom{4}{2} (0.9)^2 (0.1)^2$
3	$\binom{4}{3} (0.9)^3 (0.1)^1$
4	$\binom{4}{4} (0.9)^4 (0.1)^0$

Y	P
0	$\binom{4}{0} (0.08)^0 (0.92)^4$
1	$\binom{4}{1} (0.08)^1 (0.92)^3$
2	$\binom{4}{2} (0.08)^2 (0.92)^2$
3	$\binom{4}{3} (0.08)^3 (0.92)^1$
4	$\binom{4}{4} (0.08)^4 (0.92)^0$

$$P(X=1, Y=1) = P(ASUV) = (0.92)(0.08)(0.02)^2$$

$$P(X=2, Y=2) = P(AASS) = (0.92)^2(0.08)^2$$

X \ Y	0	1	2	3	4	$f_x(x)$
0	1.6×10^{-7}	2.56×10^{-6}	1.54×10^{-5}	4.1×10^{-5}	4.1×10^{-5}	
1	2.88×10^{-5}	3.46×10^{-4}	1.38×10^{-3}	1.84×10^{-3}		
2	1.94×10^{-3}	1.56×10^{-2}	3.11×10^{-2}			
3	5.83×10^{-2}	0.2833				$5.83 \times 10^{-2} + 0.2833$
4	0.6561					

$$a) P(Y=0 | X=3) = \frac{P(Y=0 \text{ and } X=3)}{P(X=3)} = \frac{f_{xy}(3,0)}{f_x(3)}$$

$$P(Y=0 | X=3) = \frac{5.83 \times 10^{-2}}{0.2916} = 0.2$$

$$\Sigma = 1$$

$$b) P(Y=1 | X=3) = \frac{f_{xy}(3,1)}{f_x(3)} = \frac{0.2833}{0.2916} = 0.8$$



$$P(AAAU) = 0.0458 = P(AAUA) = P(AUAA) = P(UAAA)$$

* Given that $(X=3)$, the only possible values for y are 0 or 1, $P(Y=0/X=3) + P(Y=1/X=3) = 0.2 + 0.8 = 1$

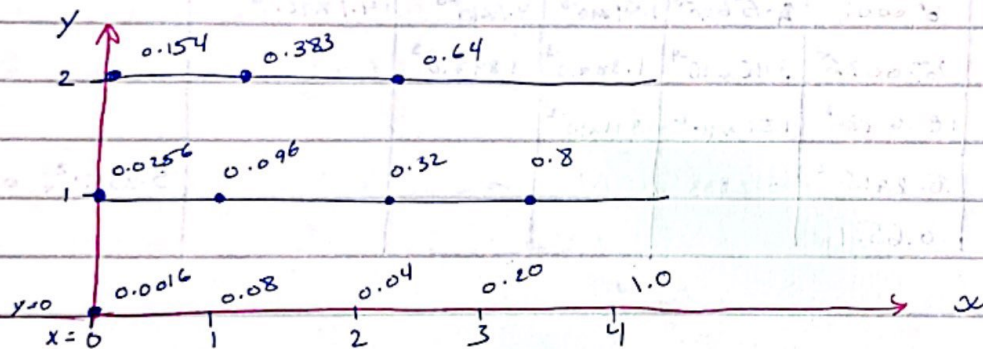
If $X=2$, $Y = \{0, 1, 2\}$

$$P(Y=0/X=2) + P(Y=1/X=2) + P(Y=2/X=2) = 1$$

$$(1) f_{Y/X}(y) \geq 0$$

$$(2) \sum f_{Y/X}(y) = 1$$

نحتاج أن نتحقق



Let R_x denote the set of all points in the range of (X, Y) for which $X=x$.

the conditional mean of Y given that $X=x$, denoted by $E(Y/X)$ or $\mu_{Y/X}$, is

$$E(Y/X) = \sum_{R_x} y f_{Y/X}(y)$$

$$E(Y/2) = 0 \times 0.04 + 1 \times 0.32 + 2 \times 0.64 = 1.6$$

$$E(Y/3) = 0.8$$

and the conditional variance of Y given $X=x$, denoted

by $\text{Var}(Y/X)$ or $\sigma^2_{Y/X}$ is $V(Y/X) = \sum_{R_x} (y_i - \mu_{Y/X})^2$

