دفتر

statistics

للدكتور : مراد احمد

2016-2017

اعداد الطالبه

وداد رضوان

Statistics

إحصاء

عمل الطالبة :- وداد رضوان

من شرح الدكتور مراد أحمد

2016/2017

Omega JU ♡

①

Types of data:-

[1] Qualitative نوعي (the type of a car)

[2] Quantitave كمي (the numbers of cars)

Discrete (countable)

countinuous (measurable)

The data comes in one of the following forms:-

[1] Raw data

1, 2, 3, 4, ...

[2] Frecuency distributions

X	1	2	3	4	5
f	100	50	25	20	19

OR

X	A	B	C	D	F
f	10	5	7	8	2

[3] stem leaf diagram

	stem	leaf
11	1	1 1 2
22	2	2 2 3

key 1|2 means 12

[4] Grouped frecuency distributions :-

I	0-4	5-9	10-14	15-19
f	7	2	8	3

Measures Of ~~Spread~~ Central tendency

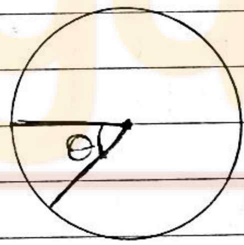
1. the mean ($\bar{X}$) متوسط حسابي
2. the median (Q_2) وسيط
3. the mode الأكثر تكراراً

Measures Of Spread

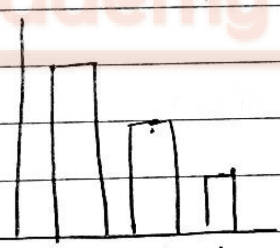
1. Range
2. The interquartile range
3. The interpercentile range
4. The standard deviation (S)
5. The varians (S^2)

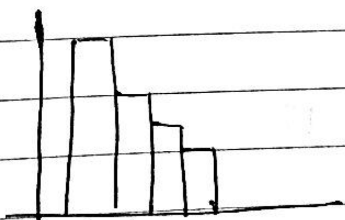
Data representation

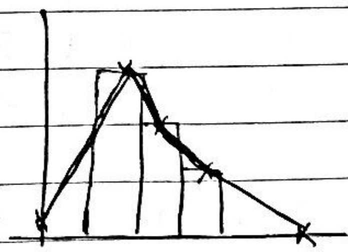
1. pie Chart
2. Bar chart
3. histagram
4. paliyon
5. C.f Curve

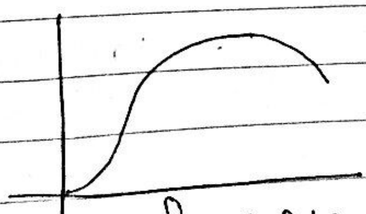


pie chart


Bar chart


Histagram


Paliyon


C.f curve

Measures of central kendray:-

[1] The Mean ($\bar{X}$)

[1] for raw data:-

e.g) find the mean for:-

$2, 7, 5, 11, 5$

sol) $\bar{X} = \dfrac{2+7+5+11+5}{5} = \dfrac{30}{5}$

$\bar{X} = 6$

$$\bar{X} = \frac{\sum_{x=1}^{n} x}{n}$$

(مجموع القيم على عددهم)

e.g) if the mean of $5, 11, 7, 2, x$ is 6 Find x

sol.)
$\bar{x} = 6$
$n = 5$

$\bar{x} = \dfrac{\sum x}{n} \Rightarrow 6 = \dfrac{5+11+7+2+x}{5}$

$30 = 25 + x$

$\boxed{x = 5}$

e.g) if the mean of $x, y, 12$ is 10 Find the mean of x, y

the mean $x, y = \dfrac{x+y}{2}$

$\bar{x} = \dfrac{x+y+12}{3}$

$30 = x + y + 12$

$x + y = 18$

$\dfrac{x+y}{2} = 9$

e.g) if the mean mark of (10) students is (12) the Student with mark (15) joined the class, find the new mean :-

Sol.) $\bar{X} = 12$, $n = 10$

$\sum x = 12 \times 10$

$\sum x = 120$

new $\sum x = 120 + 15 = 135$

new $n = 11$

new mean $= \frac{135}{11} = 12.27$

e.g) if the mean mark of (10) students boys is (12) & the mean mark of (12) girls is (10) find the mean mark of students all together :-

Boys

$n = 10$

$\bar{X} = 12$

$\sum x = 120$

Girls

$n = 12$

$\bar{X} = 10$

$\sum x = 120$

all together :-

$n = 22$

$\sum x = 240$

$\bar{X} = \frac{240}{22}$ $\bar{X} = 10.9$

[2] Mean for stem and leaf :-

e.g) find the mean of (stem and leaf)

S	L
1	1 1 2
2	2 2 4
3	2 4 6

Key 1|2 means 12

$$\bar{X} = \frac{11+11+12+22+22+24+32+34+36}{9} \approx 22.6$$

3] for frequency distributions :-

e.g)

X	1	2	3	4	5	sum
f	3	8	7	2		20
fx	3	16	21	8		48

f,x نضرب

$\bar{x} = \frac{48}{20} = 2.4$

$\bar{x} = \frac{\sum fx}{\sum f}$

$fx = f * X$

e.g) if $\bar{x} = 25$ find k :-

Sol)

X	1	2	3	4	5	sum
f	3	8	k	2	5	18
f.x	3	16	3k	8	25	52 + 3k

$25 = \frac{52 + 3k}{18}$

$k = -14$ ← ليس من المفروض تطلع سالب بس لأنه سؤال عشوائي .

4] Mean for frequency distributions (Grouped)

$\bar{x} = \frac{\sum fx}{\sum f}$ لازم نطلع (X) ⇐ نجمع الفئتين وتقسم على (2)

e.g)

I	0-4	5-9	10-14	15-20	sum
f	3	8	7	2	20
X	2	7	12	17	
fx	6	56	84	34	190

$\bar{x} = \frac{190}{20} = 9.5$

this is just an estimate for the mean because it is continuous e.g) If $\bar{X} = 9$ find k :-

I	0-4	5-9	10-14	15-19	sum
f	3	k	7	2	

Sol)

x	2	7	12	17	12+k
fx	6	7k	84	34	124+7k

$$\bar{X} = \frac{\sum fx}{\sum x}$$

$$9 = \frac{124+7k}{12+k}$$

OR $108 + 9k = 124 + 7k$

$2k = 16$

$k = 8$

$\boxed{2}$ The mode :-

The value that occurs mostly (or with highest frecuency)

e.g) 2, 7, 5, 1, 5 . mode = 5

e.g) 2, 7, 5, 11, 3, 5, 7 mode = 5, 7

e.g) 2, 7, 5, 11, 3 . No mode

e.g)

x	1	2	3	4	5
f	3	8	7	6	5

mode = 2

e.g)

I	0-4	5-9	10-14	15-19
f	3	7	8	2

modal class :- 10-14

mode = $\frac{10+14}{2} = 12$

3) The median (Q_2):-

1) for raw data:-
The middle data after arranging the data in ascending (or disascending order)

e.g) 2, 7, 5, 11, 3 .. find Q_2?
Sol.) $\cancel{2}, \cancel{3}, 5, \cancel{7}, \cancel{11}$ ①ترتب
$Q_2 = 5$ ② نأخذ الرقم الأوسط

e.g) 2, 7, 5, 11, 3, 7
Sol.) $\cancel{2}, \cancel{3}, 5, 7, \cancel{7}, \cancel{11}$
$Q_2 = \frac{5+7}{2} = 6$

Note:- $\frac{n}{2}$

(كسر) fraction → next integer (القيمة التالية)

whole no. (k) (عدد صحيح) → $\frac{k^{th} + (k^{th}+1)}{2}$ value

n → row data, stem and leaf (عدد الأرقام)
n → Frequency, grouped distribution (Σf مجموع التكرار)

e.g) $\underset{1}{2}, \underset{2}{3}, \underset{3}{5}, \underset{4}{7}, \underset{5}{8}, \underset{6}{10}, \underset{7}{11}$ (ordered)

Sol.) $\frac{n}{2} = \frac{7}{2} = 3.5 \longrightarrow$ next integer (4^{th} value)

$Q_2 = 7$

e.g) $\underset{1}{2}, \underset{2}{3}, \underset{3}{5}, \underset{4}{7}, \underset{5}{8}, \underset{6}{10}, \underset{7}{12}, \underset{8}{12}$ (ordered)

$\frac{n}{2} = \frac{8}{2} = \frac{4^{th} + 5^{th}}{2}$ value

$Q_2 = \frac{7+8}{2} = 7.5$

e.g)

S	L
1	1 1 2 2
2	2 2 3 4
3	4 4 5 6

(ordered)

sol.) $\frac{n}{2} = \frac{12}{2} = \frac{6^{th} + 7^{th}}{2}$

$Q_2 = \frac{22+23}{2} = 22.5$

[2] For frequency distribution :-
Same as the previous case but after finding the cummulicative frecuency (c.f) Where ($n = \Sigma f$).

e.g)

x	1	2	3	4	5
f	3	8	7	2	5
c.f	3	11	18	20	25
	0-3	4-11	12-18 (13)	19-20	21-25

Sum 25

$\frac{n}{2} = \frac{25}{2} = 12.5 \longrightarrow 13^{th}$ value

$Q_2 = 3$

e.g)

x	10	20	30	40
f	2	8	7	3
c.f	2	10	17	20
	0-2	3-10	11-17	18-20

sum 20

$\frac{n}{2} = \frac{20}{2} = \frac{10^{th} + 11^{th}}{2}$

$Q_2 = \frac{20+30}{2} = 25$

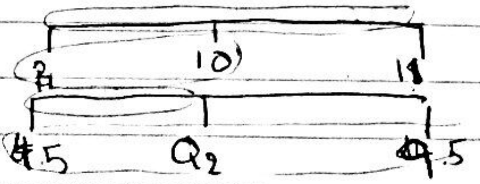
[3] for Grouped frequency distribution :-
Q_2: $(\frac{n}{2})^{th}$ value after finding the (c.f) and the upper real boundaries (U.R.B) then we use the linear inter polation method where ($n = \Sigma f$)

e.g)

	-.5	4.5	9.5	14.5	19.5
I	0-4	5-9	10-14	15-19	Sum
f	3	8	7	2	20
C.f	3	11	18	20	
U.R.B	4.5	9.5	14.5	19.5	

U.R.B:
١) الفرق بين الحدود للفئتين مقسومتين على 2
٢) نجمع نفس نهاية كل فئة .

$\frac{n}{2} = \frac{20}{2} = 10^{th}$ value ← في هذه الحالة لا نأخذ الحدين المتتاليين



$\frac{7}{8} = \frac{Q_2 - 4.5}{5}$

$Q_2 = 8.87$

9

e.g)

I	$0 \le x < 6$	$6 \le x < 12$	$12 \le x < 18$	$18 \le x < 24$	Sum
f	2	8	7	3	20
c.f	2	10	17	20	
U.R.B	6	12	18	24	

$\frac{n}{2} = \frac{20}{2} = 10^{th}$ value

$Q_2 = 12$

e.g)

I	0 - 1	1.2 - 2.2	2.4 - 3.4	3.5 - 3.6	3.5 < 4 sum
f	3	7	8	4	22
c.f	3	10	18	22	
U.R.B	1.1	2.3	3.5	3.7	

$\frac{n}{2} = \frac{22}{2} = 11^{th}$ value

$$\frac{11-10}{18-10} = \frac{Q_2 - 2.3}{3.5 - 2.3}$$

$Q_2 = 2.45$

e.g) For the following ordered set up data :-

a , 2 , 3 , 5 , b , 9 , 10 , c

The mean is 7

The mode is 2

The median is 6 find a, b, c

The mode is 2 $\rightarrow$ $a = 2$

The median is 6 $\rightarrow \frac{b+5}{2} = 6$

$12 = b + 5 \Rightarrow b = 7$

The mean is 7

$$\frac{2+2+3+5+7+9+10+c}{8} = 7$$

$c = 18$

(10)

Measures of spread مقاييس التشتت

[1] The range :-

Range = Max (U.R.B) - Min (L.R.B)

e.g) Find the range for :-

1) 2, 7, 5, 11, 3

نرتب ① 2, ~~3~~, 5, 7, 11 , ② Range = Max - Min = 11 - 2 = 9

2)

S	L		
2	2	3	4
3	2	2	3
4	3	4	5

Rang = Max - Min = 45 - 22 = 23

3)

X	1	2	3	4	5
F	1000	2000	500	40	50

Rang = 5 - 1 = 4

(Frequency) لم لا نهتم بالقيم

e.g)

I	0 - 4	5 - 9	10 - 14	15 - 19
F	3	8	7	2
U.R.B	4.5	9.5	14.5	19.5

-.5

Rang = 19.5 + .5 = 20

e.g) if $a<b<c$
The median is 7
The range is 8
The mean is 7
find a, b, c

sol.) median $\boxed{b=7}$

$c-a=8 \quad \ldots (1)$ $\qquad$ $c-a=8$

$\frac{a+7+c}{3}=7$ $\qquad$ $c+a=14$

$\qquad\qquad\qquad\quad 2c=22$

$a+c=14 \quad \ldots (2)$ $\qquad$ $\boxed{c=11}$

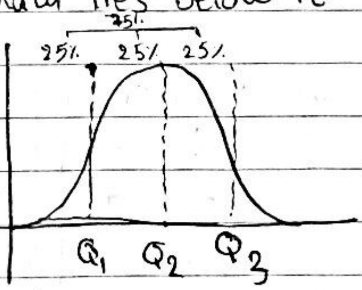
$11-a=8 \rightarrow \boxed{a=3}$

[2] The inter-quartile range : (IQR)

$IQR = Q_3 - Q_1$; where

Q_1: lower quartile (25%) of the data lies below it

Q_3: Upper quartile (75%) of the data lies below it



Note:- We can find Q_1 and Q_3 as Q_2 but we use $Q_1: \left(\frac{n}{4}\right)^{th}$ value

$Q_3: \left(\frac{3n}{4}\right)^{th}$ value

e.g) find IQR for

(1) 2, 7, 5, 11, 5, 8, 10

رتب (1) 2, 5, 5, 7, 8, 10, 11

$Q_1: \frac{n}{4} = \frac{7}{4} = 1.75 \rightarrow 2^{th}$ value

$Q_1 = 5$

$Q_3 = \frac{3n}{4} = \frac{3 \times 7}{4} = 5.25 \rightarrow 6^{th}$

$Q_3 = 10$

$IQR = 10 - 5 = 5$

e.g)

S	L
1	1 1 2 2
2	1 2 2 3
3	1 1 2 3

key
1|2 means 12

$n = 12$

$Q_1 : \frac{n}{4} = \frac{12}{4} = \frac{3^{rd} + 4^{th}}{2}$

$Q_1 = \frac{12+12}{2} = 12$

$Q_3 : \frac{3n}{4} = \frac{3 \times 12}{4} = \frac{9^{th} + 10^{th}}{2}$

$Q_3 = \frac{31+31}{2} = 31$

$IQR = Q_3 - Q_1 = 31 - 12 = 19$

e.g)

x	1	2	3	4	5	sum
f	3	8	7	2	5	25
cf	3	11	18	20	25	
	0-3	4-11	12-18	19-20	21-25	

$Q_1 : \frac{n}{4} = \frac{25}{4} = 6.25 \rightarrow 7^{th}$ value

$Q_1 = 2$

$Q_3 = \frac{3n}{4} = \frac{3 \times 25}{4} = 18.75 \rightarrow 19^{th}$ value

$Q_3 = 4$

$IQR = Q_3 - Q_1 = 4 - 2 = 2$

e.g)

I	0-4	5-9	10-14	15-19	sum
f	3	8	7	2	20
cf	3	11	18	20	
UB	4.5	9.5	14.5	19.5	

$Q_1 : \frac{n}{4} = \frac{20}{4} = 5^{th}$

$Q_3 = \frac{3n}{4} = 15$

3 5 11

4.5 Q_1 9.5

11 15 18

9.5 Q_3 14.5

$\frac{5-3}{11-3} = \frac{Q_1 - 4.5}{9.5-4.5}$

$Q_1 = 5.75$

$\frac{4}{7} = \frac{Q_3 - 9.5}{5}$

$Q_3 = 12.35$

$IQR = 6.6$

KoshinColor

3) The inter-percentile range:-

The r^{th} percentile is defined by:

$P_r = \frac{r}{100} \times n$

* r% of the data below pr.
* the mth% to nth% interpercentile range:

$P_n - P_m$

e.g) find the 35% to 65% interpercentile for:

1) 2, 3, 5, 7, 7, 8, 8, 9, 10, 12 لازم نتأكد إنهم مرتبين

$P_{35} = \frac{35}{100} \times 10$ $P_{35} = 3.5 \rightarrow 4^{th}$ value $P_{35} = 7$

$P_{65} = \frac{65}{100} \times 10$ $P_{65} = 6.5 \rightarrow 7^{th}$ value $P_{65} = 8$

$P_{65} - P_{35} = 8 - 7 = 1$

2)

x	f	c.f	
1	3	3	0 - 3
2	8	11	4 - 11
3	7	18	12 - 18
4	2	20	19 - 20
5	5	25	21 - 25

sa. sum 25

$P_{35} = \frac{35}{100} \times 25 = 8.75 \rightarrow 9^{th}$ value

$P_{35} = 2$

$P_{65} = \frac{65}{100} \times 25 = 16.25 \rightarrow 17^{th}$ value

$P_{65} = 3$

$P_{65} - P_{35} = 3 - 2 = 1$

3)

I	F	C.F	U.R.B
$0 \leq x < 6$	2	2	6
$6 \leq x < 12$	8	10^{7}	$\overset{P_{35}}{12}$
$12 \leq x < 18$	7	17^{13}	$\overset{P_{65}}{18}$
$18 \leq x < 24$	3	20	24
Sum	20		

$P_{35} = \frac{35}{100} \times 20 = 7^{th}$

2 — 7 — 10

6 — P_{35} — 12

$\frac{5}{8} = \frac{P_{35} - 6}{6}$

$P_{35} = 9.75$

$P_{65} = \frac{65}{100} \times 20 = 13^{th}$

10 — 13 — 17

12 — P_{65} — 18

$\frac{3}{7} = \frac{P_{65} - 12}{6}$

$P_{65} = 14.57$

$P_{65} - P_{35} = 4.82$

$\boxed{4}$ Standard deviation (s) and Variance (s^2)

(i) For row data

$$S^2 = \frac{\Sigma (x - \bar{x})^2}{n-1} \quad \text{OR} \quad S^2 = \frac{\Sigma x^2}{(n-1)} - \frac{(\Sigma x)^2}{n(n-1)}$$

e.g.) find S and S^2

(1) 2, 7, 5, 11, 5 $\rightarrow$ X

$\bar{x} = \frac{30}{5} = 6$

$x - \bar{x}$	-4	1	-1	5	-1	Sum
$(x - \bar{x})^2$	16	1	1	25	1	44

$\Sigma (x - \bar{x}) = 0$

$S^2 = \frac{44}{4} = 11$

$S = \sqrt{11}$

OR

						sum
x	2	7	5	11	5	30
x^2	4	49	25	121	25	224

$$S^2 = \frac{\sum x^2}{n-1} - \frac{(\sum x)^2}{n(n-1)}$$

$$S^2 = \frac{224}{4} - \frac{900}{5(4)} = 11$$

$$S = \sqrt{11}$$

E.g) If the mean mark of (10) boys is (12) and the standard deviation is (2), and the mean mark of (12) girls is (10) and the standard deviation is (3); Find the mean Mark and the standard deviation of the students all together?

Boys

$n = 10$

$\bar{x} = 12$

$S = 2 \longrightarrow S^2 = 4$

$\sum x = 120$

$$S^2 = \frac{\sum x^2}{n-1} - \frac{(\sum x)^2}{n(n-1)}$$

$$4 = \frac{\sum x^2}{9} - \frac{(120)^2}{10(9)}$$

$$\boxed{\sum x^2 - 1476}$$

Girls

$n = 12$

$\bar{x} = 10$

$S = 3 \longrightarrow S^2 = 9$

$\sum x = 120$

$$S^2 = \frac{\sum x^2}{n-1} - \frac{(\sum x)^2}{n(n-1)}$$

$$9 = \frac{\sum x^2}{11} - \frac{(120)^2}{12(11)}$$

$$\boxed{\sum x^2 = 1299}$$

All together

$n = 22$

$\sum x = 240$

$\sum x^2 = 2775$

the new $\bar{x} = \frac{240}{22} = 10.9$

the new: $S^2 = \frac{\sum x^2}{n-1} - \frac{(\sum x)^2}{n(n-1)}$

$$= \frac{2775}{21} - \frac{(240)^2}{22(21)}$$

$S^2 = 7.45$

~~$S = 1.029$~~

$S = 2.73$

e.g) if the mean mark of (11) students is (12) and the standard deviation. is (2)..
A Student with mark (15) joined the class, Find the mean and the new standard deviation.