$f_{Y/X}(y)$

$$\text{Var}(y/2) = (0 - 1.6)^2 (0.04) + (1 - 1.6)^2 (0.32) + (2 - 1.6)^2 (0.64) = 0.32$$



29-8-2022

* expectation and covariance :-

- expected value : let (X, Y) be a two-dimensional random variable with joint density f_{xy} . let $H(X, Y)$ be a random variable then the expected value of $H(X, Y)$ denoted by $E(H(X, Y))$ is given by :

① for (X, Y) discrete variable $E(H(X, Y)) = \sum_{\text{all } x} \sum_{\text{all } y} H(x, y) \cdot f_{xy}(x, y)$

② for (X, Y) continuous variable $E(H(X, Y)) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} H(x, y) \cdot f_{xy}(x, y) dx dy$

example

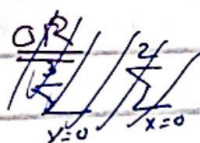
$x \backslash y$	0	1	2	3	$f_x(x)$
0	0.84	0.03	0.02	0.01	0.9
1	0.06	0.01	0.008	0.002	0.08
2	0.01	0.005	0.004	0.001	0.02
$f_y(y)$	0.91	0.045	0.032	0.013	1.00

$X \rightarrow$ random variable denoted the number of welds

$Y \rightarrow$ " " " " " : improperly tightened bolts; find.

② $E(X) = \sum_{x=0}^{\infty} \sum_{y=0}^{\infty} x \cdot f_{xy}(x, y)$

$$= 0(0.84) + 0(0.03) + 0(0.02) + 0(0.01) + 1(0.06) + 1(0.008) + 1(0.002) + 2(0.01) + 2(0.005) + 2(0.004) + 2(0.001) = 0.12$$



OR

$$E(x) = \sum_{x=0}^2 f_x(x) \cdot x = 0(0.9) + 1(0.08) + 2(0.02) \\ = 0.12$$

$$\textcircled{B} E(y) = \sum_{y=0}^3 \sum_{x=0}^2 y \cdot f_{xy}(x, y)$$

$$= 0 [0.84 + 0.06 + 0.01] + 1 [0.03 + 0.01 + 0.005] + \dots \\ + 3 [0.01 + 0.002 + 0.001] = 0.148$$

$$\text{or} \\ E(y) = \sum_{y=0}^3 y \cdot f_y(y) = 0(0.91) + 1(0.045) + 2(0.032) + 3(0.013) \\ = 0.148$$

$$E(x+y) = \sum_{x=0}^2 \sum_{y=0}^3 (x+y) \cdot f_{xy}(x, y)$$

$$= (0+0)(0.84) + (0+1)(0.03) + (0+2)(0.02) + (0+3)(0.01) \\ + (1+0)(0.06) + (1+1)(0.01) + \dots + (2+0)(0.01) \\ + (2+1)(0.005) + (2+2)(0.004) + (2+3)(0.001) = \underline{\underline{0.268}}$$

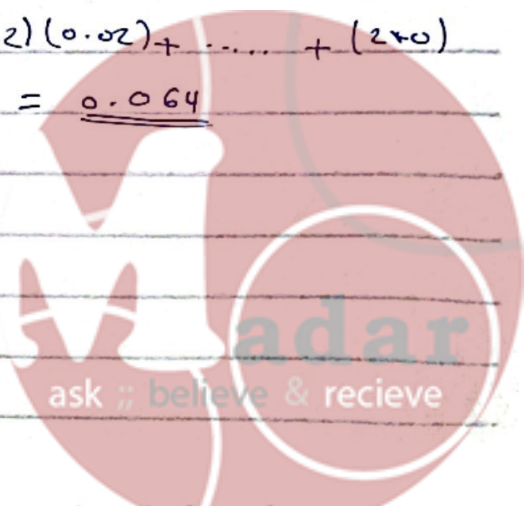
$$E(x) = 0.12, \quad E(y) = 0.148$$

$$\text{or} \quad E(x+y) = E(x) + E(y) = 0.12 + 0.148 = \underline{\underline{0.268}}$$