Sol.) $n = 11$
$\bar{x} = 12$
$\sum x = 132$
$S = 2$

the new:
$n = 12$
$\bar{x} = ??$
$\sum x = 132 + 15 = 147$
$\sum x^2 = 1849$
new $\bar{x} = \frac{135}{2} = 11.25$

$$S^2 = \frac{\sum x^2}{n-1} - \frac{(\sum x)^2}{n(n-1)}$$

$$4 = \frac{\sum x^2}{10} - \frac{(132)^2}{11(10)}$$

$$\sum x^2 = 1624$$

$$\text{new } \sum x^2 = 1624 + (15)^2 = 1849.$$

new s^2:-

$$S^2 = \frac{\sum x^2}{n-1} - \frac{(\sum x)^2}{n(n-1)}$$

$$= \frac{1849}{11} - \frac{(147)^2}{12(11)}$$

$$S^2 = 9.7$$

$$S = 3.11.$$

ii) for grouped and frequency distributions:-

$$S^2 = \frac{\sum f(x - \bar{x})}{\sum f - 1} \quad \text{OR} \quad S^2 = \frac{\sum f x^2}{\sum f - 1} - \frac{(\sum f x)^2}{\sum f(\sum f - 1)}$$

e.g)

X	1	2	3	4	Sum
F	3	8	7	2	20
x^2	1	4	9	16	
fx	3	16	21	8	48
fx^2	3	32	63	32	130

$$S^2 = \frac{\sum f x^2}{\sum f_x - 1} - \frac{(\sum f x)^2}{\sum f(\sum f_x - 1)} \rightarrow \frac{130}{19} - \frac{(48)^2}{20(19)}$$

$$S^2 = 0.77$$

$$S = 0.88$$

e.g)

I	0-4	5-9	10-14	15-19	Sum
f	3	8	7	2	20
x	2	7	12	17	
x^2	4	49	144	289	
fx	6	56	84	34	180
fx^2	12	392	1008	578	1990

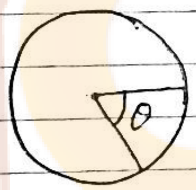
$$S^2 = \frac{1990}{19} - \frac{(180)^2}{20(19)}$$

$$= 19.47$$

$$S = 4.41$$

*** Graphical representation

[i] Pie Chart



$$\theta = \frac{f}{\Sigma f} * 360$$

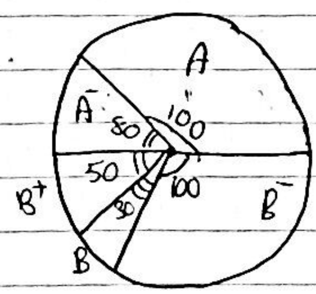
$$f = \frac{\theta}{360} \Sigma f$$

e.g)

A	A^-	B^+	B	B^-	Sum
10	8	5	3	10	36

$$A \rightarrow \theta = \frac{10}{36} * 360 = 100^\circ$$

$$A^- \rightarrow \theta = \frac{8}{36} * 360 = 80^\circ$$

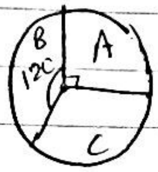


$$B^+ \rightarrow \theta = \frac{5}{36} * 360 = 50^\circ$$

$$B \rightarrow \theta = \frac{3}{36} * 360 = 30^\circ$$

$$B^- \rightarrow \theta = \frac{10}{36} * 360 = 100^\circ$$

e.g) $n = 18$ find A, B, C



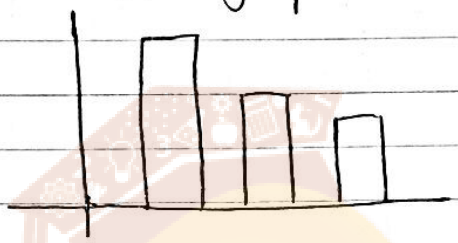
$$A = \frac{90}{360} * 18 = 4.5$$

$$B = \frac{120}{360} * 18 = 6$$

$$C = \frac{150}{360} * 18 = 7.5$$

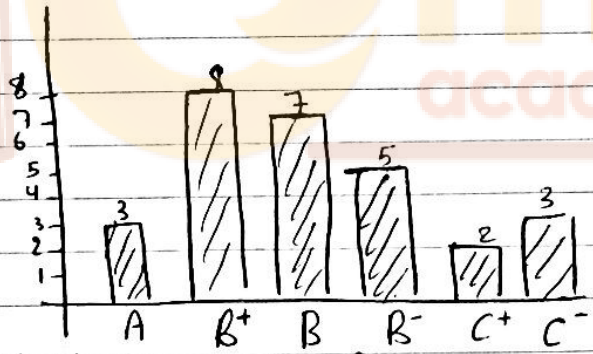
A	B	C
4.5	6	7.5

[2] Bar graph: Discrete or qualitative:-

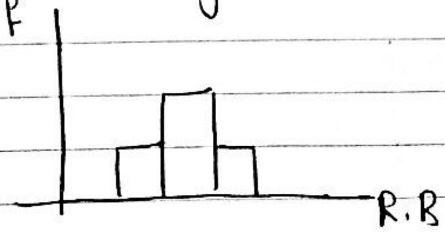


e.g)

A	B^+	B	B^-	C^+	C^-
3	8	7	5	2	3



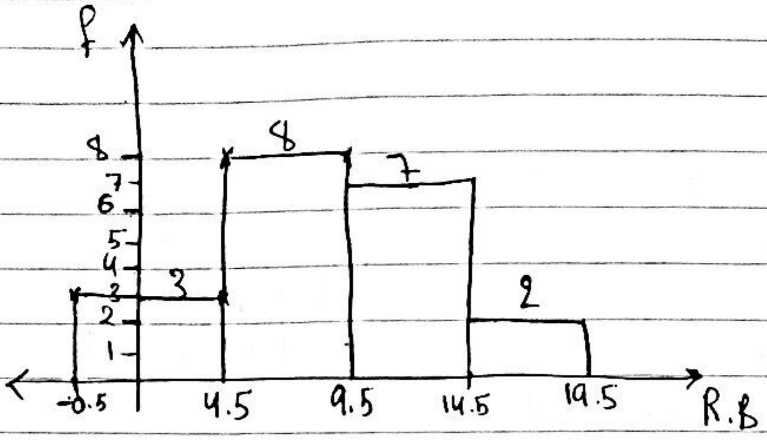
[3] Histogram:- For continuous data:-



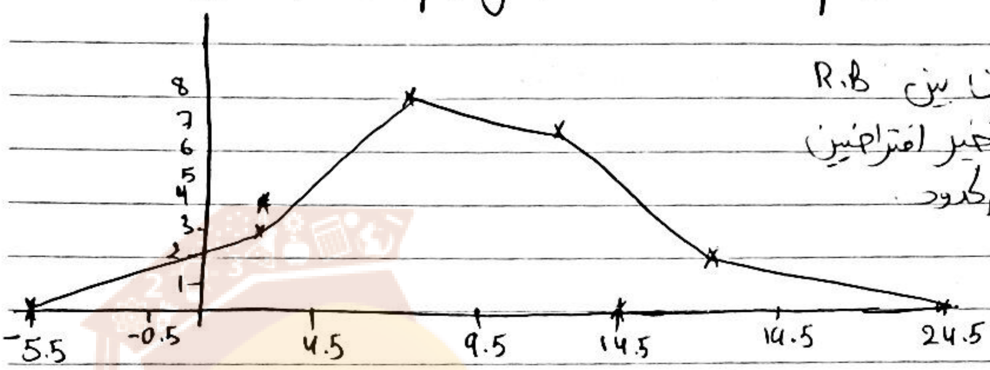
e.g) draw histogram for

-0.5 4.5 9.5 14.5 19.5

I	0-4	5-9	10-14	15-19	Sum
F	3	8	7	2	20

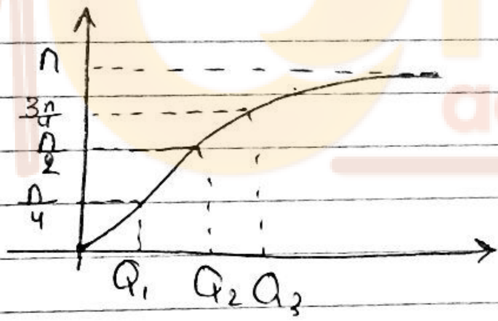


Note: the paliyon for the previous example is :-



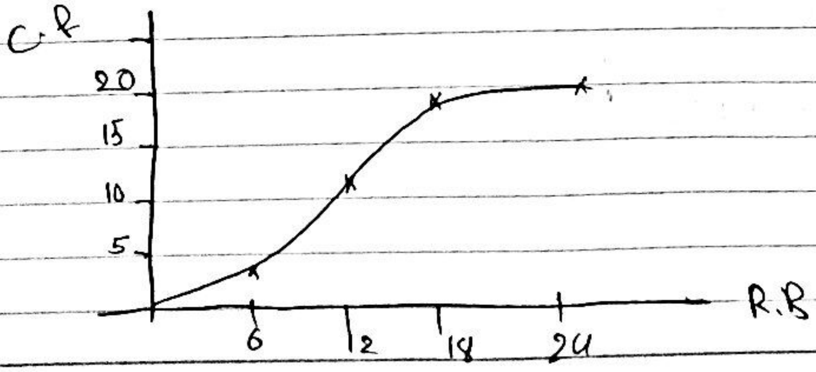
نأخذ نقطة المنتصف بين R.B
وتعين حد أول و أخير افتراضيين
حسب الفرق بين الحدود

C.f curve :-

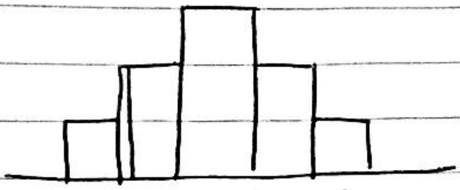


e.g) draw c.f curve for:-

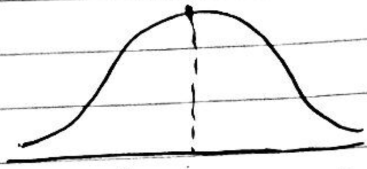
	$0\leq x\leq 6$	$6\leq x\leq 12$	$12\leq x\leq 18$	$18\leq x\leq 24$	sum
f	3	8	7	2	20
C.f	3	11	18	20	
U.R.B	6	12	18	24	



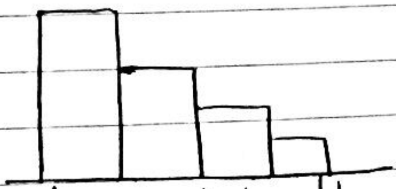
Skewness :-



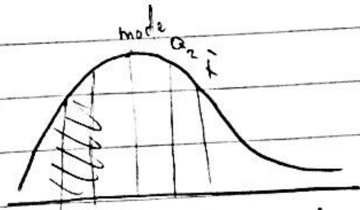
Symmetrical



mode = mean = Q_2

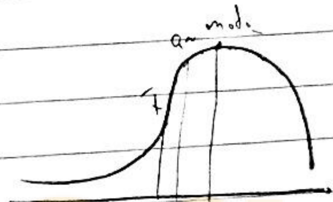


skewed to the right
(positively skewed)



$\bar{X} > Q_2 >$ mode

skewed to the left
negatively skewed



mode $> Q_2 > \bar{X}$

e.g) $\bar{X} = 5$
mode = 7
skewed to the left

e.g) $\bar{X} = 7$
mode = 5
skewed to the right

e.g) mean = 5
mode = 5
median = 5
symmetrical

Comparing 2 collections:-

[1] The Z score:

$$Z = \frac{X - \bar{X}}{S}$$

e.g)

	Section I	section II
$\bar{x}$	60	70
S	3	5
x	65	68

The Z score for X_{65} is

$$Z = \frac{65 - 60}{3} = 1.67$$

The Z score for X_{68} is:-

$$Z = \frac{68 - 70}{5} = -0.4$$

Z score for 65 > Z score for 68

∴ the mark 65 in section I is better than the mark 68 in section II

[2] Cofficient of variation (c.v)

$$C.V = \frac{S}{\bar{X}} * 100\%$$

e.g)

	Section I	Section II
$\bar{x}$	60	70
S	4.5	5

C.V for section I $= \frac{4.5}{60} * 100\%$

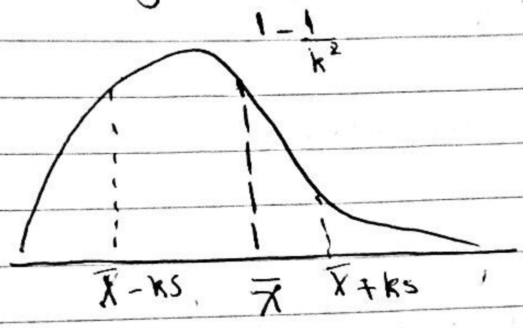
$= 7.5\%$

C.V for section II $= \frac{5}{70} * 100\%$

$= 7.14\%$

The variability of section I is higher than the variability of section II

* Chebyshev's theorem:



For any $k>1$ the proportion of data that lies k standard deviation about the mean is at least $1-\frac{1}{k^2}$

$$p(\bar{x}-ks < x < \bar{x}+ks) \geq 1-\frac{1}{k^2}$$

e.g) $\bar{x}=5$, $s=2$

what is the interval that at least 75% of the data lies below it?

sol). $1-\frac{1}{k^2}=0.75$

$1-\frac{1}{4}=\frac{1}{k^2} \Rightarrow k=2$

$\bar{x}+ks=5+4=9$

$\bar{x}-ks=5-4=1$

at least 75% of the data lies ~~below~~ in $(1,9)$

e.g) $\bar{x}=5$, $s=2$

Find the percentage of data that lies within 2 standard deviations about the mean?

$\bar{x}+ks=5+4=9$

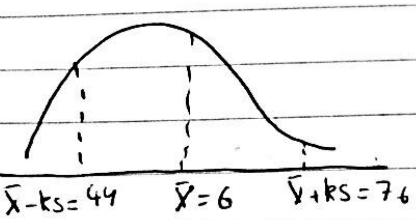
$\bar{x}-ks=5-4=1$

$=1-\frac{1}{k^2}=1-\frac{1}{4}=0.75$

75% lies in $(1,9)$

e.g) suppose that $\bar{x} = 60$, $S = 8$ and the no. of observations ~~are between 44 an 76~~ is $n = 100$:

[A] at least how many observations are between ~~55~~ 44 and 76:



$\bar{X} - ks = 44$

$60 - 8k = 44$

$\boxed{k = 2}$

$1 - \frac{1}{k^2} = 1 - \frac{1}{4} = 0.75$

$0.75 * 100\% = 75\%$

* at least 75% of observations are between (44 , 76)

[B] At most, how many observations are ~~less~~ more than 36 and ~~more~~ less than 84 ?

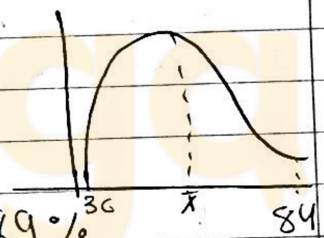
$\bar{X} + ks = 84$

$60 + 8k = 84$

$\boxed{k = 3}$

$1 - \frac{1}{k^2} = 1 - \frac{1}{9} = .89$

$0.89 * 100\% = 89\%$



at least 89% of observations are less than 84

[C] find an interval that contains at least 60% of the observations?

$1 - \frac{1}{k^2} = 0.6$

$\frac{1}{k^2} = \frac{1}{k^2} = \frac{2}{5}$

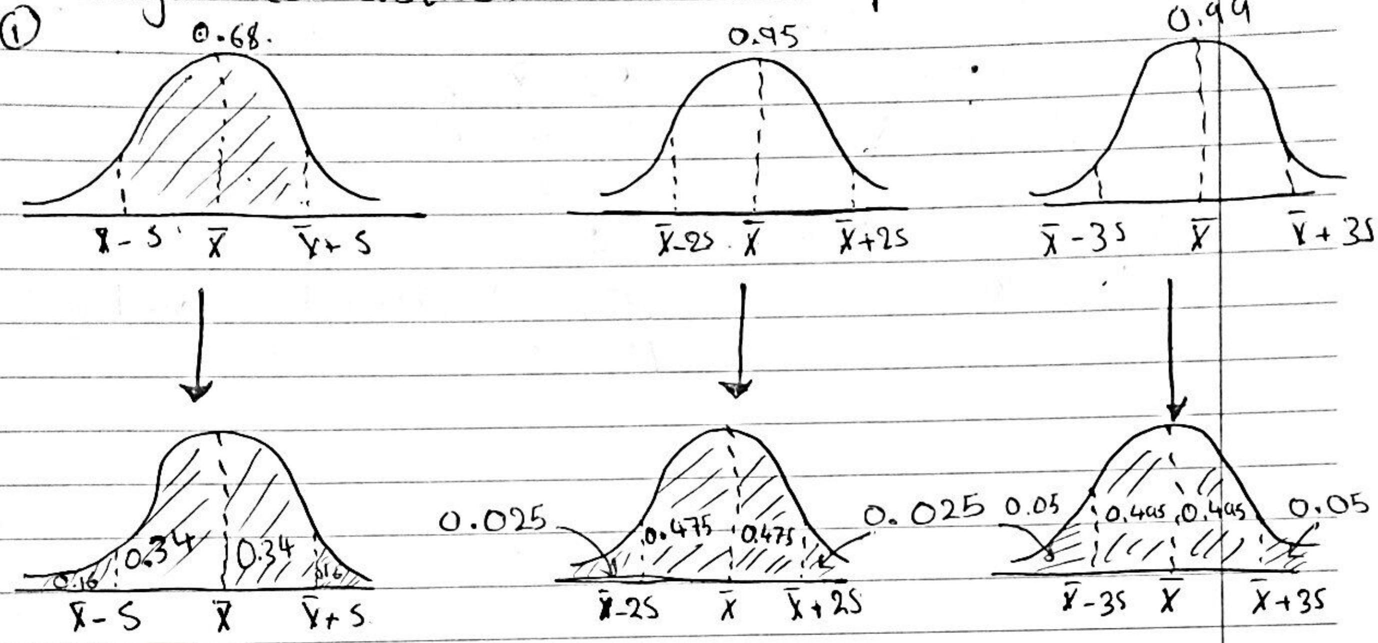
$k^2 = \frac{5}{2} \Rightarrow k = 1.58$

$\bar{x} + ks = 60 + (8)(1.58) = \underline{72.6}$

$\bar{x} - ks = ~~72~~ \; 60 - (8)(1.58) = \underline{47.3}$

For symetrical distribution (or bell-shape)

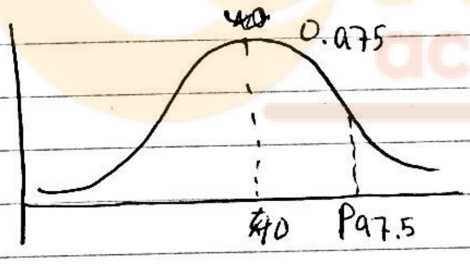
①



e.g) Suppose that the data is bell-shape :-

$\bar{x} = 40 \quad , S = 4$

ⓐ Find the 97.5 percentile?

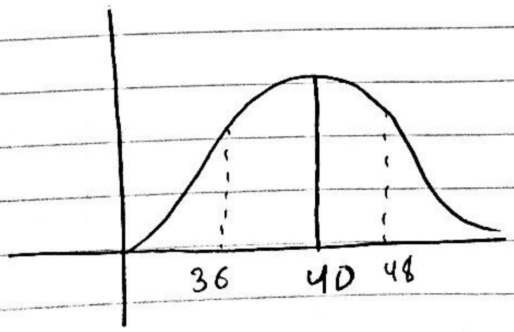


$P_{97.5} = \bar{x} + 2s$

$= 40 + 2(4)$

$P_{97.5} = 48$

ⓑ The percentage of the observation that are between 36 and 48



$48 = 40 + ks$

$8 = 4k$

$\boxed{k = 2}$

% of observation between 40 & 48 is 34% since k=2

$36 = 40 - ks$

$4 = 4k$

$\boxed{k = 1}$

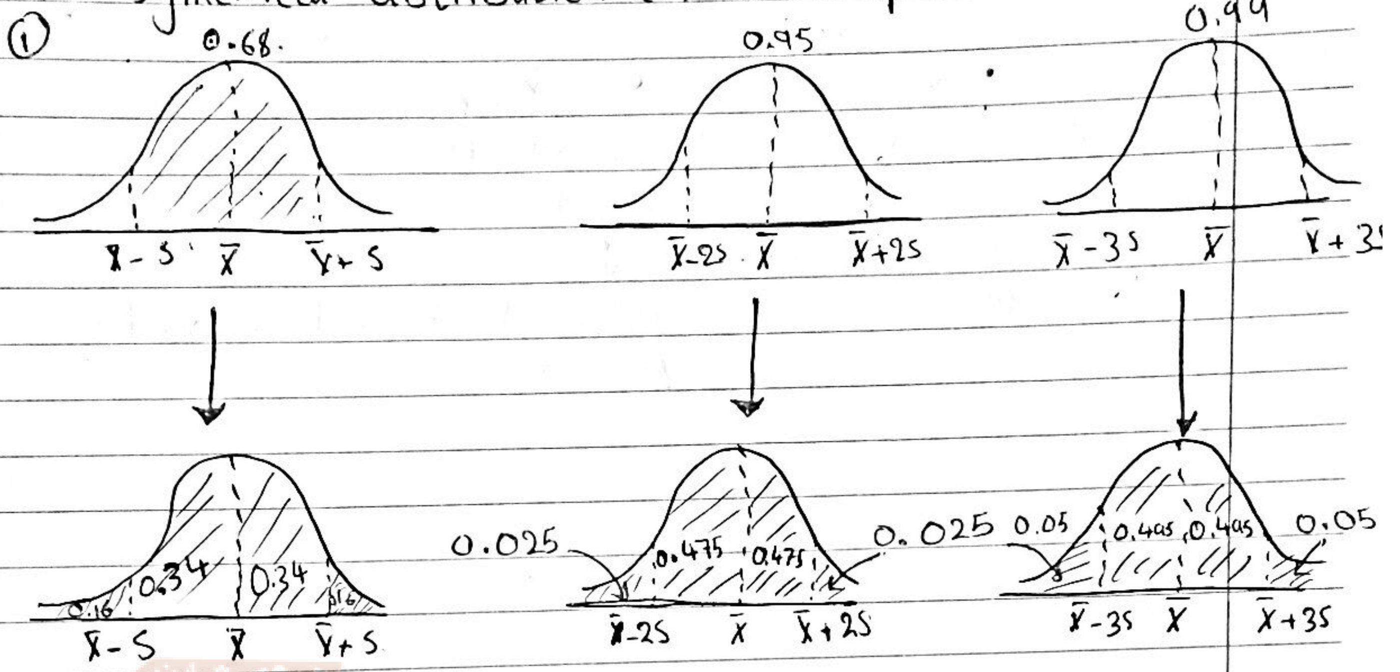
% of observations between 36 & 40 is 47.5 % since $k=1$

then the no. of observations between 36 & 48 is (34 + 47.5) %

= 81.5 %

For symetrical distribution (or bell-shape)

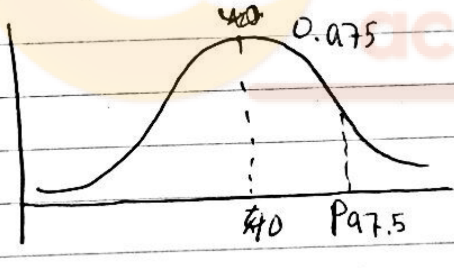
(1)



e.g) Suppose that the data is bell-shape :-

$\bar{X}=40$, $S=4$

(a) Find the 97.5 percentile?