$$E(xy) = \sum_{x=0}^2 \sum_{y=0}^3 xy \cdot f_{xy}(x, y)$$

$$= (0 \times 0)(0.84) + (0 \times 1)(0.03) + (0 \times 2)(0.02) + \dots + (2 \times 0) \\ (0.01) + \dots + (2 \times 3)(0.001) = \underline{\underline{0.064}}$$

$$E(xy) \neq E(x) \times E(y)$$



* Covariance:-

It is a common measure of the relationship between two random variable.

→ Let x & y be random variable with means μ_x and μ_y .

The covariance between x and y , $\text{cov}(x, y)$ or σ_{xy} is given by:

$$\text{cov}(x, y) = E((x - \mu_x)(y - \mu_y))$$

$$\text{cov}(x, y) = E(xy) - E(x)E(y)$$

* From the previous exam:-

$$E(x) = 0.12$$

$$E(y) = 0.148$$

$$E(xy) = 0.064$$

$$\text{cov}(x, y) = 0.064 - (0.12)(0.148) = 0.04624$$

* A covariance of 0 implies that the knowledge that x assumes a value above its mean gives no indication as the value of y relative to its mean (μ).

* If small values of x tend to be associated with small values of y and large values of x tend to be associated with large values of y

then $(x - \mu_x)$ and $(y - \mu_y)$ will usually have the same algebraic signs. This implies that $(x - \mu_x)(y - \mu_y)$ will be positive, $\text{cov}(x, y)$ will be positive.



* If small values of x tend to be associated with large values of y and vice versa, this results in -ve value for $(x - \mu_x)(y - \mu_y)$ yielding a -ve covariance.

• How x and y vary relative to one another.

* It can be shown that whenever we have the random variable x and y are independent the covariance between x and y is zero.

* Let (x, y) are two-dimensional random variables with joint density f_{xy} . If x and y are independent then

$$E(xy) = E(x) * E(y) \rightarrow \text{cov}(x, y) = 0$$

* Correlation:

It is another relationship between two random variables, easier to interpret than covariance.

* Let x and y be random variables with means μ_x and μ_y and variances σ_x^2 and σ_y^2 , respectively. The correlation, ρ_{xy} a person coefficient, between x and y is given by

$$\rho_{xy} = \frac{\text{Cov}(x, y)}{\sqrt{(\text{Var } x)(\text{Var } y)}} = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$$

* possibilities of ρ_{xy} :

① positive: the correlation between x and y is positive.

② Negative: " " " " " " " " negative.

③ Zero: " " " " " " " " zero.



* The correlation coefficient ρ_{xy} for any two random variables X and Y lies between -1 and $+1$ inclusive

$$-1 \leq \rho_{xy} \leq +1$$

If $\rho_{xy} = 1$, then X and Y have perfect positive correlation

This implies that $Y = \beta_0 + \beta_1 X$

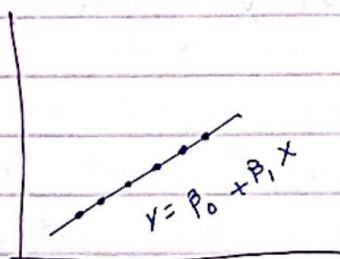
$\beta_1 > 0$ i.e. small values of X are associated with small values of Y , large values of X are associated with large values of Y .

* If $\rho_{xy} = -1$, then X and Y have perfect -ve correlation

$$Y = \beta_0 + \beta_1 X, \text{ where } \beta_1 < 0$$

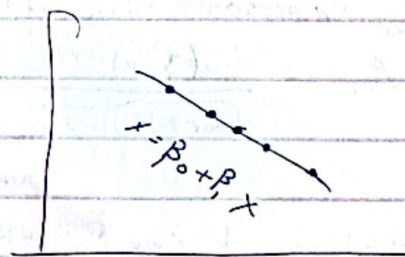
* values of ρ_{xy} near -1 or $+1$ do occur and indicate a linear trend between them

* If $\rho_{xy} = \text{zero}$, we say that X and Y are uncorrelated. but not related, If there is a relationship between them, then it's not linear.