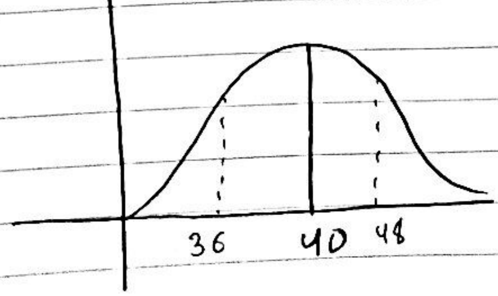


$P_{97.5}=\bar{X}+2s$

$=40+2(4)$

$P_{97.5}=48$

(b) The percentage of the observation that are between 36 and 48



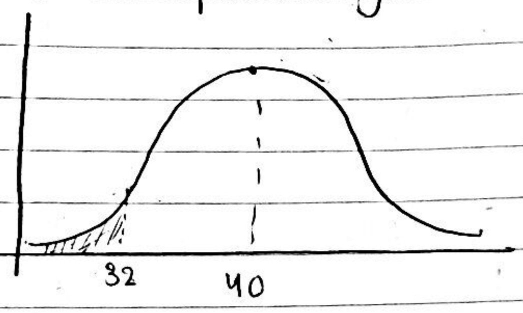
$48=40+ks$
$8=4k$
$\boxed{k=2}$
$\downarrow$
% of observation between 40 & 48 is 34% since k=2

$36=40-ks$
$4=4k$
$\boxed{k=1}$
$\downarrow$
% of observations between 36 & 40 is 47.5 % since $\boxed{k=1}$

then the no. of observations between 36 & 48 is (34+47.5) %
= 81.5 %

25

[c] the percentage of observation that are less than 32.



$32 = \bar{X} - ks$

$32 = 40 - 4k$

$\boxed{k = 2}$

$\Downarrow$

25% of observation are less than 32.

Coding (transformation) :-

If $y = ax + b$, then

$\boxed{1}$ $\bar{y} = a\bar{x} + b$

$\boxed{2}$ $S_y = S_x |a|$.

$\boxed{3}$ $var(y) = a^2 \, var(x)$

$\boxed{4}$ $Q_2(y) = (a) Q_2 + b$.

$\boxed{5}$ Range of $y = |a| *$ Rang of x

$\boxed{6}$ IQR of $y = |a| *$ IQR of x

$\boxed{7}$ P_r of $y = a(P_r \text{ of } x) + b \quad a > 0$

P_r of $y = a(P_{100-r} \text{ of } x) + b \quad a < 0$

e.g) If $y = 2 - 3x$, $\bar{y} = 50$, Q_1 of $x = -8$

$S_y = 10$ IQR of $x = 20$

Find: (a) $\bar{X}$ (b) Q_1 of y.

(a) $\bar{y} = 2 - 3\bar{X}$

$50 = 2 - 3\bar{X}$

$\bar{X} = -16$

(b) $IQR_y = |a|\ IQR_x$

$= |-3|\ 20$ ~~X~~

$IQR_y = 60$

$IQR\ of_x = Q_{3_x} - Q_{1_x}$

$20 = Q_{3x} + 8$

$\boxed{Q_{3x} = 12}$

$Q_1\ of\ y = a\ 2 - 3(Q_{3x})$

$= 2 - 3(12)$

$= -34$

e.g.) min of $x = 2$

max of $x = 20$ $\quad$ $P(65)$ of $x = 13$

Q_1 of $x = 5$

Q_2 of $x = 10$ $\quad$ $\boxed{\text{If } y = 3 - 2x}$

Q_3 of $x = 15$

$\bar{x} = 11$

$S_x = 3$

Find:- (a) Rang of y $\quad$ (d) P_{35} of y

(b) IQR of y $\quad$ (e) $\bar{y}$

(c) S_y and S^2_y $\quad$ (f) Q_1 of y & Q_3 of y.

Sol.)

(a) Range of $y = |-2| * (18)$

$= 36$

(b) $IQR_y = |-2|\ IQR\ of_x$

$= 2 * (15 - 5) = 20$

(c) $S_y = |a| S_x = 2(3) = 6$

$S^2_y = 36$

(d) $P_{35y} = b - a\,(P_{65x})$

$= 3 - 2(13)$

$= 3 - 26$

$= -23$

(e) $\bar{y} = 3 - 2\bar{x}$

$= 3 - 22$

$= -19$

(f) $Q_{1y} = 3 - 2(Q_{3x})$

$= 3 - 30 = -27$

$Q_{3y} = 3 - 2(Q_{1x})$

$= 3 - 2(5) = -7$

29

e.g)

X	1000	2000	3000	4000	5000	Sum
f	3	2	7	8	5	25

Use the Code:

$y = \dfrac{X - 3000}{1000}$ to find the mean & standard deviation

y	-2	-1	0	1	2	Sum
f	3	2	7	8	5	25
fy	-6	-2	0	8	10	10

$\bar{y} = \dfrac{\Sigma fy}{\Sigma f} = \dfrac{10}{25}$ $\leftrightarrow$ $\cancel{\bar{x} = \bar{y} \dfrac{-3000}{1000}}$

$\bar{y} = \dfrac{\bar{x} - 3000}{1000}$

$0.4 = \dfrac{\bar{x} - 3000}{1000}$

$400 + 3000 = \bar{x} \rightarrow \bar{x} = 3400.$

$S_y = \dfrac{\Sigma fy^2}{\Sigma f - 1} - \dfrac{(\Sigma fy)^2}{\Sigma f(\Sigma f - 1)}$

y^2	4	1	0	1	4	Sum
~~fy~~	~~36~~	~~4~~	~~0~~			
fy^2	12	2	0	8	20	42

$S_y^2 = \dfrac{42}{24} - \dfrac{(10)^2}{(25)(24)}$

$S_y^2 = 1.58$

$S = 1.26$

$S_y = |a| \times (S_x)$

$\cancel{= \dfrac{1}{1000}(1.26)}$

$1.26 = \dfrac{1}{1000}(S_x)$

$S(x) = 1260.$

Note:- Relative frequency:-

$$r.f = \frac{f}{\sum f}$$

e.g)

X	1	2	3	4	5	sum
f	3	8	7	2	5	25
rf	$\frac{3}{25}$	$\frac{8}{25}$	$\frac{7}{25}$	$\frac{2}{25}$	$\frac{5}{25}$	
	0.12	0.32	0.28	0.08	0.20	
rf * x	0.36	0.64	0.84	0.16	1	$\boxed{2.16}$

$\bar{X}$

Note:- $\bar{X} = \sum(r.f * X)$

e.g) For the following data:-

34, 56, 45, 34, 23, 12, 39,
23, 53, 66, 77, 88, 99, 90,
45, 56, 65, 78, 87, 98, 89,
23, 12, 21, 32, 35, 48

(a) Construct a frequency table of this data using 4 classes of equal length

(b) Construct a frequency table with unequal lengths but with the following cut points:-
12, 38, 46, 66, 91, 99

Sol.) a) Range = 99 - 12 = 87

class length = $\frac{87}{4}$ = 21.75 → 22 is the length of the class

I	12 - 33	34 - 56	57 - 79	80 - 99	sum
f	7	10	4	6	27

b)

I	11.5 - 38.5	38.5 - 46.5	46.5 - 66.5	66.5 - 91.5	91.5 - 99.5	Sum
f	11	2	6	6	2	27

Chapter 2 :- Probability :-

* Sample space (Ω) : الفضاء العيني

The set of all possible out comes :-

e.g) Find the sample space for the following experemints

① tossing a fair coin 1 time

Sol $\Omega = \{H, T\}$

$n(\Omega) = 2$

② tossing an unbiased coins 2 times:

$$\Omega = \left\{ \begin{matrix} (H,H), (H,T) \\ (T,T), (T,H) \end{matrix} \right\} \qquad n(\Omega) = 4 = 2^2$$

③ tossing a fair coin 3 times:

$$\Omega = \left\{ \begin{matrix} (H,H,H), (H,H,T), (H,T,H), (T,H,H) \\ (T,T,T), (T,T,H), (T,H,T), (H,T,T) \end{matrix} \right\}$$

$n(\Omega) = 8 = 2^3$

Note: when tossing a fair coin k times we have $n(\Omega) = 2^k$

④ Throwing a fair die 1 time:-

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

$n(\Omega) = 6$

⑤ Throwing 2 fair die 1 time:

Ω	1	2	3	4	5	6
1	(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
2	(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
3	(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
4	(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
5	(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
6	(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

$n(\Omega) = 36 = 6^2$

Note: when throwing a fair die k-times we have $n(\Omega) = 6^k$

★ 6) [1] [2] [3] [4]   2 cards are drawn :-

a) with replacement   مع ارجاع
b) without replacemen   = بدون

a) $\Omega = \{(1,2)(1,3)\cancel{(3,4)}(2,3)(2,4)\cancel{(3,4)}\}$ $\Omega = \{(1,1),(1,2),(1,3),(1,4),(2,1)(2,2),(2,3),(2,4)(3,1)(3,2)(3,3)(3,4)(4,1)(4,2)(4,3)(4,4)\}$

b) $\Omega = \{(1,2)\ (1,3)\ (1,4)\ (3,4)\ (2,4)\ (2,3)\}$ ~~(2,3) (2,4)~~

★ The event :-
Any subset of the sample space

* types of events :-

1) Simple : Consists of 1 element of (Ω)

2) Composite (combined) : Consists of more than 1 element of (Ω).

3) Impossible (null) :- Consists of no elements of Ω

4) Certain (sure) : Consists of all elements of Ω

e.g) when throwing a fair die 1 time. Define :
$\Omega = \{1,2,3,4,5,6\}$
A: Getting a no. divisible by 5   ★ عدد عناصر A / عدد عناصر Ω
$A = \{5\} \longrightarrow$ simple event $\rightarrow p(A) = \frac{n(A)}{n(\Omega)} = \frac{1}{6}$

B: Getting a no. which is a prime no. (عدد أولي)

$B = \{2,3,5\} \longrightarrow$ Combined event $\rightarrow p(B) = \frac{3}{6}$

C: Getting a no. which is more than 6

$C = \{\} = \phi \longrightarrow$ impossible event (null) $\rightarrow p(C) = 0$

33

D: Getting a no. less than 7 :-

D = {1, 2, 3, 4, 5, 6} $\longrightarrow$ Certain event

Note: the probability of an event (A) is given by:

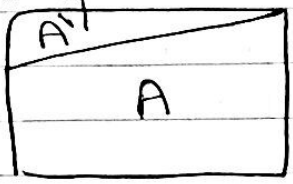
$p(A) = \frac{n(A)}{n(\Omega)}$

$\boxed{0 \leq p(A) \leq 1}$

$p(\Phi) = 0$

$p(\Omega) = 1$

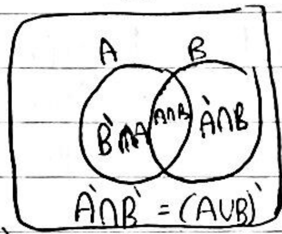
Rules of probability



[1] $p(A') = 1 - p(A)$

[2] 1) $p(A \cap B') = p(A) - p(A \cap B)$

2) $p(A' \cap B) = p(B) - p(A \cap B)$



[3] $p(A \cup B) = p(A) + p(B) - p(A \cap B)$

[4] De morgan's laws:-

i) $p(A' \cap B') = p(A \cup B)' = 1 - p(A \cup B)$

ii) $p(A' \cup B') = p(A \cap B)' = 1 - p(A \cap B)$

[5] A & B are said to be mutually exclusive (or disjoint) if: (a) $p(A \cap B) = 0$

(b) $p(A \cup B) = p(A) + p(B)$

[6] A & B are said to be independent if they don't influence each other.

Mathmatically, $p(A \cap B) = p(A) . p(B)$

[7] Conditional probability :-

$p(A \backslash B) = p\frac{(A \cap B)}{B}$

Given

Note:- If A & B are independant; then

$p(A/B) = p\frac{(A \cap B)}{B} = \frac{p(A) . \cancel{p(B)}}{\cancel{p(B)}} = p(A)$

e.g) If $p(A) = 0.8$ $p(B) = 0.7$ & $p(A \cap B) = 0.6$

Find:-

i) $p(A')$
ii) $p(B')$
iii) $p(A \cap B')$
iv) $p(A' \cap B)$
v) $p(A \cup B)$
vi) $p(A' \cup B')$
vii) $p(A' \cap B')$
xviii) $p(A \cup B')$
~~ix) p k p A~~
ix) $p(A' \cup B)$
x) $p(A' | B)$
xi) $p(A' | B')$
xii) $p(A | B')$
xiii) $p(A' | B')$
xiv) $p(A | B)'$

i) $p(A') = 1 - p(A)$
$= 1 - 0.8$
$= 0.2$

ii) $p(B') = 1 - p(B)$
$= 1 - 0.7$
$= 0.3$

iii) $p(A \cap B') = p(A) - p(A \cap B)$
$= 0.8 - 0.6 = 0.2$

iv) $p(A' \cap B) = p(B) - p(A \cap B)$
$= 0.7 - 0.6 = 0.1$

v) $p(A \cup B) = p(A) + p(B) - p(A \cap B)$
$= 0.8 + 0.7 - 0.6$
$= 0.9$

vi) $p(A' \cup B') = p(A \cap B)' = 1 - p(A \cap B)$
$= 1 - 0.6$
$= 0.4$

vii) $p(A' \cap B') = p(A \cup B)' = 1 - p(A \cup B)$
$= 1 - 0.9$
$= 0.1$

viii) $p(A \cup B') = p(A) + p(B') - p(A \cap B')$

$= 0.8 + 0.3 - 0.2$

$= 0.9$

ix) $p(A' \cup B) = p(A') + p(B) - p(A' \cap B)$

$= 0.2 + 0.7 - 0.1$

$= 0.8$

x) $p(A|B) = \dfrac{p(A \cap B)}{p(B)}$

$= \dfrac{0.6}{0.7} = \dfrac{6}{7}$

xi) $p(A'|B) = \dfrac{p(A' \cap B)}{p(B)}$

$= \dfrac{0.1}{0.7} = \dfrac{1}{7}$

xii) $p(A|B') = \dfrac{p(A \cap B')}{p(B')}$

$= \dfrac{0.2}{0.3} = \dfrac{2}{3}$

$\cancel{xiii)\ p(A'|B')' = 1 - p(A|B)}$

$\cancel{= 1 - \frac{6}{7}}$

$\cancel{= \frac{1}{7}}$

xiii) $p(A'|B') = \dfrac{p(A' \cap B')}{p(B')}$

$\dfrac{0.4}{0.3} = \cancel{6}\ \dfrac{1}{3}$

xiv) $p(A|B)' = 1 - p(A|B)$

$= 1 - \dfrac{6}{7}$

$= \dfrac{1}{7}$

e.g) if $p(A) = 0.6$, $p(B) = 0.5$ & $p(A \cup B) = 0.8$

Are A & B mutually exclusive, independent or neither?

sol) $p(A \cup B) = p(A) + p(B) - p(A \cap B)$

$P(A \cap B) = 0.6 + 0.5 - 0.8$
$P(A \cap B) = 0.3$

Now. $P(A) . P(B) = 0.3$

Since $p(A \cap B) = p(A) . p(B)$
then ~~p(A∩B) is~~ A & B are independent

e.g) if A & B are independent
$p(A) = 2p(B)$ & $p(A \cup B) = 0.8$ find $p(A)$

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

$P(A \cup B) = p(A) + p(B) - p(A \cap B)$

$0.8 = 2x + x - p(A) . p(B)$
$0.8 = 3x - 2x^2$

$2x^2 - 3x + 0.8 = 0$

$D = b^2 - 4ac$
$= 9 - 4(2)(0.8)$
$= 2.6$

$x = \dfrac{-b \pm \sqrt{D}}{2a} = \dfrac{3 \pm \sqrt{2.6}}{4}$

$\cancel{x = 1.15}$ $\quad$ $x = 0.347$

$p(B) = 0.347$
$P(A) = 0.693$

e.g) If A & B are independent then show that :-

(i) A & B' are independent
(ii) A' & B are independent
(iii) A' & B' are independent.

Sol.) given $(A \cap B) = p(A) . p(B)$

(i) $L.H.S = p(A \cap B')$
$= p(A) - p(A \cap B)$

$= p(A) - p(A) . p(B)$
$= p(A)\ (1 - p(B))$
$= p(A) . p(B') \Rightarrow$ Which show that A & B' is independent.

(ii) $(A' \cap B) = p(A') . p(B)$ ⇐ show that

$L.H.S = p(A' \cap B)$
$= p(B) - p(A \cap B)$

$= p(B) - p(A) . p(B) \Rightarrow = p(B)\ (1 - p(A)$

$= p(B) . p(A')$ #

(iii) $p(A' \cap B') = p(A') . p(B')$

$R.H.S \rightarrow p(A') . p(B') = (1 - p(A))\ (1 - p(B))$

$= 1 - p(B) - p(A) + p(A) p(B)$

$= 1 - (p(B) + p(A) - p(A \cap B))$

$= 1 - p(A \cup B) = p(A \cup B)'$

$= p(A' \cap B') \longrightarrow L.H.S$

#

e.g) Show that :-

i) $p(A' | B) = p(A|B)'$

L.H.S $\longrightarrow p(A'|B)$

$= \cancel{p}\, \dfrac{p(A' \cap B)}{p(B)}$

$= \dfrac{p(B) - p(A \cap B)}{p(B)}$

$= \dfrac{\cancel{p(B)}}{p(B)} - \dfrac{p(A \cap B)}{p(B)}$

$= 1 - \dfrac{p(A \cap B)}{p(B)}$

$= 1 - p(A|B)$

$= p(A|B)' =$ R.H.S #

e.g) $p[(A \cup B)|C] = p(A|C) + p(B|C) - p(A \cap B|C)$

L.H.S $\rightarrow p[A \cup B|C] = \dfrac{p[(A \cup B) \cap C]}{p(C)}$

$= \dfrac{p(A \cap C) \cup p(B \cap C)}{p(C)}$

$= \dfrac{p(A \cap C) + p(B \cap C) - p((A \cap C) \cap (B \cap C))}{p(C)}$

$= \dfrac{p(A \cap C)}{p(C)} + \dfrac{p(B \cap C)}{p(C)} - \dfrac{p[(A \cap C) \cap (B \cap C)]}{p(C)}$

$= p(A|C) + p(B|C) - \dfrac{p(A \cap B \cap C)}{p(C)}$

$= p(A|C) + p(B|C) - p(A \cap B|C)$ #

R.H.S

e.g) show that

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

L.H.S $\rightarrow P(A\cup B\cup C)$ let $D = B\cup C$

$$= P(A\cup D)$$

$$= P(A)+P(D)-P(A\cap D)$$

$$= P(A)+P(B\cup C)-\cancel{P(A\cup)}\;P(A\cap(B\cup C))$$

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

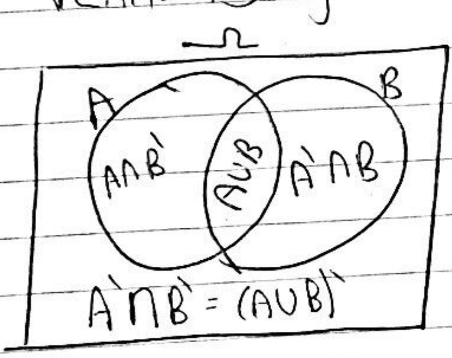
$$= P(A)+P(B)+P(C)-P(B\cap C)-P\left[P(A\cap B)+P(A\cap C)-P((A\cap B)\cap(A\cap C))\right]$$

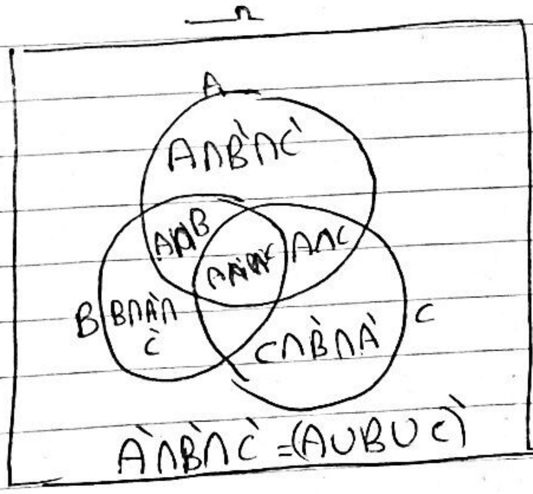
$$= P(A)+P(B)+P(C)-P(B\cap C)-P(A\cap B)-P(A\cap C)+P[(A\cap B)\cap(A\cap C)]$$

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

R.H.S. //

Venn-Diagram:





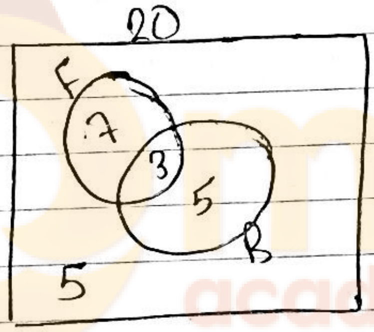
e.g) A class of 20 students.

10 play football
8 play basket ball
5 play neither

A student is selected Randomly. Find the probability that this student plays :-

i) football only
ii) exactly 1 sport
iii) at least 1 sport
iv) at most 1 sport
v) football given that the student plays basketball.
vi) football given that the student plays exactly 1 sport.

Sol.)



1) $p(F) = \frac{7}{20}$

2) $\frac{7+5}{20} = \frac{12}{20} = \frac{6}{10} = \frac{3}{5}$

3) $\frac{7+3+5}{20} = \frac{15}{20} = \frac{3}{4}$

4) $\frac{7+5+5}{20} = \frac{17}{20}$

5) $p(f|b) = \frac{3}{8}$

6) $p(f \mid 1 \text{ sport}) = \frac{7}{12}$

e.g) A class of 30 students

15 play football
12 play basketball
5 play vollyball only
7 play football & basket ball
6 play football & vallyball
3 play basketball & " only
3 play the 3 sports

A student is sellected at random, find the probability that this student:-

(a) plays exactly 1 sport

(b) plays exactly 2 sports

(c) plays at least 1 sports

(d) plays at ~~most~~ least 2 sport

(e) plays at ~~least~~ most 1 sport

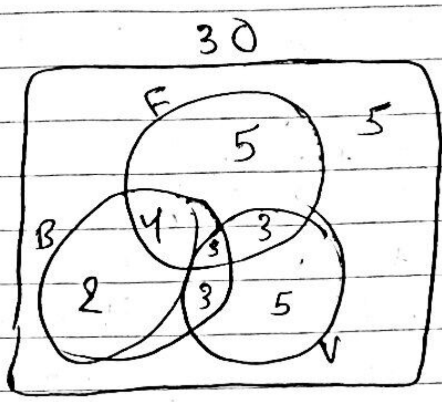
(f) plays at most 2 sports

(g) plays football given that the student plays basketball.