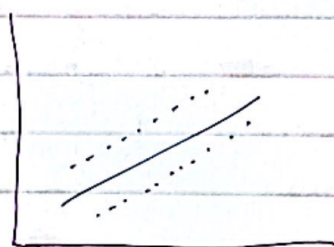
perfect positive

$$\rho_{xy} = +1$$



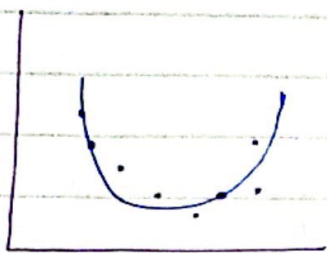
perfect negative

$$\rho_{xy} = -1$$



linear

$$-1 \approx \rho_{xy} \approx +1$$



uncorrelated.

$$\rho_{xy} = 0$$

example

Car welding x

untightened y

$$\rho_{xy} = \frac{\text{cov}(x, y)}{\sqrt{\underset{\text{var}}{\text{var } x}} \sqrt{\underset{\text{var}}{\text{var } y}}} = \frac{0.064}{\sqrt{0.146} \sqrt{0.268}} = \underline{0.23} < +1$$

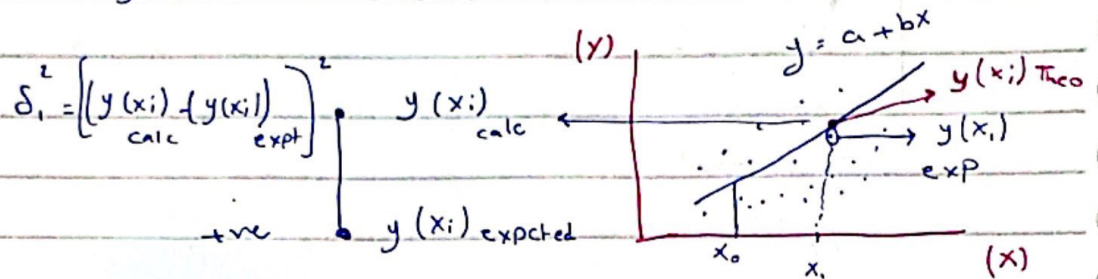
$$\text{var}(x) = E(x^2) - (E(x))^2 = 0.16 - (0.12)^2 = 0.146$$

$$\text{var}(y) = E(y^2) - (E(y))^2 = 0.29 - (0.148)^2 = 0.268$$



30-8-2022

* Simple linear regression and Correlation :-



$$G_2^2 = \left([y(x_0)_{\text{calc}} - y(x_0)_{\text{exp}}] \right)^2$$

; true

$$S = \sum_{i=0}^{\infty} [y(x_i)_{\text{cal}} - y(x_i)_{\text{given}}]^2$$

* linear model : $A + Bx_i$

$$\text{error} \leftarrow S^2 = \sum_{i=0}^{\infty} \left([A + Bx_i]_{\text{calculated}} - y(x_i)_{\text{experimental}} \right)^2 = \sum_{i=0}^{\infty} [y(x_i)_{\text{exp}} - A - Bx_i]^2$$

→ To minimize the $\sum$ of squares of errors, we've to find the best parameters $\{A, B\}$ values that make the $\sum$ minimum.

$$\frac{\partial S}{\partial A} = 0, \quad \frac{\partial S}{\partial B} = 0$$

$$= 2 \sum_{i=0}^{\infty} [y(x_i)_{\text{exp}} - A - Bx_i] \cdot (-1)$$

$$= 2 \sum_{i=0}^{\infty} [y(x_i)_{\text{exp}} - A - Bx_i] \cdot (-x_i)$$



$$\sum_{i=0}^{\infty} y(x_i)_{cal} - \sum_{i=0}^{\infty} A - \sum_{i=0}^{\infty} B x_i = 0 \quad \dots (1)$$

$$\sum_{i=0}^{\infty} x_i y(x_i)_{cal} - \sum_{i=0}^{\infty} A x_i - \sum_{i=0}^{\infty} B x_i^2 = 0$$

$$\sum_{i=0}^{\infty} y(x_i)_{exp} = A \sum_{i=0}^{\infty} (1) + B \sum_{i=0}^{\infty} x_i$$

$$\sum_{i=0}^{\infty} x_i y(x_i)_{exp} = A \sum_{i=0}^{\infty} x_i + B \sum_{i=0}^{\infty} x_i^2$$

x_i	$y(x_i)_{exp}$
x_0	$y(x_0)$
x_1	$y(x_1)$
x_2	$y(x_2)$
x_3	$y(x_3)$
x_4	$y(x_4)$

$$\begin{bmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^2 \end{bmatrix} \begin{Bmatrix} A \\ B \end{Bmatrix} = \begin{Bmatrix} \sum y(x_i) \\ \sum x_i y(x_i) \end{Bmatrix}$$

System of linear equation

Solving this system to find A and B, these parameters are the best values that represent the linear with a minimum $\sum$ of squares of error.

(2) polynomial of 2nd degree $y = A + Bx_i + Cx_i^2$

$$S = \sum_{i=1}^n [y(x_i)_{exp} - A - Bx_i - Cx_i^2]^2$$

$$\frac{\partial S}{\partial A} = \frac{\partial}{\partial A} \sum_{i=1}^n (A + Bx_i - f(x_i)) = 0$$

$$\frac{\partial S}{\partial B} = \frac{\partial}{\partial B} \sum_{i=1}^n (A + Bx_i - f(x_i)) (x_i) = 0$$

$$\frac{\partial S}{\partial A} = 0$$

$$\frac{\partial S}{\partial B} = 0$$

ask :: believe & recieve

$$\frac{\partial S}{\partial c} =$$

$$\begin{bmatrix} n & \sum x_i & \sum x_i^2 \\ \sum x_i & \sum x_i^2 & \sum x_i^3 \\ \sum x_i^2 & \sum x_i^3 & \sum x_i^4 \end{bmatrix} \begin{Bmatrix} A \\ B \\ C \end{Bmatrix} = \begin{Bmatrix} \sum y(x_i) \exp \\ \sum y(x_i) \exp(x_i) \\ \sum y(x_i) \exp(x_i^2) \end{Bmatrix}$$

* Sys of linear eq $\rightarrow$ by solving the sys, the sys can have the best values of A, B, C

* Non-linear equation:

$$\text{I} \quad y = ab^x \rightarrow \ln y = \ln a + x \ln b$$

$$y = A + xB$$

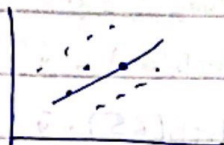
x_i	$y(x_i)$	$\ln y(x_i)$
x_1	$y(x_1)$	$\ln y(x_1)$
x_2	$y(x_2)$	$\vdots$
x_3	$y(x_3)$	$\vdots$
x_4	$y(x_4)$	$\vdots$
$\sum x_i$	$\sum y(x_i)$	$\sum \ln y(x_i)$

$$\rightarrow \begin{bmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^2 \end{bmatrix} \begin{Bmatrix} A \\ B \end{Bmatrix} = \begin{Bmatrix} \sum y \\ \sum xy \end{Bmatrix}$$