(h) plays exactly one sport given that the student plays at least one sport.

12

Sol.)



a) $\frac{5+5+2}{30} = \frac{12}{30} = \frac{2}{5}$

(b) $\frac{3+4+3}{30} = \frac{10}{30} = \frac{1}{3}$

(c) $\frac{25}{30} = \frac{5}{6}$

(d) $\frac{13}{30}$

(e) $\frac{5+5+2+5}{30} = \frac{17}{30}$

(f) $p(F|B) = \frac{7 \rightarrow \text{plays } f}{12 \rightarrow B \text{ عدد الي بلعبوا}}$

(h) $p(\text{exactly 1} / \text{at least 1}) = \frac{12 \rightarrow \text{exactly 1}}{25 \rightarrow \text{at least 1}}$

* Tree diagram:-

Usually use it when we select more than 1 item

e.g.)

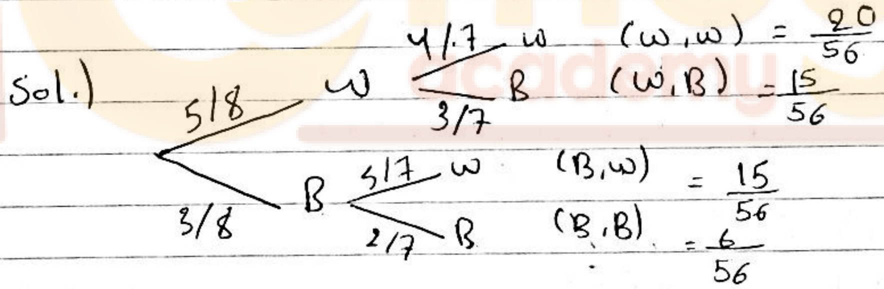
5	3
w	B

2 balls are selected random : (i) without replacement
(ii) with replacement.

Find the probability that:-

(a) the two balls are white
(b) the two balls are of the same color
(c) the two balls are of different color
(d) the first ball is white given that the 2^{nd} is black
(e) the 1^{st} ball is white given that the 2 balls are of the same color
(f) the seconed ball is white given that th 1^{st} is white
(g) the 2^{nd} is white given that the 2 balls are of different colors.

Sol.)



إذا لم يذكر لي في سؤال بإستبدال أو بدون نعتبر إنه بدون إستبدال

(a) $p(w,w) = \frac{20}{56}$

(b) $p(w,w) + p(B,B) = \frac{20}{56} + \frac{6}{56} = \frac{26}{56}$

(c) $p(w,B) + p(B,w) = \frac{15}{56} + \frac{15}{56} = \frac{30}{56}$

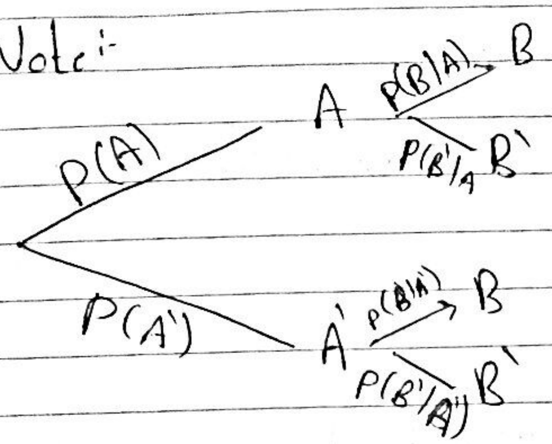
(d) $p(1^{st} w \mid 2^{nd} B) = \frac{p(w,B)}{p(2^{nd} B)} = \frac{p(w,B)}{p(w,B) + p(B,B)} = \frac{\frac{15}{56}}{\frac{15}{56} + \frac{6}{56}} = \frac{15}{21} = \frac{5}{7}$

(e) $p(1^{st} w \mid \text{same color}) = \frac{p(w,w)}{p(\text{same})} = \frac{p(w,w)}{p(w,w) + p(B,B)} = \frac{20}{26} = \frac{10}{13}$

(f) $p(2^{nd} w \mid 1^{st} w) = \frac{p(w,w)}{p(w,w) + p(B,w)}$ w,B $= \frac{20}{35} = \frac{4}{7}$

(g) $p(2^{nd} w \mid \text{diff}) = \frac{p(B,w)}{p(\text{diff})} = \frac{p(B,w)}{p(B,w) + p(w,B)} = \frac{15}{30} = \frac{1}{2}$

Note:-



e.g) $p(A) = 0.6$

$p(B|A)$ $\cancel{p(B|A)} = 0.3$

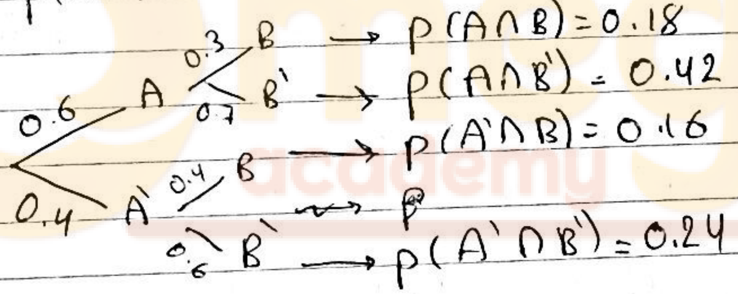
$p(B|A') = 0.4$

Find:- i) $P(A \cap B)$

ii) $p(B)$

iii) $p(A|B)$

Sol.)



i) $p(A \cap B) = 0.18$

ii) $p(B) = p(A \cap B) + p(A' \cap B)$ ← total probability

$= p(B|A)\,p(A) + p(B|A').\,p(A')$

$= 0.18 + 0.16$

$= 0.34$

iii) $p(A|B) = \dfrac{p(A \cap B)}{p(B)}$

$= \dfrac{0.18}{0.34} = \dfrac{9}{17}$

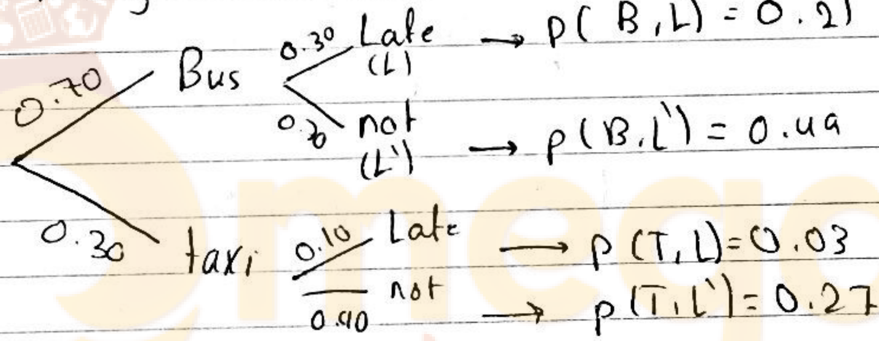
e.g) Emad comes to the university by either taxi or bus.

If the probability that Emad comes by ~~taxi~~ bus is 0.70 and by taxi is 0.30. If he comes by bus, then the prob. that he will be late is 0.3 & if he comes by taxi, then the prob. that he will be late is 0.10

a) Find the prob. that Emad will be late in a given school day.

b) If he is late, what is the prob. that Emad come by taxi that day

sol.)

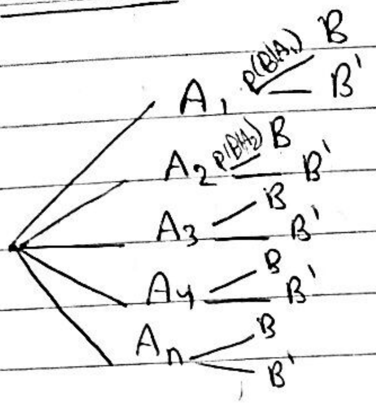


a) $p(L) = 0.21 + 0.03 = 0.24$

b) $p(T|L) = \dfrac{p(T, L)}{p(L)}$

$= \dfrac{0.03}{0.24} = \dfrac{3}{24} = \dfrac{1}{8}$

Note:



$$p(B) = p(B|A_1)p(A_1) + p(B|A_2)p(A_2) + \dots$$

$$p(B) = \sum_{k=1}^{n} p(B|A_k)\,p(A_k)$$

also, $$p(A_k|B) = \frac{p(A_k \cap B)}{p(B)}$$

$$= \frac{p(B|A_k).\,p(A_k)}{\sum_{k=1}^{n} p(B|A_k)\,p(A_k)}$$

$\downarrow$

Baye's theorem

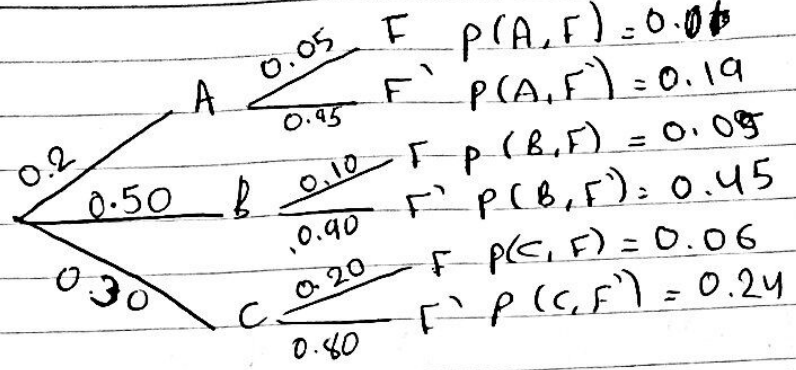
e.g) A Factory produces 3 types of Calculators (A, B, C) if it produces 20% of type A, 50% of type B & the rest one of type C. It is found that 5%, 10%, 20% are faulty calculators from A & B & C respectively. A Random calculator is selected at random.

a) Find the prob. that it is faulty

b) If it is found to be faulty, what is the probability that it is from type B.

Q7

Sol)



a) $p(F) = 0.01 + 0.05 + 0.06 = 0.12$

b) $p(B|F) = \dfrac{p(B,F)}{p(F)} = \dfrac{0.05}{0.12} = \dfrac{5}{12}$

* Counting Principles :-

$\boxed{1}$ The factorial (!) :-

$n! = n(n-1)(n-2)(n-3) \ldots\ldots (2)(1)$

when $n \in N$ (عدد طبيعي)

e.g) $5! = 5(4)(3)(2)(1) = 120$

$4! = 4(3)(2)(1) = 24$

$3! = 3(2)(1) = 6$

$2! = 2(1) = 2$

$1! = (1) = 1$

Note :- $\boxed{0! = 1}$

e.g) simplify :-

i) $\dfrac{10!}{8!} = \dfrac{10(9)\cancel{(8!)}}{\cancel{(8!)}} = 90$

ii) $\dfrac{5!}{7!} = \dfrac{\cancel{5!}}{7(6)\cancel{(5!)}} = \dfrac{1}{42}$

e.g) Solve:

i) $\dfrac{n!}{(n-2)!} = 6$

$\dfrac{n(n-1)\cancel{(n-2)!}}{\cancel{(n-2)!}} = 6 \rightarrow n^2 - n - 6 = 0$

$(n-3)(n+2) = 0$

$\boxed{n=3}$ $n = -2$ X

ii) $\dfrac{(n+2)!}{n} = 12$

$\dfrac{(n+2)(n+1)\cancel{(n)!}}{\cancel{n!}} = 12 \rightarrow n^2 + 3n + 2 = 12$

$n^2 + 3n - 10 = 0$

$(n+5)(n-2) = 0$

X $n = -5$ $\boxed{n=2}$

iii)

Note:- $n!$, the no. of ways to arrange n-objects in n-places

e.g) In how many ways can we arrange 3 students in 3 places in a row:-

Sol) | 3 | 2 | 1 |

$3 * 2 \times 1 = 3! = 6$

* The permination:-

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

e.g) Evaluate:-

i) ${}^{10}P_{2} = \frac{10!}{8!} = \frac{10(9)(\cancel{8!})}{\cancel{8!}} = 90$

ii) ${}^{10}P_{0} = \frac{10!}{10!} = \frac{10!\cancel{(4!)}}{10!\cancel{(4!)}} = 1$

iii) ${}^{10}P_{9} = \frac{10!}{1!} = 10!$

(iv) ${}^{10}P_{10} = \frac{10!}{0!} = \frac{10!}{1} = 10!$

Note:- $\boxed{{}^{n}P_{0} = 1}$

$\boxed{{}^{n}P_{n} = n! = {}^{n}P_{(n-1)}}$

$\boxed{{}^{n}P_{1} = n}$

e.g) in how many ways can we arrany 5 books in 3 places in a shelf.

e.g)

5	4	3

$(5)(4)(3) = 60$

$$= {}^{5}P_{3} = \frac{(5)(4)(3)\cancel{(2)(1)}}{\cancel{(2)(1)}} = 60$$

$\star$ ${}^nP_r$ is the no. of arrangements of ~~n-obje~~ n-objects in r-places

e.g) solve for n:-

i) ${}^nP_{n-2} = 3$

sol) $\frac{n!}{2!} = 3$

$n! = 6$

$n! = 3! \Rightarrow \boxed{n = 3}$

ii) ${}^nP_2 = 6 \Rightarrow \frac{n!}{(n-2)!} = 6$

$\frac{n(n-1)\cancel{(n-2)!}}{\cancel{(n-2)!}} = 6$

$n^2 - n - 6 = 0$

$\boxed{n = 3}$ $n = -2$ X

iii) The Combination:-

$${}^nC_r = \binom{n}{r} = \frac{n!}{r!(n-r)!}$$

Note:- $\binom{n}{r} = \frac{{}^nP_r}{r!}$

OR ${}^nP_r = r! \binom{n}{r}$

e.g) ${}^{10}C_2 = \frac{10!}{2!(8!)} = \frac{(10)(9)\cancel{(8!)}}{(2)\cancel{(8!)}} = \frac{90}{2} = 45$

e.g) $\binom{10}{0} = \frac{\cancel{10!}}{0!(\cancel{10!})} = 1$

(51)

e.g) $\binom{10}{10} = \frac{\cancel{10!}}{\cancel{10!}(0!)} = 1$

~~e.g) (1~~

Note $\boxed{\binom{n}{0} = \binom{n}{n} = 1}$

e.g) $\binom{10}{1} = \frac{10!}{1!(9!)} = \frac{10(\cancel{9!})}{1(\cancel{9!})} = 10$

e.g) $\binom{10}{1} = \frac{10!}{(1!)(9!)} = \frac{10(\cancel{9!})}{1(\cancel{9!})} = 10$

Note $\boxed{\binom{n}{1} = \binom{n}{n-1} = n}$

e.g) $\binom{1000}{999} = 1000$

e.g) $\binom{5}{3} = \frac{5!}{3!(2!)} = 10$

$2+3=5 \rightarrow \binom{5}{3} = \binom{5}{2}$

e.g) $\binom{5}{2} = \frac{5!}{2!(3!)} = 10$

Note: $\binom{n}{r} = \binom{n}{n-r}$

بإختصار إذا كان مجموع العددين اللي تحت = العدد اللي فوق .. إذن العددين اللي فوق متساوين والقوسين متساوين

Solve:-

1) $\binom{n}{2} = 6 \Rightarrow \frac{n!}{2!(n-2!)} = 6$

$\frac{n(n-1)\cancel{(n-2)!}}{2\cancel{(n-2)!}} = 6$

$n^2 - n - 12 = 0$

$(n-4)(n+3) = 0$

$\boxed{n=4}$ $\quad n=-3$ ✗

e.g) ii) $\binom{n}{n-2} = 6$

$$\frac{n!}{2!(n-2)!} = 6$$

$$\frac{n(n-1)\cancel{(n-2)!}}{2\cancel{(n-2)!}} = 6$$

$$n^2 - n - 12 = 0$$

$$n = -3 \times \quad \boxed{n = 4}$$

Note:

$\binom{n}{r}$: the no. of ways for selecting r-objects out of (n)-object at the same time.

e.g) In how many ways can we select 2 students from a class of 10 students?

$$\binom{10}{2} = \frac{10!}{2!(8!)} = 45 \text{ ways.}$$

e.g)

10	20
w	B

* إذا السؤال ما حدد سواء بإرجاع أو بدون إرجاع نعتبر إنه بدون إرجاع دائماً

6 balls are selected at random. Find the probability that:

i) 4 balls are white

ii) at least 1 ball is white

iii) at least 4 balls are white

iv) all balls are white

sol.)

i) $p(4w, 2B) = \frac{\binom{10}{4} \cdot \binom{20}{2}}{\binom{30}{6}} = 0.067$

ii) $p(\text{at least } 1w) = 1 - p(\text{no white})$

$$= 1 - p(0w, 6B)$$

$$= 1 - \frac{\binom{10}{0} \cdot \binom{20}{6}}{\binom{30}{6}} =$$

iii) $p(4w, 2B) + p(5w, 1B) + p(6w, 0B) =$

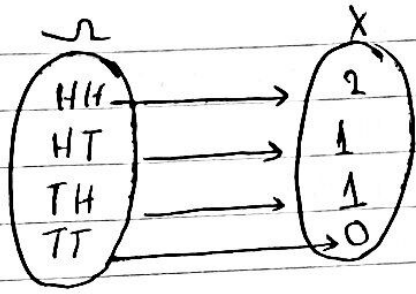
iv) $p(6w, 0B) =$

(53)

Random Variable :- a function from the sample space (Ω) to a set of real numbers.

e.g) when tossing an unbaised coin 2 times; Define the random variable (X) to be the no. of head obtained.

Sol) $\Omega = \begin{Bmatrix} H,H & H,T \\ T,T & T,H \end{Bmatrix}$



$X(H,H) = 2$

$X(HT), X(TH) = 1$

$X(TT) = 0$

Note:- If the range is acountable set:

like $\{x_1, x_2, x_3, \ldots x_n\}$

or $\{x_1, x_2, x_3, \ldots\}$

then the random variable X is called:- Discrete random variable, but if the range is an interval (or union of intervals), then the random variable is called continuous random variable

Discreat r.v :-

• probability density function (p.d.f)

a function $F(x) = p(X=x)$ is called a p.d.f. if :-

① $F(X) \geq 0$ for all x

② $\sum f(x) = 1$.

e.g) in the previous example:-

x	0	1	2
$p(X=x)$	$\frac{1}{4}$	$\frac{2}{4}$	$\frac{1}{4}$

$F(x)$.

since $p(x) \geq 0$ for all x and $\sum f(x) = 1$

then $F(x)$ is a p.d.f

e.g) Consider the following p.d.f

X	1	2	3	4
$p(X=x)$	0.2	a	$2a$	0.2

Find a?

$a \leq 0$

$0.2 + a + 2a + 0.2 = 1$

$0.4 + 3a = 1$

$3a = 0.6$

$\boxed{a = 0.2}$

e.g)

X	1	2	3	4
$p(x)$	0.3	0.2	0.1	0.4

Find (i) $p(X=1) = 0.3$

(ii) $p(X=1.5) = 0$

(iii) $p(X \leq 2) = p(X=1) + p(X=2)$
$= 0.3 + 0.2 = 0.5$

(iv) $p(1 < X \leq 4) = p(X=2) + p(X=3) + p(X=4)$
$= 0.2 + 0.1 + 0.4 = 0.7$

(v) $p(\underset{\text{given}}{X < 3 \mid X > 1}) = \dfrac{p(X<3)}{p(X>1)} = \dfrac{p(X=2)}{1 - p(\leq 1)} = \dfrac{0.2}{1-0.3} = \dfrac{0.2}{0.7} = \dfrac{2}{7}$

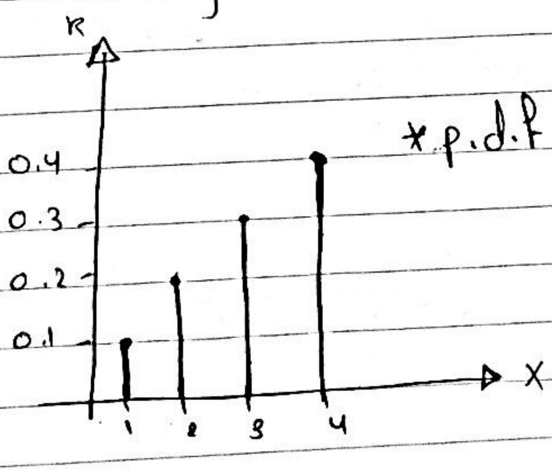
e.g) $F(X) = \begin{cases} k.x, & x = \{1,2,3,4\} \\ 0, & \end{cases}$ is a pdf Find the value of k.

X	1	2	3	4	Sum
$p(X=x)$	k	$2k$	$3k$	$4k$	1

$k + 2k + 3k + 4k = 1$

$10k = 1$

$k = \dfrac{1}{10} = 0.1$



*p.d.f

55

The comulative distribution Function (C.d.F)

e.g

x	1	2	3	4
$p(X=x)$	0.2	0.1	0.4	0.3

Find :-

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

(i) $F(1) = p(X \leq 1) = 0.2$

(ii) $F(2) = p(X \leq 2) = p(x=1) + p(x=2) = 0.2 + 0.1 = 0.3$

(iii) $F(3) = p(X \leq 3) = p(X=1) + p(X=2) + p(X=3) = 0.2 + 0.1 + 0.4 = 0.7$

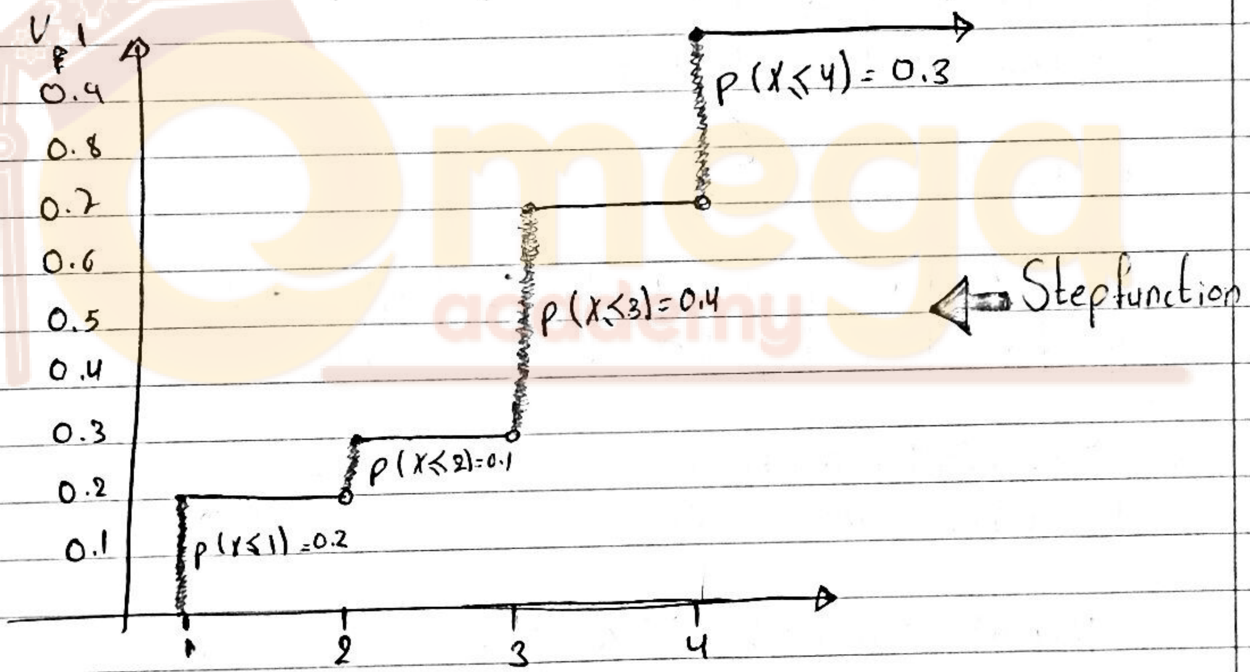
(iv) $F(4) = p(X \leq 4) = 0.2 + 0.1 + 0.4 + 0.3 = 1$

(v) $F(5) = p(X \leq 5) = 1$

(vi) $F(0) = p(X \leq 0) = 0$

(vii) $F(1.2) = p(X \leq 1.2) = p(X=1) = 0.2$

(viii) $F(1.7) = p(X \leq 1.7) = p(X=1) = 0.2$



$$F(x) = \begin{cases} 0, & X < 1 \\ 0.2, & 1 \leq X < 2 \\ 0.3, & 2 \leq X < 3 \\ 0.7, & 3 \leq X < 4 \\ 1, & X \geq 4. \end{cases}$$

notes:-

① $F(X < \text{smallest}) = 0$

② $F(X = \text{smallest}) = p(X = \text{smallest})$

③ $F(X \geq \text{largest}) = 1$

④ $F(X_n) - F(X_{n-1}) = p(X = x_n)$

* The expected value :-

$$\boxed{\mu \cdot E(X) = \sum_{x} x \, p(X=x)}$$

e.g) Find $E(X)$ for :-

X	1	2	3	4	sum
$p(X=x)$	0.4	0.2	0.3	0.1	1

Sol. : $E(x) = 1(0.4) + 2(0.2) + 3(0.3) + 4(0.1)$
$= 0.4 + 0.4 + 0.9 + 0.4$
$= 2.1$

Find $E(x)$

e.g)

X	1	2	3
$p(X=x)$	a	b	a

$E(x) = 2 \Rightarrow$ the mean.

e.g)

x	1	2	3	4
$p(X=x)$	0.2	a	b	0.4

If $E(x) = 2.8$ find a & b

pdf $\rightarrow$ $0.2 + 0.4 + a + b = 1$
$\boxed{a + b = 0.4}$ ... ①

$1(0.2) + 2a + 3b + 1.6 = 2.8$
$\boxed{2a + 3b = 1}$ ... ②

$-2a - 2b = -0.8$
$2a + 3b = 1$
$\boxed{b = 0.2}$

$a + 0.2 = 0.4$
$\boxed{a = 0.2}$

e.g) When throwing 2 pair die 2 times. Define the random variable S to be the sum of two numbers obtained. Fine the expected sum.

sol.)

S	2	3	4	5	6	7	8	9	10	11	12
$p(S=s)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

من الجدول ← عند الاحتمالات تبع رمي حجرين نرد .. الخط الأفقي (١) يدل على عدد الأحوال اللي مجموعها العدد

$E(X) = 7 \implies$ Symmetrical.

e.g) $p(X=x) = 0.1x \quad , \quad x \in \{1,2,3,4\}$

Find $E(x)$

x	1	2	3	4
$p(X=x)$	0.1	0.2	0.3	0.4

$E(x) = 0.1 + 0.4 + 0.9 + 1.6 = 3$

Properties of $E(X)$:-

i) $E(a) = a$

ii) $E(aX) = a\,E(X)$

iii) $E(X \mp Y) = E(X) \mp E(Y)$

iv) $E(g(X)) = \sum_x g(x) \cdot p(X=x)$

e.g)

x	1	2	3	4
$p(X=x)$	0.4	0.3	0.2	0.1

Find :- (i) $E(X) = 0.4 + 0.6 + 0.6 + 0.4 = 2$

(ii) $E(X^2) = (1)^2(0.4) + (2)^2(0.3) + (3)^2(0.2) + (4)^2(0.1) = 5$

(iii) $E(X^3) = (1)^3(0.4) + (2)^3(0.3) + (3)^3(0.2) + (4)^3(0.1) = 14.2$

(iv) $E\left(\frac{1}{X}\right) = \left(\frac{1}{1}\right)(0.4) + \frac{1}{2}(0.3) + \frac{1}{3}(0.2) + \frac{1}{4}(0.1) =$

e.g) If $E(X)=4$ then find:-

(i) $E(7)$ (iii) $E(E(X))=$

(iii) $E(3-2X)$ (iv) $E(\frac{X}{2}+1)$

(i) $E(7)=7$

(ii) $E(E(X))= E(4)=4$

(iii) $E(3-2X)= E(3)-2E(X)$

$= 3-8=-5$

(iv) $E(\frac{X}{2}+1)= \frac{1}{2}E(X)+E(1)$

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

* The variance :-

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

or

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

Thus, the Standard deviation is:-

sigma $\sigma_X=\sqrt{var(X)}$

e.g) Find Var(X) for :-

X	1	2	3	4
P(X=x)	0.3	0.4	0.1	0.2

Sol.) $E(X)= 1(0.3)+2(0.4)+3(0.1)+4(0.2)=2.20$

$E(X^2)=1(0.3)+4(0.4)+9(0.1)+16(0.2)=6$

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

$=6-(2.2)^2$

$=1.16$

* Properties of var(x) :

$\boxed{1}$ $Var(a) = 0$

$\boxed{2}$ $Var(aX) = a^2 \, Var(X)$

$\boxed{3}$ $Var(aX+b) = a^2 \, Var(X)$

$\boxed{4}$ $E(X^2) = Var(X) + (E(X))^2$

or $E(X^2) = \sigma_x^2 + \mu^2$

e.g) If $E(X) = 5$ & $Var(X) = 4$

Find : (i) $Var(3)$ (ii) $Var(E(X))$

(iii) $Var(Var(X))$ (iv) $E(Var(X))$

(V) $Var(2x+3)$ (Vi) $Var(3-2x)$

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

Sol.) (i) $Var(3) = 0$

(ii) $Var(E(X)) = Var(5) = 0$

(iii) $Var(Var(X)) = Var(4) = 0$

(iv) $E(Var(X)) = E(4) = 4$

(V) $Var(2X+3) = 4(Var(X)) = 4(4) = 16$

(Vi) $Var(3-2X) = 4(Var(X)) = 4(4) = 16$

(Vii) $E(X(X-2)) = E(X^2 - 2X) = E(X^2) - 2E(X)$

$= Var(X) + (E(X))^2 - 2E(X)$

$= 4 + 25 - 2(5)$

$= 29 - 10 = 19$

e.g) prove that :

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

L.H.S $\rightarrow$ $E(X-\mu^2)^2 = E(X^2 - 2\mu X + \mu^2)$

$= E(X^2) - 2E(\mu X) + E(\mu^2)$

$= E(X^2) - 2\mu E(X) + \mu^2$

$= E(X^2) - 2\mu(\mu) + \mu^2$

$= E(X^2) - \mu^2 \longrightarrow$ R.H.S

#

* Bivariate Distribution :-
joint distribution

[1] The joint p.d.f of X & Y (say $f(x,y)$) :

(i) $f(x,y) \geq 0$ for all x & y
(ii) $\sum_{y} \sum_{x} f(x,y) = 1$

e.g)

Y\X	1	2	3	
1	0.2	0.05	0.15	0.4
2	0.3	0.1	0.2	0.6
2	0.5	0.15	0.35	1

→ joint p.d.f of X & Y

i Find the marginal p.d.f of X & Y.

X	1	2
$P(X=x)$	0.4	0.6

→ marginal p.d.f of X

Y	1	2	3
$P(Y=y)$	0.5	0.15	0.35

e.g) ii) find $E(X)$, $E(X,Y)$, $E(Y)$

$$E(Y) = 1(0.5) + 2(0.15) + 3(0.35)$$
$$= 1.85$$

$$E(X) = 1(0.4) + 2(0.6)$$
$$= 0.4 + 1.2 = 1.6$$

$$E(X,Y) = (1)(1)(0.2) + (1)(2)(0.05) + (1)(3)(0.15) + (2)(1)(0.3) + (2)(2)(0.1)$$
$$+ 2(3)(0.2) = 2.85$$

The Covariance:-

$$Cov(X,Y) = E(X,Y) - E(X).E(Y).$$

Note:- $Cov(X,X) = Var(X)$.

e.g) In the previous example:-

$$Cov(x,y) = E(X,Y) - E(X).E(Y)$$
$$= 2.95 - (1.6)(1.85)$$
$$= -0.01$$

The correlation cofficient:-

$$\rho = \frac{Cov(X,Y)}{\sigma_X \sigma_Y} = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$$

e.g) In the previous example:

$$E(X^2) = (1)^2(0.4) + (2)^2(0.6) = 0.4 + 2.4 = 2.8$$
$$Var(X) = E(X^2) - (E(X))^2 = 2.8 - (1.6)^2 = 0.24.$$

$$\sigma_X = \sqrt{0.24} = 0.490$$

$$E(Y^2) = (1)^2(0.5) + (2)^2(0.15) + (3)^2(0.35) = 4.25$$

$$\cancel{\sigma_Y = \sqrt{4.25} = 0.90}.$$

$$Var(Y) = E(Y^2) - (E(Y))^2$$
$$= 4.25 - (1.85)^2 = 0.83$$

$$\sigma_Y = \sqrt{0.83} = 0.90$$

$$\rho = \frac{-0.01}{(0.49)(0.9)} = -0.022.$$

$$\boxed{\text{Note: } -1 \leq \rho \leq 1}$$

e.g

$X \backslash Y$	1	2	3	
1	0.3	0.1	0.2	0.6
2	0.1	0.2	0.1	0.4
	0.4	0.3	0.3	1

$$P(x|y) = \frac{f(x,y)}{f(y)} = \frac{P((X=x),(Y=y))}{P(Y=y)}$$

find: i) $f(X=1 \mid Y=1)$

x	1	2
$P(X=x)$	0.6	0.4

y	1	2	3
$P(Y=y)$	0.4	0.3	0.3

$$f(X=1 \mid Y=1) = \frac{P(X=1, Y=1)}{P(Y=1)} = \frac{0.3}{0.4} = 0.75$$

$$\text{(ii) } f(X=2 \mid Y=1) = \frac{P(X=2, Y=1)}{P(Y=1)} = \frac{0.1}{0.4} = 0.25$$

~~(iii)~~ Note:

X	1	2
	0.75	0.25

→ conditional p.d.f of $X|y=2$.

$$\text{(iii) } f(X=1 \mid Y=2) = \frac{0.1}{0.3} = \frac{1}{3}$$

$$\text{(iv) } f(X=2 \mid Y=2) = \frac{0.2}{0.3} = \frac{2}{3}$$

x	1	2
$P(X|y=2)$	$\frac{1}{3}$	$\frac{2}{3}$

→ conditional p.d.f of $X|y=2$.

$$\text{(v) } f(X=1 \mid Y=1) = \frac{0.3}{0.6} = \frac{3}{6} = 0.5$$

$$\text{(vi) } f(X=1 \mid y=2) = \frac{0.1}{0.6} = \frac{1}{6}$$

$$\text{(vii) } f(X=1 \mid y=3) = \frac{0.2}{0.6} = \frac{2}{6}$$

~~x~~ y	1	2	3
~~p(x|y)~~ $p(y|x=1)$	$\frac{3}{6}$	$\frac{1}{6}$	$\frac{2}{6}$

→ Conditional p.d.f of $f(y|X=1)$

63

viii) find the conditional p.d.f of $X|Y=3$

x	1	2
$p(x)_{y=3}$	$\frac{2}{3}$	$\frac{1}{3}$

ix) find the conditional p.d.f of $Y|X=2$.

y	1	2	3
$p(y\|x=2)$	$\frac{1}{4}$	$\frac{2}{4}$	$\frac{1}{4}$

e.g)

X \ Y	1	2	3	
1	0.1	0.2	0.1	0.4
2	0.3	0.1	0.2	0.6
	0.4	0.3	0.3	1

Find :- (i) the marginal p.d.f of X & Y
(ii) $Cov(X,Y)$
(iii) the Correlation Coefficient (ρ)
(iv) the conditional p.d.f of $Y|X=1$
(v) $p(Y>1.2 | X=1)$.
(v) $E(Y|X=1)$
(vi) $E(Y^2|X=1)$
(viii) ~~E &~~ $Var(Y^2|X=1)$

(i)

x	1	2
$p(X=x)$	0.4	0.6

y	1	2	3
$p(Y=y)$	0.4	0.3	0.3

(ii) $E(X) = 1(0.4) + 2(0.6) = \cancel{0.5}\ 0.4 + 1.2 = 1.6$

$E(Y) = 1(0.4) + (2)(0.3) + 3(0.3)$

$= 0.4 + 0.6 + 0.9 = 1.9$

$E(X,Y) = 0.1 + 2(0.2) + 3(0.1) + 2(0.3) + 4(0.1) + 6(0.2)$

$= 3$

$CoV = E(X,Y) - E(X)E(Y)$

$= 3 - (1.6)(1.9) = -0.04$

(iii) $Corr = \frac{Cov(X,Y)}{\sigma_x \sigma_y} = \cancel{\frac{-0.04}{\sqrt{1.8}\sqrt{1.9}}} =$

$E(X^2) = 0.4 + 2.4 = 2.8$

$E(Y^2) = 0.4 + 1.2 + 2.7 = 4.3$

64

$Var(X) = E(X^2) - (E(X))^2 = 2.8 - (1.6)^2 = 0.24$

$Var(Y) = E(Y^2) - (E(Y))^2 = 4.3 - (1.9)^2 = 0.69$

$$\rho = \frac{Cov(X,Y)}{\sigma_X \sigma_Y} = \frac{-0.04}{\sqrt{0.24}\sqrt{0.64}} = -0.[illegible]$$

(iv)

y	1	2	3
$p(y \mid x=1)$	$\frac{1}{4}$	$\frac{2}{4}$	$\frac{1}{4}$

(v)

y	2	3
$p(y>1.2 \mid x=1)$	$\frac{2}{4}$	$\frac{1}{4}$

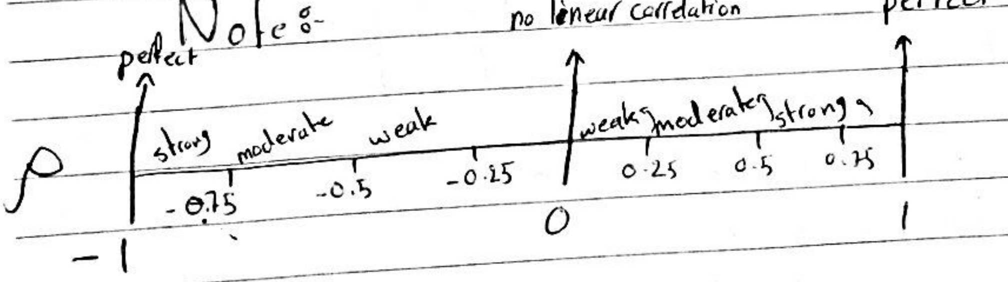
$$p((y>1.2) \mid x=1) = \frac{2}{4} + \frac{1}{4} = \frac{3}{4} = 0.75$$

(vi) $E(Y \mid X=1) = \frac{1}{4} + \frac{4}{4} + \frac{3}{4} = \frac{8}{4} = 2$

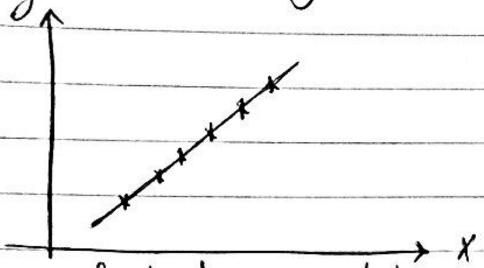
(vii) $E(Y^2 \mid X=1) = \frac{1}{4} + \frac{8}{4} + \frac{9}{4} = \frac{18}{4} = 4.5$

(viii) $Var(y \mid x=1) = E(y^2 \mid x=1) - (E(y \mid x=1))^2$
$= 4.5 - 4 = 0.5$

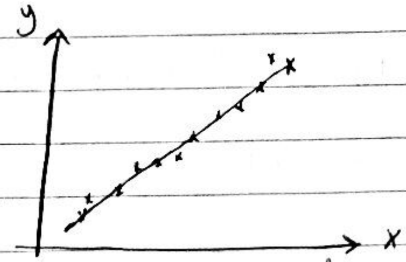
Note:-



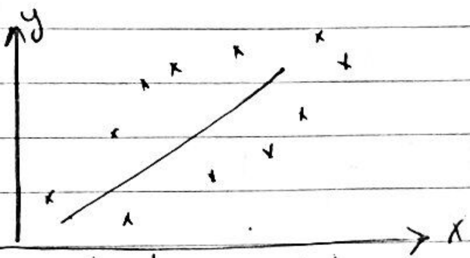
* Scatter Diagram :-



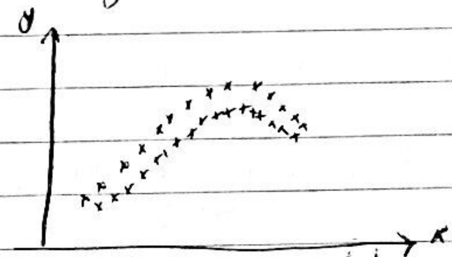
perfect +ve correlation



Strong +ve correlation



Weak +ve correlation



No linear correlation

* Independent random variable :-

Deffenition:- X & Y are independent if :-

$$f(x,y) = f_x(x) \cdot f_y(y)$$

or $p(X=x, Y=y) = p(X=x) \cdot p(Y=y)$

e.g)

$X \backslash Y$	1	2	
1	0.3	0.4	0.7
2	0.2	0.1	0.3
	0.5	0.5	1

Are X & Y independent?

Sol.)

x	1	2
$p(X=x)$	0.7	0.3

y	1	2
$p(Y=y)$	0.5	0.5

$p(X=1, Y=1) = 0.3$

$p(X=1) = 0.7$

$p(y=1) = 0.5$

$p(X=1) \cdot p(y=1) = 0.7(0.5) = 0.35$

Since $p(X=1, Y=1) \neq p(X=1) \cdot p(y=1)$

then X & Y are not independent

Notes:- If X & Y are independent, then

(i) $p(X=x, \& Y=y) = p(X=x) \cdot p(Y=y)$

(ii) $p(X=x \mid Y=y) = p(X=x)$

(iii) $Cov(X,Y) = 0$

(iv) $\rho = 0$

Note:- $Cov(ax+b, cy+d) = ac\, Cov(x,y)$

proof: $= E[(ax+b)(cy+d)] - E(ax+b) \cdot E(cy+d)$

$= E[adx + acxy + bcy + bd] - [aE(x)+b][cE(y)+d]$

$= \cancel{adE(x)} + ac\,E(x,y) + \cancel{bc\,E(y)} + \cancel{bd} - (acE(x)E(y) + \cancel{adE(x)} + \cancel{bcE(y)} + \cancel{bd})$

$= ac\,E(x,y) - ac\,E(x)\,E(y)$

$= ac\,[E(x,y) - E(x)\,E(y)]$

$= ac\,Cov(x,y)$ #

e.g) If $E(X) = E(Y) = 5$ & $E(X,Y) = 6$ } Find $Cov(2x+3, 3y-5)$

sol.) $Cov(x,y) = E(x,y) - E(x)(E(y)$
$= 6 - 25$
$= -19$.

$Cov(2x+3, 3y-5) = 6\,(Cov(x,y))$
$= 6\,(-19) = -114$.

$$\rho(ax+b, cy+d) = \frac{ac}{|ac|} \cdot \rho(x,y).$$

$$\rho(ax+b, cy+d) = \frac{Cov(ax+b, cy+d)}{\sigma_{ax+b} \quad \sigma_{cy+d}}$$

$$= \frac{ac \; Cov(x,y)}{\sqrt{a^2 Var(x)} \; \sqrt{c^2 Var(y)}}$$

$$= \frac{ac}{|a||c|} \frac{Cov(x,y)}{\sqrt{Var(x) \, Var(y)}}$$

$$= \frac{ac}{|ac|} \rho(x,y).$$

Note: If $y = ax+b$, then,

$\rho = 1$ if $a > 0$
$\rho = -1$ if $a < 0$

e.g) $Cov(x,y) = 20$, $Var(X) = 16$
$Var(y) = 64$, $E(X) = -3$, $E(Y) = 7$

calculate: (a) $Corr(X,Y)$ & (b) $E(X \cdot Y)$
(c) $Cov(3x-2, 7-4y)$ d) $Corr(9x+6, 7-8Y)$

(a) $\rho(x,y) = \frac{Cov(x,y)}{\sqrt{Var(x)}\sqrt{Var(y)}} = \frac{20}{\sqrt{64}\sqrt{16}} = 0.625 = \frac{5}{8}$

(b) $E(x,y) = \cancel{Var}\, Cov(X,y) + E(x)\, E(y)$

$= 20 + -21$

$E(x,y) = -1$

(c) $Cov(3x-2, 7-4y) = -12 \; Cov(x,y)$

$= -12 \; (20)$

$= -240$

(d) $Corr(4x+6, 7-8y) = - Corr(x,y)$

$$= -\frac{5}{8}$$

e.g) Find :

(i) $Corr(X, 2X+3) = 1$

(ii) $Corr(X, 5-6X) = -1$

(iii) $Corr(X,Y)$ when X & Y are independent.