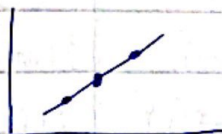


$$[2] \quad y = \frac{1}{a+bx} \rightarrow \frac{1}{y} = a + bx = y = a + bx$$

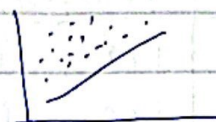
$R \rightarrow$ coefficient of regression ($0 \rightarrow 1$)



$$R^2 = 0.8$$



$$R^2 = 1$$



$$R^2 \approx 0$$

big difference and
poor model

$0 \leq R^2 \leq 1 \rightarrow$ exactly overlap the points



31.8.2022

→ Regression :-

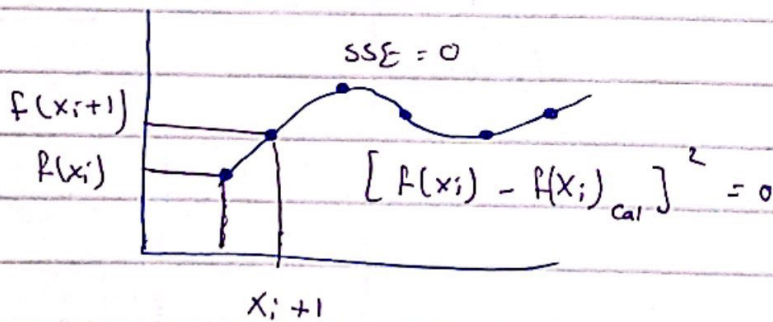
x_i	$y(x_i)$	$[y_i - y_{cal}(x_i)]^2$	$[y(x_i) - \bar{y}]^2$	$[y_{cal}(x_i) - \bar{y}]^2$
60	1.69	$\{1.69 - [A+B(60)]\}^2$	$\{1.69 - \bar{y}\}^2$	$\{A+B(60) - \bar{y}\}^2$
65	1.60	$\{1.6 - [A+B(65)]\}^2$	$\{1.6 - \bar{y}\}^2$	$\{A+B(65) - \bar{y}\}^2$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\bar{y} = \frac{\sum y(x_i)}{n}$		$\sum [y(x_i) - y_{cal}(x_i)]^2$ SSE	$\sum [y(x_i) - \bar{y}]^2$ SST	$\sum [y_{cal}(x_i) - \bar{y}]^2$ SSR

$$y(x_i) = A + Bx_i$$

Coefficient
of regression

$$r^2 = \left\{ \frac{SST - SSE}{SST} \right\} \quad \text{or} \quad r^2 = \left\{ \frac{SSR}{SST} \right\}$$

$$SST = SSE + SSR$$

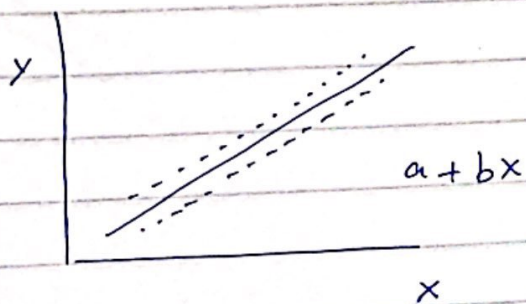


$$r^2 = \frac{SST}{SST} = 1$$

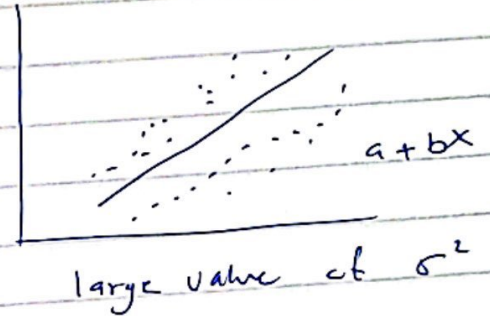


• estimating σ^2 :-

- The parameter σ^2 is a measure of the amount of variability inherent in the regression model.
- A large value of σ^2 will lead to observed (x_i, y_i) that are quite spread out the regression line.



Small value of σ^2



Large value of σ^2

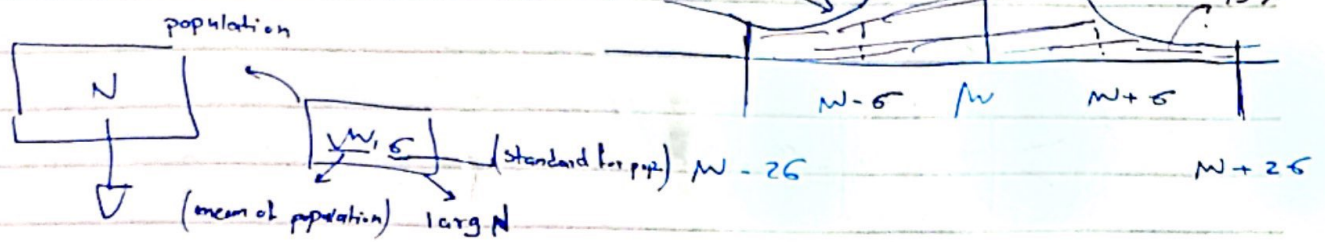
• When σ^2 is small the observed points will tend to fall very close to the true line.

$$\sigma^2 = \frac{SSE}{n-2} = \frac{\sum (y_i - y_{cal})^2}{n-2}$$



Confidence intervals:-

→ شوي الفترة اللي احاطتكون
بناطرين اها رح تكون فيها



sample → small n

$\bar{X}$, s (standard deviation)
→ (mean for the sample)

* When $n \approx N \rightarrow \left(\bar{X} \approx \mu, s \approx \sigma \right) \rightarrow \text{large } (n)$
estimators

* Estimators $\rightarrow (\bar{X}, s)$

① unbiased : $\bar{X} \approx \mu$

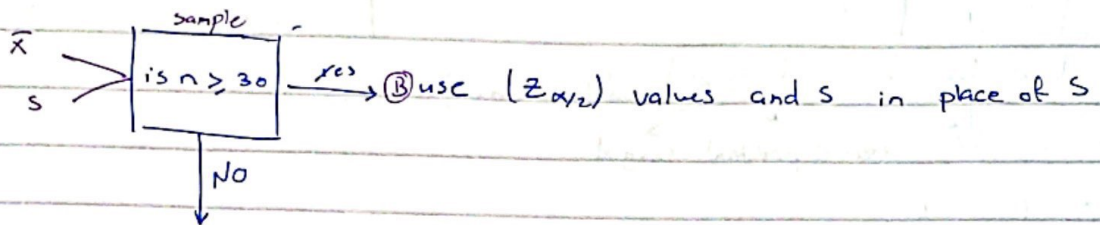
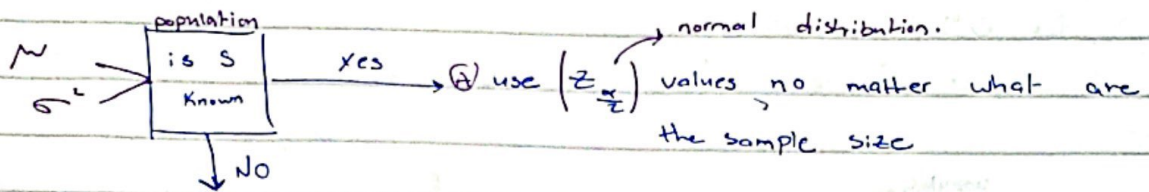
② consistent & as n increases : $\bar{X} \rightarrow \mu$

③ Relatively efficient estimators .

* smallest variance and hence standard deviation - s very small
with respect to the mean value.



① Confidence interval for mean.



③ use $t_{\frac{\alpha}{2}}$ values and S in the formula.

② Confidence interval for variance.

③ Confidence interval for standard deviation.

① Confidence interval for mean → for case A & B we use these formulas to obtain the confidence interval.

$$\bar{X} - \left(Z_{\frac{\alpha}{2}} \right) \left(\frac{\sigma}{\sqrt{n}} \right) < \mu < \bar{X} + \left(Z_{\frac{\alpha}{2}} \right) \left(\frac{\sigma}{\sqrt{n}} \right) \quad (\text{population})$$

S is \sigma & \mu is sample \mu formula \mu is \mu

maximum error of estimate

error → $E = \left(Z_{\frac{\alpha}{2}} \right) \left(\frac{\sigma}{\sqrt{n}} \right)$