$\hookrightarrow = 0$.

Note :- $Var(ax+by+c) = a^2 Var(X) + b^2 Var(y) + 2ab Cov(x,y)$

proof :-

$Var(ax+by)$

$$= E[(ax+by) - (a\mu_x + b\mu_y)]^2$$

$$= E[a(X-\mu_x) + b(Y-\mu_y)]^2$$

$$= E[a^2(X-\mu_x)^2 + b^2(Y-\mu_y)^2 + 2ab(X-\mu_x)(Y-\mu_y)]$$

$$= a^2 E(X-\mu_x)^2 + b^2 E(Y-\mu_y)^2 + 2ab\, E(X-\mu_x)(Y-\mu_y)$$

$$= a^2 Var(X) + b^2 Var(y) + 2ab\, Cov(x,y)$$

Note : If X & Y are independent, then

$$Var(ax+by) = a^2 Var(X) + b^2 Var(y)$$

e.g) $Var(2X+3Y+5) =$ if $Var(X) = 4$ $Var(y) = 5$ $Cov(x,y) = 3$

$4 Var(X) + 9 Var(y) + 2(6) Cov(x,y)$

$4(4) + 9(5) + 12(3) = 97$

(6a)

*

e.g)

$x \backslash y$	1	2	3	
1	0.3	0.2	0.3	0.8
2	0.1	0.05	0.05	0.2
	0.4	0.25	0.35	1

Find: (a) the marginal p.d.f of x & y

(b) the conditional p.d.f of $x|y=2$

(c) $E(x|y=2)$ & $Var(x|y=2)$

(d) $Cov(x,y)$

(e) $Corr(x,y)$

(f) $Cov(2x+3, 4-5y)$

(g) $Corr(2x+3, 4-5y)$

(h) $Var(2x+5y-3)$

(a)

x	1	2
$p(X=x)$	0.8	0.2

y	1	2	3
$p(Y=y)$	0.4	0.25	0.35

(b)

x	1	2
$p(x\|y=2)$	$\frac{0.2}{0.25}$	$\frac{0.05}{0.25}$

(c) $E(x|y=2) = 1\left(\frac{4}{5}\right) + 2\left(\frac{1}{5}\right)$

$= \frac{6}{5} = 1.2$

(d) $Cov(x,y) = E(x,y) - E(x)E(y)$

$E(x) = (1)(0.8) + 2(0.2) = 1.2$

$E(y) = 0.4 + 0.5 + 3(0.35) = 1.95$

$E(x,y) = 0.3 + 0.4 + 0.9 + 0.2 + 0.2 + 0.3$

$= 2.3$

$Cov(x,y) = 2.3 - (1.2)(1.95) = -0.04$

(e) $Corr(x,y) = \frac{Cov(x,y)}{\sigma_x \sigma_y} = \frac{-0.04}{\sqrt{(0.4)(0.86)}} = -0.116.$

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

$= (0.8 + 0.8) - (1.2)^2$

$= 1.6 - (1.2)^2$

$= 0.16$

$Var(y) = E(y^2) - (E(y))^2$

$= 4.55 - (1.95)^2$

$= 0.7475$

$\sigma_y = 0.86$

$\sigma_x = 0.4$

70

(f) $Cov(2x+3, 4-5y) = -10\,(Cov(x,y))$

$= -10\,(\cancel{0.23})(0.04)$

$= -\cancel{23}\ -0.4$

(g) $Corr(2x+3, 4-5y) = -\ Cor(x,y) = 0.116.$

(h) $Var(2x+5y-3) = 4\,Var(x) + 25\,Var(y) + 2(2)(5)\,Cov(x,y)$

$= 4(0.16) + 25(0.86) + 20\ (2.3)$

$= 68.14.$

Q 24 - 29
p: 84

X \ Y	0	1	2	
-1	$\frac{1}{12}$	$\frac{1}{6}$	$\frac{1}{12}$	$\frac{4}{12}$
0	$\frac{1}{12}$	$\frac{1}{6}$	$\frac{1}{12}$	$\frac{4}{12}$
1	$\frac{1}{12}$	$\frac{1}{6}$	$\frac{1}{12}$	$\frac{4}{12}$
	$\frac{3}{12}$	$\frac{3}{6}$	$\frac{3}{12}$	1

Let $E(Y) = 1$ $\quad E(Y^2) = 1.5$ $\quad E(X,Y) = 0$.

24) $P(X+Y=1) = P(X=0, Y=1) + P(X=-1, Y=2) + P(X=1, Y=0)$

$$= \frac{1}{6} + \frac{1}{12} + \frac{1}{12} = \frac{2}{6} = \frac{1}{3}$$

25) $P(X<1 \mid Y=1) = \dfrac{P(X<1, Y=1)}{P(Y=1)} = \dfrac{P(X=-1, Y=1) + P(X=0, Y=1)}{P(Y=1)}$

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

26) $E[2X - 4Y^2 - 3XY + 1] = 2E(X) + 4E(Y) - 3E(X,Y) + 1$

$$= \cancel{2(0)} - 4(1.5) - \cancel{3(0)} + 1$$

$$= \cancel{}3 = -6 + 1 = -5$$

$$\boxed{E(X) = \frac{-4}{12} + 0 + \frac{4}{12} = 0}$$

27) $E(X^2) = \frac{4}{12} + 0 + \frac{4}{12} = \frac{8}{12} = \frac{2}{3}$

28) $Var(2X - 4Y - 1) = Var(2X - 4Y)$

$$= 4\,Var(X) + 16\,Var(Y) + -16\,Cov(X,Y)$$

$Var(X) = E(X^2) - (E(X))^2$
$= \left(\frac{4}{12} + 0 + \frac{4}{12}\right) - 0$
$Var(X) = \frac{2}{3}$

$Var(Y) = E(Y^2) - (E(Y))^2$
$= 1.5 - 1$
$Var(Y) = 0.5$

$Cov(X,Y) = E(X,Y) - E(X)E(Y)$
$= 0 - 0 = 0$

$$Var(2X - 4Y - 1) = 4\left(\frac{2}{3}\right) + 16(0.5) - 16(0)$$

$$= \frac{8}{3} + 8 = \frac{32}{3}$$

29) $Corr(2X - 1, 1 - 3Y) = -Corr(X, Y)$

$$= -\frac{Cov(X,Y)}{\sigma_x \sigma_y} = 0$$

$Var(y) = E(y^2) - (E(y))^2$

$\boxed{Q_{22} - 23}$

$E(X) = E(Y) = 0$

$Var(X) = 1$

$Var(y) = 4$

$Corr(X,Y) = \rho = -\frac{1}{2}$

22) $E(3x - 4y + 1) = 3E(x) - 4E(y) + 1$

~~E(y) =~~

$= 3(0) - 4(0) + 1$

$= 1$

23) $Var(2x - 3y + 1) = 4 Var(x) + 9 Var(y) \cancel{+12} - 6 Cov(x,y)$

$Cov(x,y) = Corr(x,y)\, \sigma_x \sigma_y$

$= -\frac{1}{2} \times 1 \times 2$

$\boxed{Cov(x,y) = -1}$

$= 4(1) + 9(4) - 6(-1)(2)$

$= 4 + 36 + 12 = \cancel{46}\ 52.$

$\boxed{Q_{30} \to 31}$

$E(X) = 3$ $\quad Var(X) = 4$ $\quad Cov(X,Y) = -1$

$E(Y) = 5$ $\quad Var(y) = 9$

30) $Cov(-X+1, 2y+3) = -2\, Cov(X,Y)$

$= -2(-1) = 2.$

~~$= -2[E(X,Y) - E(X)E(Y)]$~~

31) $Corr(-X, 2Y) = -Corr(X,Y)$

$= -\left(\frac{-1}{(3)(2)}\right) = \frac{1}{6}$

Q32 - 34

X \ Y	1	2	3	
0	0.05	0.15	0	0.2
1	0.1	0.2	0.2	0.5
2	0.05	0.25	0.	0.3
	0.2	0.6	0.2	1

Q_{32}: $p(X - Y > 0) = p(X=2, Y=1) = 0.05$

Q_{33}: $p(X \geq 1 \mid Y = 2) = \dfrac{p(X \geq 1, Y=2)}{p(Y=2)}$

$$= \frac{p(X=1, Y=2) + p(X=2, Y=2)}{p(Y=2)}$$

$$= \frac{0.2 + 0.25}{0.6} = \frac{0.45}{0.6} = \frac{3}{4}$$

Q_{34}: $E(X^2 - 2Y - 1.7) =$

$E(X^2) = 0 + 0.5 + 1.2$

$= 1.7$

$E(Y) = 0.2 + \cancel{1.8}^{1.2} + 0.6 = \cancel{0.7} = 2$

$E(X^2 - 2Y - 1.7) = E(X^2) - 2E(Y) - 1.7 = \cancel{1.7} - 2(2) - \cancel{1.7}$

$= -4.$

Q35 - 36 $E(X) = 2.5$ $E(Y) = 4$ $Var(X) = 4$ $Var(Y) = 1$

$E(X,Y) = 12.$

Q_{35}: $Cov(2X - 1, 5 - 3Y) = -6\, COV(X, Y)$

$= -6\,(E(X,Y) - E(X)E(Y)) = -6\,(12 - (2.5)(4))$

$= -6\,(12 - 10)$

$= -12.$

Q_{36}: the standard deviation of $(X - 3Y + 1)$

$Var(X - 3Y + 1) = Var(X) + 9\, Var(Y) + -6\, Cov(X, Y)$

$= 4 \cancel{25} + 9(1) - 6\,[E(X,Y) - E(X)E(Y)]$

$= \cancel{4.25} + 13 - 6\,(12 - 2.5(4))$

$= \cancel{22.25} = 13 - 12 = 1$

$\sigma_{(X - 3Y + 1)} = \sqrt{1} = 1$

74

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

Q29 X is a (r.v) such that
$E(X) = \mu$, $Var(X) = \sigma^2$

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

show that $E(Z) = 0$, $Var(Z) = 1$.

$$E(Z) = E\left(\frac{X-\mu}{\sigma}\right)$$
$$= E\left(\frac{X}{\sigma} - \frac{\mu}{\sigma}\right)$$
$$= E\left(\frac{X}{\sigma}\right) - E\left(\frac{\mu}{\sigma}\right) = \frac{1}{\sigma} E(x) - \frac{\mu}{\sigma}$$
$$= \frac{\mu}{\sigma} - \frac{\mu}{\sigma} = 0 \quad \#$$

Now $Var(Z) = Var\left(\frac{X-\mu}{\sigma}\right)$
$$= Var\left(\frac{X}{\sigma}\right)$$
$$= \frac{1}{\sigma^2} Var(X) = \frac{1}{\sigma^2} \sigma^2 = 1$$

Q30 $\mu = 100$
$\sigma^2 = 25$

(a) $E(-2x + X^2) = -2E(X) + E(X^2)$
$$= -2\mu + Var(X) + (E(X))^2$$
$$= -2\mu + \sigma_x^2 + \mu^2$$
$$= -200 + 25 + 10000$$
$$= 9825.$$

(b) $E\left(\frac{X - \overset{\mu}{100}}{\underset{\sigma}{5}}\right) = E\left(\frac{X}{5} - \frac{100}{5}\right)$

Z score
$$= \frac{1}{5} E(X) - 20$$
$$= \frac{1}{5}(100) - 20$$
$$= 0$$

(75)

* Some special distributions:-

- Discrete:-
  1. Binomial
  2. poisson
  3. Geometric
  4. Hypergeometric

- Continuous:-
  1. Normal
  2. t-distribution
  3. χ^2-distribution
  4. F-distribution

  (2)–(4): مهمين

* The Binomial Distribution:-

If (i) the ~~distribution~~ experements are independent
(ii) the outcomes in each trial are success or fail only
(iii) we will repeat the experiment n-times $(n \geq 3)$

Then the (iv) random variable (x) is said to follow a binomial distribution. This can be written as: $X \sim Bin(n,p)$; where
n: no. of trials
p: probability of successful in each trial

- The probability of successful in x trials is given by:-

$$f(x) = P(X=x) = \binom{n}{x} p^x \cdot q^{n-x} \text{ ; where}$$

$q = 1-p$ ← مكملة

$x \in \{0,1,2,3,\dots,n\}$

This is the p.d.f for the r.v (X) if it follows a binomial distribution.

Note: $E(X) = np$
$Var(X) = npq$
$S.D(X) = \sqrt{npq}$

e.g) A fair coin is tossed 10 times. Find:-

(a) the prob. of getting:

(i) exactly 7 heads

(ii) at least 9 heads

(iii) at least 2 heads

(iv) at most 1 head

(v) at most 8 heads

(vi) at least 4 given that at most 1

(b) the expected no. of heads, the variance & the standard deviation

Sol) $n = 10$

$p = \frac{1}{2}$ $\quad X \sim Bin(10, \frac{1}{2})$

$q = 1 - p = \frac{1}{2}$

(a) (i) $p(X=7) = \binom{10}{7}(\frac{1}{2})^7(1-\frac{1}{2})^3$

$= (120)(\frac{1}{2})^{10} = \frac{120}{1024} = \frac{15}{128}$

(ii) $p(X \geq 9) = p(X=9) + p(X=10)$

$= \binom{10}{9}(\frac{1}{2})^9(\frac{1}{2}) + \binom{10}{10}(\frac{1}{2})^{10}(\frac{1}{2})^0$

$= \frac{10}{1024} + \frac{1}{1024} = \frac{11}{1024}$

(iii) $p(X \geq 2) = 1 - p(X < 2) = 1 - [p(X=1) + p(X=0)]$

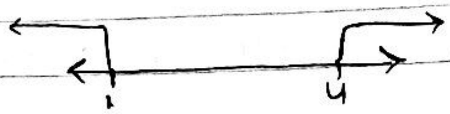
$= 1 - [\binom{10}{1}(\frac{1}{2})(\frac{1}{2})^9 + \binom{10}{0}(\frac{1}{2})^0(\frac{1}{2})^{10}]$

$= 1 - \frac{11}{1024} = \frac{1013}{1024}$

(iv) $p(X \leq 1) = p(X=1) + p(X=0) = \frac{11}{1024}$

(v) $p(X \leq 8) = 1 - p(X > 8) = 1 - [p(X=9) + p(X=10)] = 1 - \frac{11}{1024} = \frac{1013}{1024}$

(vi) $p(X \geq 4 \mid X \leq 1) = \dfrac{p(\emptyset)}{p(X \leq 1)} = 0$



~~(exit)~~ (b) $E(X) = np = 10\left(\frac{1}{2}\right) = 5$

$Var(X) = npq = 10\left(\frac{1}{2}\right)\left(\frac{1}{2}\right) = 2.5$

$S.D(X) = \sqrt{2.5} = 1.58$

e.g) If $X \sim Bin(3, p)$ & $p(X \geq 1) = \frac{19}{27}$, Find $E(X)$.

Sol.) $p(X \geq 1) = \frac{19}{27}$

$1 - p(X < 1) = \frac{19}{27}$

$1 - p(X=0) = \frac{19}{27} \rightarrow p(X=0) = 1 - \frac{19}{27} = \frac{8}{27}$

$p(X=0) = \binom{3}{0}(p)^0(1-p)^3 = \frac{8}{27}$

$= \sqrt[3]{1(1-p)^3} = \sqrt[3]{\frac{8}{27}}$

$1 - p = \frac{2}{3} \rightarrow p = \frac{1}{3}$

$E(X) = np = 3\left(\frac{1}{3}\right) = 1$

e.g) $X \sim Bin(n, p)$, mean $= 2$, Variance $= 1.6$

Find: (a) $p(X \leq 4)$ (b) $p(X < 4)$.

Sol.) $E(X) = np = 2$

$Var(X) = npq = 1.6$ $\div$

$\dfrac{1.6}{2} = \dfrac{\cancel{np}q}{\cancel{np}} \Rightarrow q = 0.8$

$p = 0.2$

$n(0.2) = 2 \quad \therefore \quad n = 10$

a) $p(X \leq 4) = p(X=2) + \cdots + p(X=4)$

Or from the table
$= 0.967$

في الجداول :
~~تحت~~ القيم لازم تكون
(عند $X \leq$)
عشان تاخد الأرقام

b) $p(X<4) = p(X \leq 3) = 0.879$

c) $p(X>3) = 1-(p \leq 3) = 1 - 0.879 = 0.121$

d) $p(X \geq 4) = 1 - p(X<4) = 1 - p(X \leq 3) =$
$= 1 - 0.879 = 0.121$

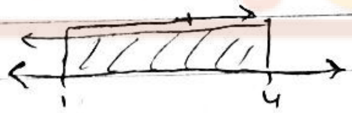
e) $p(2 \leq X \leq 5)$
$p(X \leq 5) - p(X \leq 1) =$

f) $p(2 < x \leq 5) = p(X \leq 5) - p(X \leq 2)$
$=$

g) $p(2 < x < 5) = p(X \leq 4) - p(X \leq 2)$

(h) $p(2 \leq x < 5) = p(x \leq 4) - p(x \leq 1)$

(i) $p(X \leq 4 \mid X \geq 1)$



$= \dfrac{p(4 \geq x \geq 1)}{p(X \geq 1)} = p\ \dfrac{(X \leq 4) - p(X=0)}{1 - p(X=0)}$

Note: If $X \sim Bin(n,p)$ then,
$p(X \leq 0) = p(X=0)$

~~(ii) $p(X \geq 2) = 1 - p(X<2) = 1 - [p(X=1) + p(X=0)]$~~

p:116 part 2 | Q_1: If the probability of a surgical operation will be successful is 0.8 In a certain day, 3 operations will be performed, what is the probability that:

(a) all of them will be successful
(b) at least 2 will be successful
(c) not more than 1 will not be successful.

Sol) $n=3$
$p=0.8$ $\quad X \sim Bin(3, 0.8)$
$q=0.2$

a) $p(X=3) = \binom{3}{3}(0.8)^3(0.2)^0$
$= (0.8)^3 = 0.512$

(b) $p(X \geq 2) = 1 - p(X<2)$
$= 1 - p(X \leq 1)$ "from the table"
$= 0.896.$

(c) $p(X \geq 2)$ same $= 1 - p(X<2)$
$= 1 - p(X \leq 1) = 1 - 0.104 = 0.896.$

Q_3: $X \sim Bin(25, 0.1)$; Find:

(a) $E[X(X+1)]$
(b) $E[3X + X^2]$

(a) $E(X^2+X) = Var(X) + (E(X))^2 + E(X)$
$= npq + (np)^2 + np$

$= 25(0.1)(0.9) + (25(0.1))^2 + (25)(0.1)$

$=$

(b) $E(3X+X^2) = E(3X) + E(X^2)$
$= 3E(X) + E(X^2)$
$E(X) + (Var(X) + (E(X))^2) = 3(2.5) + (2.5 + 0.9) + (2.5)^2$
$= 16$

Q4 :- $X \sim Bin(25, 0.2)$; Find :-

(a) $p(\mu - \sigma < X < \mu + \sigma)$

(b) $p(\mu - \sigma \leq X \leq \mu + \sigma)$

(c) $p(\mu - 2\sigma < X < \mu + 2\sigma)$

(d) $p(\mu - 2\sigma \leq X \leq \mu + 2\sigma)$

Sol.) $\mu = E(X) = np = 25(0.2) = 5$

$Var(x) = npq = 25(0.2)(0.8) = 4$

$\sigma = \sqrt{var(x)} = 2$

(a) $p(5-2 < X < 5+2)$

$p(3 < X < 7)$

$= p(X \leq 6) - p(X \leq 3)$ "from the table"

(b) $p(3 \leq X \leq 7) = p(X \leq 7) - p(X \leq 3)$ "from the table.

(c) $p(5-4 < X < 5+4) = p(1 < X < 9)$

$= p(X \leq 8) - p(X \leq 1)$ "from the table"

(d) $p(1 \leq X \leq 9) = p(X \leq 9) - p(X \leq 0)$ "from the table"

The poisson Distribution :-

- If we want to study the occurence of an event during a specific period of time. (or given distance, area, volume, ---) & If the events are independent, then we say that the random variable X follows a poisson distribution.

This is donated by ; $X \sim poi(\mu)$

where,

μ : Average (or mean) of the occurance during the given period of time.

_The p.d.f is :

$$f(x) = \frac{e^{-\mu} \cdot \mu^{x}}{x!} \qquad x = 0, 1, 2, 3, ----$$

mean $= E(X) = \mu$
Variance $= Var(X) = \mu$
Standard deviation $= \sigma = \sqrt{\mu}$

e.g) X: no. of patients admitted to the (I.C.U) at the university hospital per day. If the average no. of persons admitted to the ICU per day is 3. What is the probability that (a) in a given day there will be exactly 2 admitions.
(b) " " " " " " " at most 2 admitions.
(c) " " " weak " " " exactly 7 admitions.