Sample size $\sqrt{n} = \frac{\left(Z_{\alpha/2} \right) (\sigma)}{E}$

$n = \frac{\left(Z_{\alpha/2} \right) (\sigma)}{E}$

n depends on

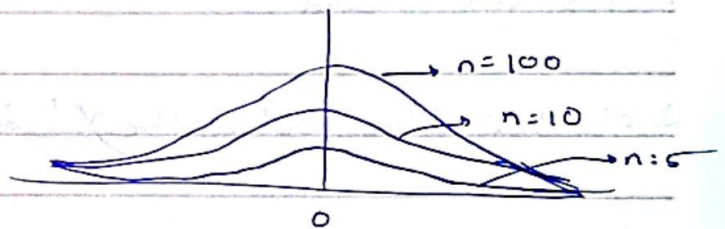
- error of the estimators (E)
- population standard deviation (σ)
- degree of confidence ($z_{\frac{\alpha}{2}}$)

* For case C, the formula we use:

$$\bar{X} - \left(t_{\frac{\alpha}{2}}\right) \left(\frac{s}{\sqrt{n}}\right) < \mu < \bar{X} + \left(t_{\frac{\alpha}{2}}\right) \left(\frac{s}{\sqrt{n}}\right)$$

∴ T curve:

→ when n is large enough
the t curve behaves
like Normal curve



② Confidence intervals for Variance:

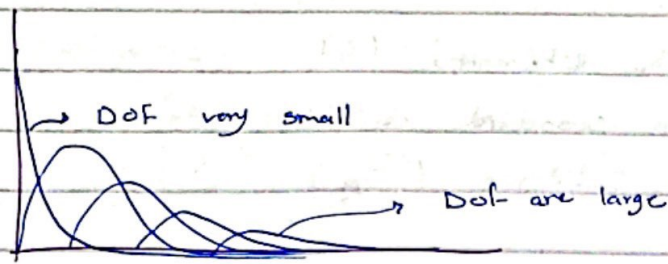
We use this formula to calculate the confidence intervals.

$$\text{DoF} \leftarrow \frac{(n-1)s^2}{\chi^2_{\text{right}}} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_{\text{left}}}$$

③ Confidence interval for standard deviation:

$$\sqrt{\frac{(n-1)s^2}{\chi^2_{\text{right}}}} < s < \sqrt{\frac{(n-1)s^2}{\chi^2_{\text{left}}}}$$





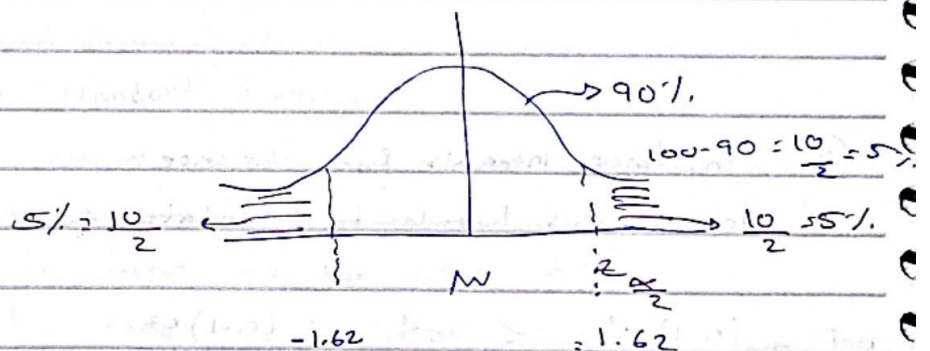
→ The area under the curve is always = 1.00 or 100 %.

* χ^2 can't be -ve

* positively skewed.

→ at about 100 Dof, the χ^2 distribution becomes somewhat symmetric

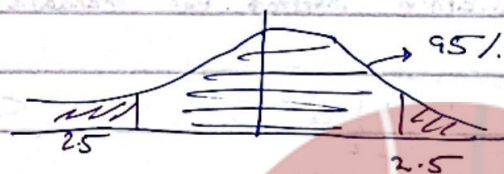
90%



95%

$$\rightarrow 100 - 95 = 5$$

$$\frac{5}{2} = \frac{5}{2} = 2.5\%$$



$$95 + 2.5 = 97.5$$

ask .. believe & recieve