Sol.) (a) $X \sim poi(3)$
$p(X=2) = \dfrac{e^{-3} \cdot (3)^2}{2!} = 0.224$. "Calculator"

(b) $X \sim poi(3)$
$p(X \leq 2) = 0.423$ "from the table"

(c) $X \sim poi(7 \times 3)$
$p(X=7) = \dfrac{e^{-21} * 21^7}{7!} =$

e.g) $X \sim p(\mu)$, If $p(X=0) = p(X=1) + p(X=2)$ find μ
$\therefore \dfrac{e^{-\mu} \cdot \mu^0}{0!} = \dfrac{e^{-\mu} \cdot \mu^1}{1!} + \dfrac{e^{-\mu} \cdot \mu^2}{2!}$

$\cancel{e^{-\mu}} \cdot \cancel{\mu^0} = \cancel{e^{-\mu}} \mu + \dfrac{\cancel{e^{-\mu}} \cdot \mu^2}{2}$

$1 = \mu + \dfrac{\mu^2}{2}$

$2 = 2\mu + \mu^2$

$\mu^2 + 2\mu - 2 = 0$

$\mu = -2 \pm \sqrt{2(2) - 4(1)(-2)}$

$= \dfrac{-2 \pm \sqrt{12}}{2}$

$\mu = -1 \pm \sqrt{3}$

$\boxed{\mu = 1 + \sqrt{3}}$ $(\mu = 1 - \sqrt{3})$ X

Notes We can approximate the binomial distribution by a poisson distribution if:-

(i) $n \geq 30$

(ii) $p < 0.10$

} Thus, $X \sim Bin(n,p) \cong poi(np)$

e.g) $X \sim Bin(100, 0.02)$

Find: (i) $p(X \leq 4)$

(ii) $p(X = 4)$

Sol.) $n = 100 > 30$

$p = 0.02 < 0.1$

$\therefore$ we can use a poisson distribution with:

$\mu = np$

$\therefore \mu = 100(0.02) = 2$

$X \sim poi(2)$

(i) $p(X \leq 4) \rightsquigarrow$ from the table $= 0.947$

(ii) $p(X = 4) \rightsquigarrow \frac{e^{-2} \cdot 2^4}{4!} = 0.0902$

OR $p(X \leq 4) - p(X \leq 3) = 0.947 - 0.857$ "from the table"

Q10 p: 118 | $X \sim Bin(100, 0.05)$

Find: (a) $p(3 \leq X < 7)$

(b) $p(5 < X < 8)$

(c) $p(X > 3)$

(d) $p(X = 6)$

Sol) $n = 100 > 30$

$p = 0.05 < 0.1$

$\therefore X \sim poi(np)$

$X \sim poi(5)$

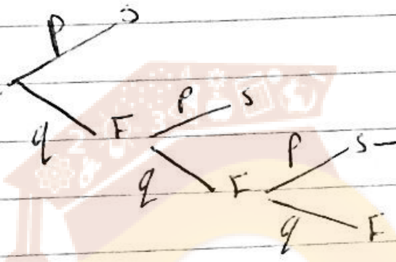
(a) $p(3 \leq X < 7) = p(X \leq 6) - p(X \leq 2)$ $\quad 0,1,\cancel{2},\boxed{3,4,5,6}$

(b) $p(5 < X < 8) = p(X \leq 7) - p(X \leq 5)$ $\quad \underbrace{0, \dots 5}_{X}, 6, 7$

(c) $p(X > 3) = 1 - p(X \leq 3)$

(d) $p(X = 6) = p(X \leq 6) - p(X \leq 5)$

* Geometric Distribution :-



$\longrightarrow p(X=3) = q^2 \cdot p$

$$p(X = x) = q^{x-1} \cdot p$$

If the trial are independent, then $X \sim Geo(p)$, where

p: probability successful in each trial

* p.d.f $\longrightarrow p(X=x) = q^{x-1} \cdot p \quad ; \quad x = 1, 2, 3, \dots$

* $E(x) = \dfrac{1}{p}$

$Var(x) = \dfrac{q}{p^2} \quad ; \quad \sigma = \dfrac{\sqrt{q}}{p}$

e.g) If the probability of hitting a target in a single shot is 0.7. If shots are to be fired at the target independently, what is the prob. that the 1st hit to the target will be in the:

(a) 3rd shot
(b) 5th shot

Sol.) $X \sim Geo\,(0.7)$ $\rightarrow p = 0.7$, $\rightarrow q = 0.3$

(a) $p(X=3) = (0.3)^2 (0.7) = 0.063$

(b) $p(X=5) = (0.3)^4 (0.7) = 0.00567$

e.g) X is Geometric with mean 4, Find :-

(a) $p(X=3)$

(b) $E(X^2)$

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

$$\left.\begin{aligned} 4 &= \frac{1}{p}, \\ p &= \frac{1}{4} \end{aligned}\right\} \quad \therefore X \sim Geo\left(\tfrac{1}{4}\right), \quad q = \frac{3}{4}$$

(a) $p(X=3) = \left(\frac{3}{4}\right)^2 \left(\frac{1}{4}\right) = 0.141$

(b) $E(X^2) = Var(X) + (E(X))^2$

$= \frac{q}{p^2} + (4)^2$

$= \frac{3/4}{(1/4)^2} + 16$

$= 28$

Q18 / p:114 | Assume the prob. of hitting a target is 0.6

[a] If 3 independent shots are fired at the target, what is the prob. that the target will be hit.

[b] What is the no. of shots need to be fired so that the prob of hitting the target is at least 90%.

Sol.) $X \sim Geo(0.6) \rightarrow p = 0.6$, $q = 0.4$.

(a) . $p(X=1) + p(X=2) + p(X=3)$

$= (0.4)^0(0.6) + (0.4)^1(0.6) + (0.4)^2(0.6)$
$= 0.936$

بناءً على الفرع السابق
(b) $X \geq 3$ (بالتجربة)

Q20) suppose that items to be testen sequentially (بالتوالي)
if the proportion of defected items is 10% , what is the probability that the first defective item will occur in the :

(a) 3^{rd} test
(b) 5^{th} test

Sol.) $X \sim Geo(0.1)$ — $p = 0.1$, $q = 0.9$

(a) $p(X=3) = (0.9)^2(0.1) = 0.081$

(b) $p(X=5) = (0.9)^4(0.1) =$

[2] * Hyper Geometric distribution :-

$$\boxed{\begin{array}{cc} \underline{A} & \underline{B} \\ M & N-M \\ x & n-x \end{array}}$$
N

(a). items to be selected without replacement

$$p(X=x) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}} \leftarrow \text{p.d.f if } X \sim Hyp(n, M, N)$$

Notes: $E(X) = n . \frac{M}{N}$

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

e.g) A box containing 10 electric bulbs, among them there are 4 defectives. If three (3) bulbs are drawn without replacement from the box, what's the probability of getting:

(a) exactly 1 defictive bulb
(b) exactly 2 defictive bulbs
(c) not more than 2 defictive bulbs.

Sol.)

D	D'
4	6

10

$X \sim Hyp(3; 4; 10)$.

(a) $p(X=1) = \dfrac{\binom{4}{1}\binom{6}{2}}{\binom{10}{3}} = 0.5$

(b) $p(X=2) = \dfrac{\binom{4}{2}\binom{6}{1}}{\binom{10}{3}} = 0.1$

(c) $p(X \leq 2) = p(X=0) + p(X=1) + p(X=2)$

$= \dfrac{\binom{4}{0}\binom{6}{3}}{\binom{10}{3}} + \dfrac{\binom{4}{1}\binom{6}{2}}{\binom{10}{3}} + \dfrac{\binom{4}{2}\binom{6}{1}}{\binom{10}{3}}$

P119
Q24

$X \sim Hyp(\overset{n}{6}; \overset{M}{9}, \overset{N}{15})$

$p(X=x) = \dfrac{\binom{9}{x}\binom{6}{6-x}}{\binom{15}{6}}$

(a) $p(X=4) = \dfrac{\binom{9}{4}\binom{6}{2}}{\binom{15}{6}} = 0.3776$

(b)

(b) $p(X \leq 2) = p(X=0) + p(X=1) + p(X=2)$

$$= \frac{\binom{9}{0}\binom{6}{6}}{\binom{15}{6}} + \frac{\binom{9}{1}\binom{6}{5}}{\binom{15}{6}} + \frac{\binom{9}{2}\binom{6}{4}}{\binom{15}{6}}$$

(c) $p(2 \leq X < 5) = p(X=2) + p(X=3) + p(X=4)$

$$= \frac{\binom{9}{2}\binom{6}{4}}{\binom{15}{6}} + \frac{\binom{9}{3}\binom{6}{3}}{\binom{15}{6}} + \frac{\binom{9}{4}\binom{6}{2}}{\binom{15}{6}}$$

Note :-

For larg N : $Hyp(n, M, N)$; then this can be approximated by binomial distribution.

i.e $X \sim Bin(n, p)$, where: $p = \frac{M}{N}$, $n = n$

e.g) If a random sample of 10 persons is drawn without replacement from a population of 1000 persons. If 200 persons own cars, what is the prob. of getting :-

(a) 3 own Cars

(b) at least 4 own cars :-

$n = 10$

$N = 1000$

$M = 200$

$X \sim Bin(n, p)$

$p = \frac{M}{N} = 0.2$

$X \sim Bin(10, 0.2)$

Sol.) a) $p(X=3) = \binom{10}{3}(0.2)^3(0.8)^7$

$= 0.201$

b) $p(X \geq 4) = 1 - p(X < 4)$

$= 1 - p(X \leq 3)$ "from the table"

$= 1 - 0.879$

$= 0.121$

88

Q13 p:118

$X \sim Bin(5, 0.25) \to p(X=x) = \binom{5}{x}(0.25)^x (0.75)^{5-x}$

$Y \sim poi(7) \longrightarrow p(Y=y) = \frac{e^{-7} \cdot 7^y}{y!}$

$W \sim Hyp(4, 10, 30) \longrightarrow p(W=w) = \frac{\binom{10}{w}\binom{20}{4-w}}{\binom{30}{4}}$

X & Y & W Are independent, Find:

(a) $p(X+Y \leq 1)$

(b) $p(W=2 | Y=5)$

(c) $Var(3X - 5W + a)$

(d) $E(X+Y)^2$

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

$Y: 0, 1, 2, 3, 4, 5, ...$

$W: 0, 1, 2, 3, 4$

(a) $p(X+Y \leq 1) = p(X=0, Y=0) + p(X=0, Y=1) + p(X=1, Y=0)$

independent

$= p(X=0) \cdot p(Y=0) + p(X=0) \cdot p(Y=1) + p(X=1) \cdot p(Y=0)$

$= \cancel{\binom{5}{0}(0.25)^0 (0.75)^5 + e^{-7} 7}$

$= \left(\binom{5}{0}(0.25)^0 (0.75)^5\right) \cdot \frac{e^{-7} \cdot 7^0}{0!} + \binom{5}{0}(0.25)^0 (0.75)^5 \cdot \frac{e^{-7} \cdot 7^1}{1!} + \binom{5}{1}(0.25)^1 (0.75)^4$

$* \frac{e^{-7} \cdot 7^0}{0!}$

(b) $p(W=2 | Y=5) = p(W=2) = \frac{\binom{10}{2}\binom{20}{2}}{\binom{30}{4}}$

(c) $Var(3X - 5W) = Var(3X) + Var(-5W)$

$= 9\, Var(X) + 25\, Var(W)$

(d) $E(X^2 + 2XY + Y^2) = E(X^2) + 2E(XY) + E(Y^2)$

$E(X^2) = Var(X) + (E(X))^2$

$= 1.25(0.75) + (1.25)^2$

$= 2.5$

(8a)

p.118 / Q.14

$X \sim Bin(3, 0.5)$

$Y \sim Bin(4, 0.4)$

X & Y are independent, Compute:-

(a) ~~p(E(X~~ $p(E(Y) \leq Y \leq 3)$ (d) $E(XY)$

(b) $p(X > 1 \mid X < 3)$ (e) $Var(X+Y)$

(c) $p(X+1 > 1)$

Sol.) $p(X=x) = \binom{3}{x}(0.5)^x(0.5)^{3-x}$

$p(Y=y) = \binom{4}{y}(0.4)^y(0.6)^{4-y}$

(a) $E(Y) = np = 4(0.4) = 1.6$

$\therefore p(1.6 \leq Y < 3)$
$= p(Y=2) + p(Y=3) = \binom{4}{2}(0.4)^2(0.6)^2 + \binom{4}{3}(0.4)^3(0.6)^1$

(b) $p(X>1 / X<3) = \frac{p(X>1 \ \& \ X<3)}{p(X<3)}$

$= \frac{p(1<X<3)}{p(X<3)} = \frac{p(X=2)}{p(X\leq 2)} \rightarrow$

(c) $p(X+Y>1) = 1 - p(X+Y \leq 1)$

$x = 0,1,2,3$

$y = 0,1,2,3,4$

$= 1 - [p(X=0, Y=0) + p(X=0, Y=1) + p(X=1, Y=0)]$

$= 1 - [p(X=0) \cdot p(Y=0) + p(X=0) \cdot p(Y=1) + p(X=1) \cdot p(Y=0)$

$= 1 - [\binom{3}{0}(0.5)^0(0.5)^3 \cdot \binom{4}{0}(0.4)^0(0.6)^4 + \binom{3}{0}(0.5)^3(0.5)^0 \cdot \binom{4}{1}(0.4)^1(0.6)^3$
$+ \binom{3}{1}(0.5)^1(0.5)^2 \cdot \binom{4}{0}(0.4)^0(0.6)^4] =$

.75
.96
1.71

95

(d) $E(X+Y) = E(X) + E(Y) = (3)(0.5) + (4)(0.4)$
$= 1.5 + 1.6 = 2.4$

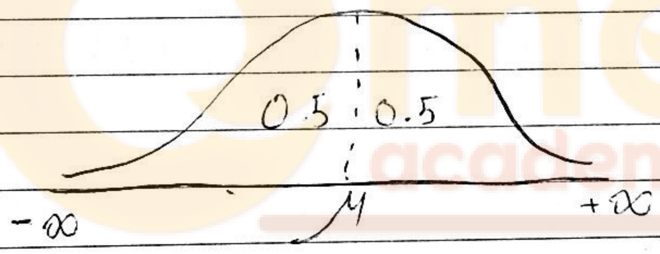
(e) $Var(X+Y) = Var(X) + Var(Y) = 3(0.5)(0.5) + 4(0.4)(0.6)$
$= 0.75 + 0.96 = 1.71$

* Special continuous distributions:-

[1] The normal distribution :-

$X \sim n(\mu, \sigma^2)$

لا نستعملها في الحل. p.d.f $\rightarrow f(x) = \dfrac{1}{\sqrt{2\pi}\,\sigma} \cdot e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$



* properties:-

[1] Symmetrical about the mean (μ)

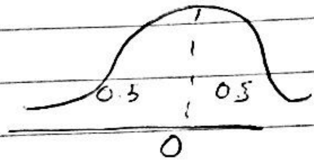
[2] μ (mean) = mode = median (Q_2)

[3] Total area under the curve = 1 unit

[4] Probability $\equiv$ Area.

* Standard Normal distribution :-

$Z \sim n(\overset{\mu}{0}, \overset{\sigma^2}{1})$



* $p(Z \leq k) \rightarrow$ from the table

continuous فقط
المساحة لاتفرق

e.g) If $Z \sim n(0,1)$. Find:

i) $p(Z \leq 1) =$

ii) $p(Z < 1)$

(iii) $p(Z \leq -1)$

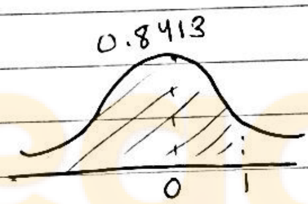
(iv) $p(Z > -1)$

~~(v) $p(Z > 1)$~~

(v) $p(Z > 1)$

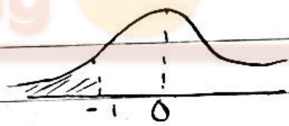
(vi) $p(-1 < Z < 1)$

Sol.)



i) $p(Z \leq 1) = 0.8413$

ii) $p(Z < 1) = p(Z \leq 1) = 0.8413$

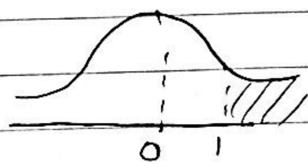


(iii) $p(Z \leq -1) = 0.1587$

(iv) $p(Z > -1) = 1 - p(Z \leq -1)$

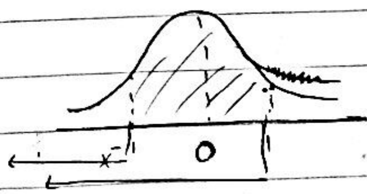
$= 1 - 0.1507$

$= 0.8413$



(v) $p(Z > 1) = 1 - 0.8413 = 0.1507$

(vi) $p(-1 < Z < 1)$



$= p(Z \leq 1) - p(Z \leq -1)$

$= 0.8413 - 0.1587$

$=$

(92)

e.g) Find k if :-

i) $p(z \le k) = 0.8413$

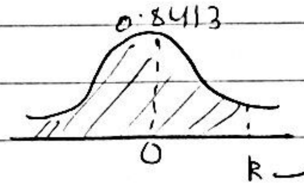
ii) $p(z \le k) = 0.1587$

(iii) $p(z > k) = 0.1587$

(iv) $p(z > k) = 0.8413$

(v) $p(-k < z < k) = 0.6826$

Sol.) (i)

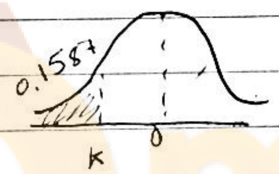


الذي يحدد k ⊕ أو ⊖ هو ال $p(z \le k)$ ..
إذا كانت > 0.5 فهو موجب
والعكس سالب .

من الجدول

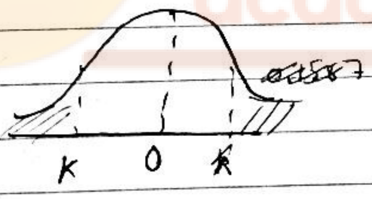
$k = 1$

(ii)



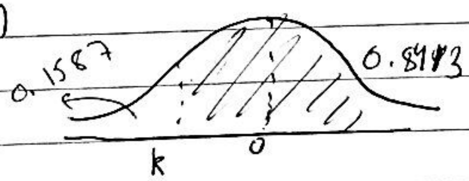
$k = -1$

(iii)

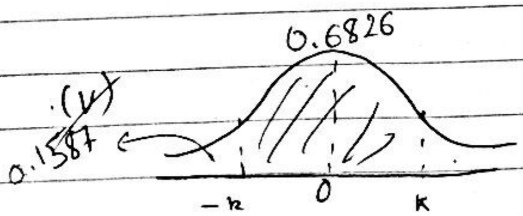


~~k=1~~ $k = -1$

(iv)



$k = -1$



$\frac{1 - .6826}{2} = 0.1587$

$-k = -1$

$k = 1$

* Standardization :-

If $X \sim n(\mu, \sigma^2)$

$$Z = \frac{X - \mu}{\sigma} \sim n(0,1)$$

eg) If $X \sim n(5,4)$ $\begin{matrix} \mu = 5 \\ \sigma = 2 \end{matrix}$

Find : (i) $p(X \leq 4)$

(ii) $p(X > 6)$

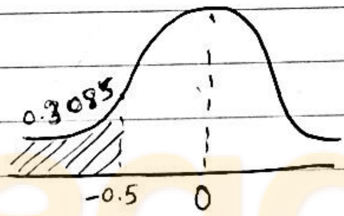
(iii) $p(4 < X < 6)$

(iv) k if $p(X \leq k) = 0.6915$

Sol.) (i) $p(X \leq 4)$

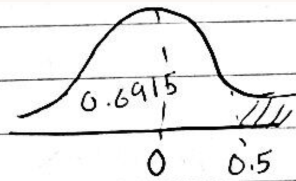
$$p\left(\frac{X - 5}{2} \leq \frac{4-5}{2}\right)$$

$$p(Z \leq -0.5) = 0.3085$$



(ii) $p(X > 6) = p\left(Z > \frac{6-5}{2}\right) = p(Z > 0.5)$

$= 1 - 0.6915 = 0.3085$

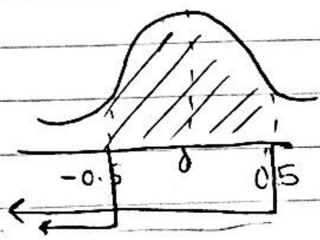


(iii) $p(4 < X < 6) = p\left(\frac{4-5}{2} < Z < \frac{6-5}{2}\right)$

$= p(-0.5 < Z < 0.5)$

$= 0.6915 - 0.3085$

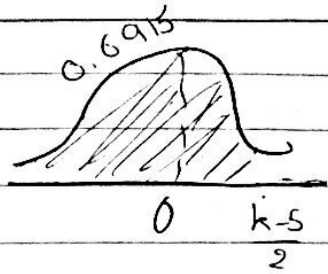
$= 0.3830$



(a4)

iv) $p(X \leq k) = 0.6915$

$p\left(Z \leq \frac{k-5}{2}\right) = 0.6915$



$\frac{k-5}{2} = \frac{0.5}{1}$ $\quad k-5 = 1$

$\boxed{k = 6}$

e.g) $X \sim n(\mu, 9)$

If $p(X > 5) = 0.3085$ find μ

Sol.) $p\left(Z > \frac{5-\mu}{3}\right) = 0.3085$

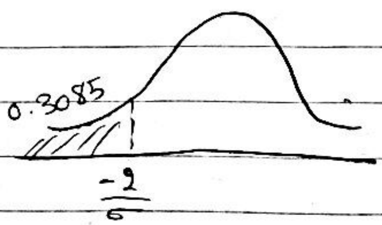
$\frac{5-\mu}{3} = 0.5 \rightarrow \mu = 3.5$

e.g) $X \sim n(10, \sigma^2)$

If $p(X < 8) = 0.3085$ find σ

Sol.) $p\left(Z < \frac{8-10}{\sigma}\right) = 0.3085$

$p\left(Z < \frac{-2}{\sigma}\right) = 0.3085$



$-\frac{2}{\sigma} = -0.5$

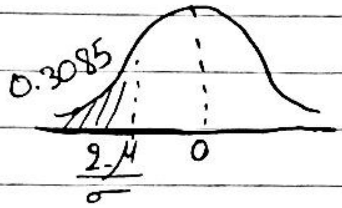
$\sigma = 4$

e.g) If $X \sim n(\mu, \sigma^2)$,

$p(x \leq 2) = 0.3085$ & $p(x \leq 8) = 0.9265$

find μ & σ

Sol.) $p(z \leq \frac{2-\mu}{\sigma}) = 0.3085$

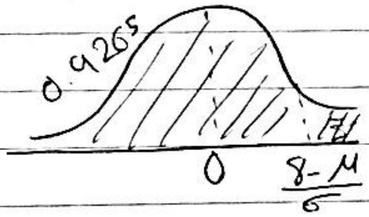


$\frac{2-\mu}{\sigma} = -0.5$

$2 - \mu = -0.5\sigma$

$\mu - 0.5\sigma = 2$ --- ①

$p(z \leq \frac{8-\mu}{\sigma}) = 0.9265$



$\frac{8-\mu}{\sigma} = 1.45 \rightarrow 8 - \mu = 1.45\sigma$

$\mu + 1.45\sigma = 8$ --- ②

$-\cancel{\mu} + 0.5\sigma = 2$

$\cancel{\mu} + 1.45\sigma = 8$

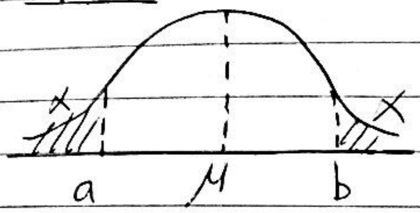
$1.95\sigma = 6$

$\boxed{\sigma = 3.07}$

$\mu + 1.45(3.07) = 8$

$\boxed{\mu = 3.54}$

Note:-



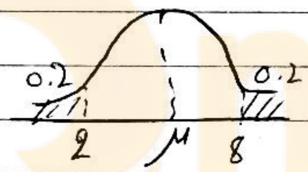
If $p(X \le a) = p(X \ge b) = \alpha$, then:-

$$\boxed{\mu = \frac{a+b}{2}}$$

e.g) $X \sim n(\mu, \sigma^2)$

If $p(X \le 2) = 0.2$ Find μ & σ
$p(X \ge 8) = 0.2$

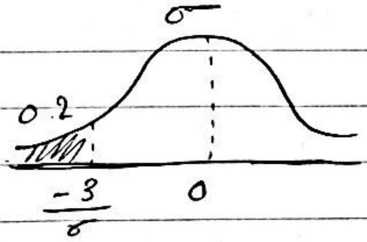
Sol.)



$$\mu = \frac{2+8}{2} = 5$$

now, $p\left(Z \le \frac{2-5}{\sigma}\right) = 0.2$

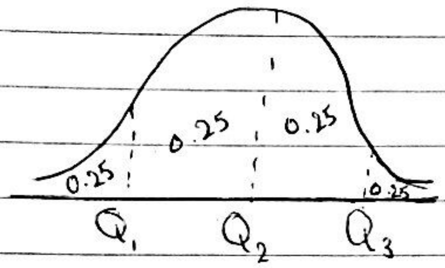
$$p\left(Z \le \frac{-3}{\sigma}\right) = 0.2$$



$$-\frac{3}{\sigma} = -8.84$$

$$\sigma = 3.57$$

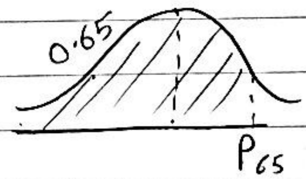
Note:-



(i) $p(X < Q_1) = 0.25$

(ii) $p(X < Q_3) = 0.75$

(iii) $p(X < P_k) = k\%$ $\longrightarrow$



e.g.) $X \sim N(10, 4)$, Find:-

the mean, the mood, Median (Q_2), standard deviation, varience, IQR & P_{85}

Sol.) mean $= \mu = 10 =$ mode $= Q_2$

$\sigma = 2 \quad \sigma^2 = 4$

Now, $p(X < Q_1) = 0.25$

$p\left(Z < \frac{Q_1 - 10}{2}\right) = 0.25$

$Q_1 = \lambda$ إذا بالآلة

$\frac{Q_1 - 10}{2} = -0.67$

$Q_1 - 10 = -1.34$

$Q_1 = 8.66$

$p(X<Q_3) = 0.75$

$p\left(Z < \frac{Q_3 - 10}{2}\right) = 0.75$

$\frac{Q_3 - 10}{2} = 0.67$

$Q_3 = 11.34$

* $IQR = Q_3 - Q_1$
$= 11.34 - 8.66 = 2.68$

Now, $p(X < P_{85}) = 0.85$

$p\left(Z < \frac{P_{85} - 10}{2}\right) = 0.85$

$\frac{P_{85} - 10}{2} = 1.04$

$P_{85} - 10 = 2.08$

$P_{85} = 12.08$

e.g. p. 126 | Suppose that the grades of a general exam are normally distributed with mean 68 & Standard deviation 10. Find:

(a) the proportion of grades that are more than 85,

(b) the proportion of grades between 60 & 90

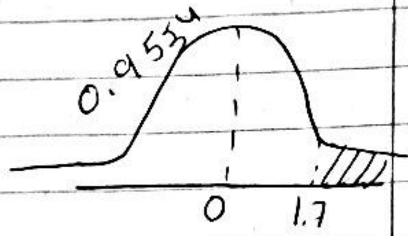
(c) the 95th percentile

(99)

Sol.) $X \sim n(68, 100)$

(a) $p(X > 85) = p\left(Z > \frac{85 - 68}{100}\right)$

$= p(Z > 1.7)$



$= 1 - 0.9554 = 0.0445$

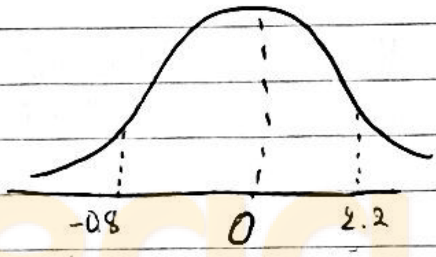
(b) $p(60 < X < 90) = p\left(\frac{60 - 68}{10} < Z < \frac{90 - 68}{10}\right)$

$= p(-0.8 < Z < 2.2)$

$= p(Z \leq 2.2) - p(Z \leq -0.8)$

$= 0.9861 - 0.2119$

$= 0.7742$



(c) $p(X \leq P_{95}) = 0.95$

$p\left(Z \leq \frac{P_{95} - 68}{10}\right) = 0.95$

$\frac{P_{95} - 68}{10} = 1.7$

$17 = P_{95} - 68 \longrightarrow P_{95} = 85$

* The normal approximation to the Binomial distribution:-

If $X \sim Bin(n, p)$

$n \geq 30$ & $p < 0.1$, then

$X \sim poi(np)$

but $n \geq 30$ & p is moderate $(p \geq 0.1)$ then,

$X \sim n(\mu, \sigma^2)$,

where $\mu = np$ $\sigma^2 = npq$ $\sigma = \sqrt{npq}$

* Continuity correction:-

$X = 1 \longrightarrow X \in (0.5, 1.5)$

$X = 2 \longrightarrow X \in (1.5, 2.5)$

$\vdots$

e.g) $X \sim Bin(100, 0.2)$, Find approximate values for the following:-

ⓐ $p(X < 26)$ ⓑ $p(18 < X \leq 26)$

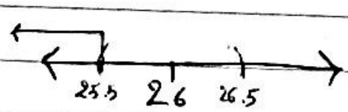
ⓒ $p(18 \leq X < 26)$ ⓓ $p(18 \leq X \leq 26)$

Sol.) since $n \geq 30$ & $p \geq 0.1$

$X \sim n(100(0.2), 100(0.2)(0.8))$

$X \sim n(20, 16)$

ⓐ $p(X < 26) = p(X < 25.5)$

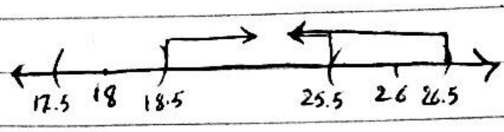


$= p(Z < \frac{25.5 - 20}{4})$

$= p(Z < 1.38)$

$= 0.9162$

(b) $p(18 \leq X \leq 26)$



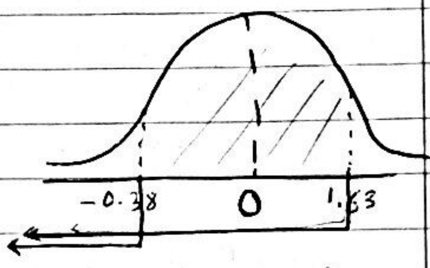
$p(18.5 \leq X \leq 26.5)$

$p\left(\frac{18.5-20}{4} \leq Z \leq \frac{26.5-20}{4}\right) = p(-0.38 \leq Z \leq 1.63)$

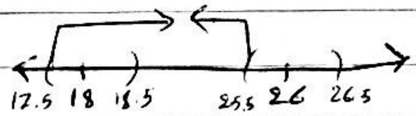
$= p(Z \leq 1.63) - p(Z \leq -0.38)$

$= 0.9452 - 0.3446$

$= 0.6006$



(c) $p(18 \leq x < 26)$



$p(17.5 \leq x < 25.5) =$

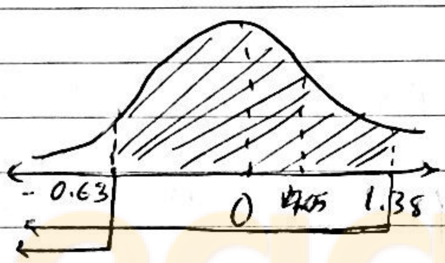
$p(\cancel{25}\, x < 25.5) - p(x < 17.5)$

$p\left(Z < \frac{25.5-20}{4}\right) - p\left(Z < \frac{17.5-20}{4}\right)$

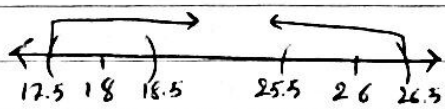
$p(Z \leq 1.38) - p(Z \leq -0.63)$

$= 0.9192 - 0.2743$

$= 0.6449.$



(d) $p(18 \leq X \leq 26)$



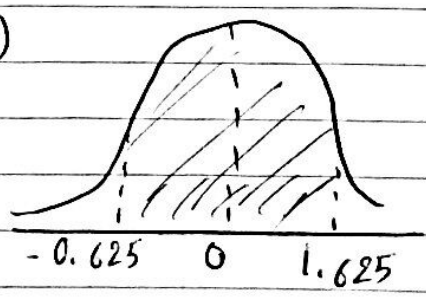
$p(17.5 \leq X \leq 26.5) = p(X \leq 26.5) - p(X \leq 17.5)$

$= p\left(Z \leq \frac{26.5-20}{4}\right) - p\left(Z \leq \frac{17.5-20}{4}\right)$

$= p(Z \leq 1.625) - p(Z \leq -0.625)$

$= 0.9452 - 0.2743$

$= 0.6709.$



e.g) $X \sim Bin(10, 0.5)$

Find $p(x \in \{\cancel{1}, 2, 3, 4\})$ using:-

(a) the Binomial table.

(b) the normal approximations.

(a) $p(X \in \{\cancel{1}, 2, 3, 4\})$

$= p(2 \leq X \leq 4) = p(X \leq 4) - p(X \leq 1)$

$= 0.377 - 0.011$

$= 0.366.$

(b) $X \sim N(5, 2.5)$

$p(2 \leq X \leq 4) = \cancel{p(X \leq 4) - p(X \leq}$

$= p(X \leq 4.5) - p(X \leq 1.5)$

$= p\left(Z \leq \frac{4.5-5}{1.58}\right) - p\left(Z \leq \frac{1.5-5}{1.58}\right)$

$\cancel{= p(Z \leq -2.22) - p(Z \leq -0.32)}$

$= p(Z \leq 0.316) - p(Z \leq -2.22)$

$= 0.3821 - 0.0139$

$= 0.3682.$

e.g p:130 | Suppose that 10% of heavy smokers will suffer from lung cancer after the age of forty. In a sample of 100 heavy smokers, what is the prob. that:-

(a) at least 12 will have lung cancer
(b) not more than 14 will have lung cancer
(c) exactly 12 will have lung cancer.

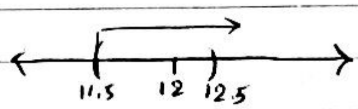
Sol.) $X \sim Bin(100, 0.1)$

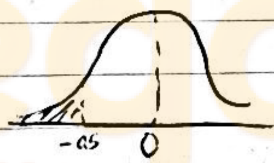
$X \approx n(10, 9)$

$\mu = np = 10$
$\sigma^2 = npq = 10(0.9) = 9.$
$\sigma = \sqrt{9} = 3.$

(a) $p(X \geq 12) = p(X \geq 11.5)$

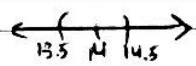
$= 1 - p(X < 11.5)$
$= 1 - p\left(Z < \frac{11.5 - 10}{3}\right)$
$= 1 - p(Z < 0.5)$
$= 1 - \cancel{0.0668}\ 0.3085)$
$= 0.6914.$

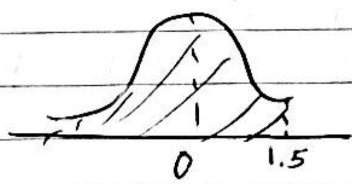




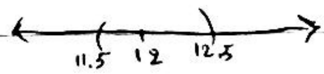
(b) $p(X \leq 14) = p(X \leq 14.5)$

$= p\left(Z \leq \frac{14.5 - 10}{3}\right)$
$= p(Z \leq 1.5) = 0.9332.$



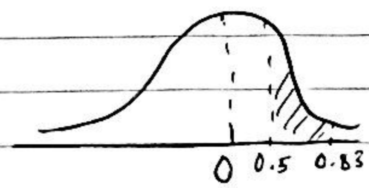


(c) $p(X = 12) = p(11.5 \leq X \leq 12.5)$



$= p\left(\frac{11.5 - 10}{3} < Z < \frac{12.5 - 10}{3}\right)$

$= p(0.5 \leq Z \leq 0.83)$
$= p(Z \leq 0.83) - p(Z \leq 0.5)$



$= 0.7881 - 0.6915$
$= 0.0966.$

(04)

Central limit theorem:-

If $x_1, x_2, x_3, \ldots\ldots x_n$ are drawn from a population whose mean is μ & Variance σ^2, then for larg n; the Sample mean $(\bar{X})$ is apparantly follows a normal distribution with mean μ & Variance $\frac{\sigma^2}{n}$

$$\bar{X} \sim n\left(\mu, \frac{\sigma^2}{n}\right)$$

e.g) suppose that a random sample of size $n=100$ is drawn from a population with mean 70 & standard deviation 20. what is the probability that the sample mean $(\bar{x})$ will be:-

(a) more than 70

(b) less than 73

Sol.) $\mu = 70$

$\sigma = 20$

$\sigma^2 = 400$

$n = 100$

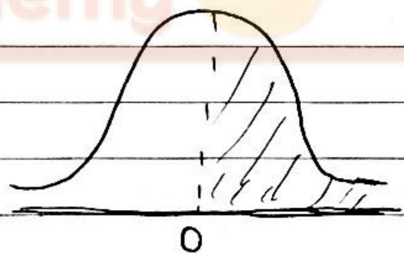
$$\bar{x} \sim n\left(70, \frac{400}{100}\right)$$

$$\bar{x} \sim n(70, 4)$$

(a) $p(\bar{x} > 70)$

$p\left(z > \frac{70 - 70}{2}\right) = p(z > 0)$

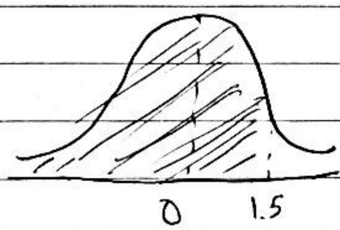
$= 0.5$



(b) $p(\bar{x} < 73)$

$p\left(z < \frac{73 - 70}{2}\right) = p(z < 1.5)$

$= 0.9332$



e.g | p.134 | Suppose that the mean weight & standard deviation of orange boxes are 10 & 2 kg, respectively. If 100 boxes are to be loaded in a car with threshold 1000 kg. what is the probability that the car will break down?.

Sol.)

$\mu = 10$, $\sigma = 2$, $\sigma^2 = 4$, $n = 100$; $x_1, x_2, x_3, \ldots, x_{100}$

find $p\left(\sum_1^{100} x_i > 1000\right)$

$p\left(\frac{\sum_1^{100} x_i}{100} > \frac{1000}{100}\right)$

$p(\bar{X} > 10) = \cancel{p(z > 10)}$ But $\bar{X} \sim n(10, \frac{4}{100})$

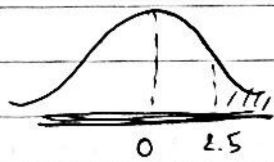
$\bar{X} \sim n(10, 0.04)$

$p\left(z > \frac{10-10}{\sqrt{0.04}}\right) = p\, z > 0$

$= 0.5$

e.g) Solve the previous example if the threshold is 1050 kg:

Sol.) $p\left(\sum_{i=1}^{100} x_i > 1050\right)$



$p(\bar{X} > 10.5)$

$p\left(Z > \frac{10.5 - 10}{\sqrt{0.04}}\right) = p(Z > 2.5)$

$= 1 - p(Z \le 2.5) = 1 - 0.993$

$= 0.0062$

Q9 | p.140 | Ex

$X \sim n(10, 16)$ If X & Y are independent,

$Y \sim n(8, 9)$ then compute $p(X > Y)$

Sol.) $p(X > Y) = p(X - Y > 0)$

let $w = X - Y \sim n(\mu, \sigma^2)$

$\mu = E(w) = E(x-y) = E(x) - E(y)$

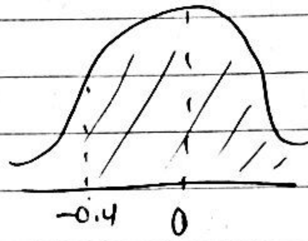
$= 10 - 8 = 2$ $\therefore \boxed{\mu = 2}$

$$\sigma^2 = Var(w) = Var(x-y) = Var(x) + Var(y) + \cancel{2cov(x,y)}$$
$$= 16 + 9 = 25$$
$$\sigma^2 = 25$$

$$W = X - Y \sim n(2, 25)$$

$\therefore\ p(w > 0)$

$$p\left(Z > \frac{0-2}{5}\right) = p(Z > -0.4)$$



$$= 1 - p(Z \leq 0.4)$$
$$= 1 - 0.3446 = 0.6554.$$

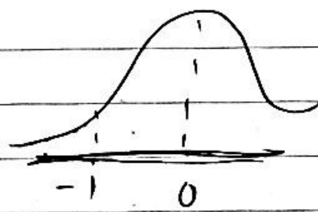
Q8 P:170	$X \sim n(10, 4)$ $Y \sim n(8, 16)$	If X & Y are independent, then compute $p(\underbrace{3X-2Y}_{w} > 4)$

Sol.) let $w = 3X - 2Y \sim n(\mu, \sigma^2)$

$$\mu = E(w) = E(3X - 2Y)$$
$$= 3E(X) - 2E(Y)$$
$$= 3(10) - 2(8) = 30 - 16 = \boxed{14 = \mu}$$

$$\sigma^2 = Var(3X - 2Y) = 9\,Var(X) + 4\,Var(Y) + \cancel{12\,cov(X,Y)}$$
$$= 9(4) + 4(16) = \boxed{100 = \sigma^2}$$

then $w \sim n(14, 100)$ $\sigma = 10$.

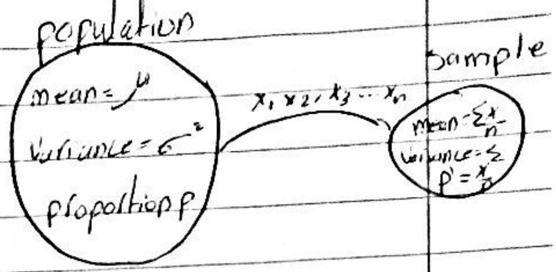


$$p(w > 4)$$
$$p\left(Z > \frac{4-14}{10}\right) = p(Z > -1)$$

$$= 1 - p(Z \leq -1)$$
$$= 1 - 0.1587$$
$$= 0.8413$$

(107)

Ch5: Sampling distributions



[1] $\bar{X} \sim n(\mu, \frac{\sigma^2}{n})$ for $n \geq 30$

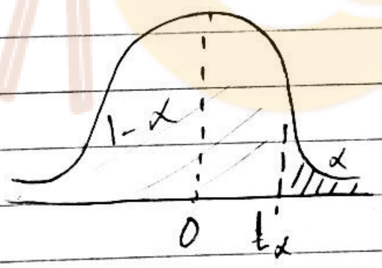
$$\frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim n(0,1)$$

[2] If σ^2 is unknown, then we can use S^2, but in this case,

$$\frac{\bar{X} - \mu}{\frac{s}{\sqrt{n}}} \sim t_{(n-1)}$$ "t-distribution with $(n-1)$ degrees of freedom (d.f)"

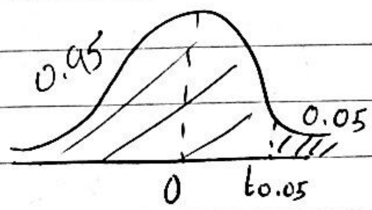
Note:- If $n \geq 30$, then $\frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} \sim n(0,1)$

t-distribution



→ جداول t-distribution
تقرأ المساحة على يمين القيمة
عكس ال normal

e.g) $P(t \leq 1.860) = 0.95$
with df = 8



e.g) Suppose that the weights of newborn babies are normally distributed with mean 3kg! A random sample of size 10 is taken & showed that it's standard deviation is 2.

(a) Find the prob. that the sample average is below 4.16 kg

(b) what is the 90^{th} percentile of the distribution of $\bar{X}$

Sol.) $\mu = 3$ kg

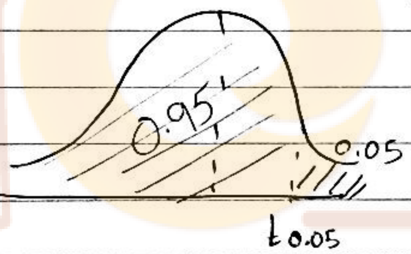
$n = 10 \quad S = 2$

(a) $p(\bar{X} < 4.16)$

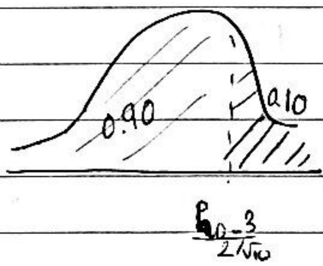
$p\left(\frac{\bar{X} - 3}{2/\sqrt{10}} < \frac{4.16 - 3}{2/\sqrt{10}}\right) = p(t < 1.834)$

with d.f $= n-1 = 10-1 = 9$

$= 0.95$



(b) $P_{90} \Rightarrow p(X < P_{90}) = 0.90$



$p\left(t < \frac{P_{90} - 3}{2/\sqrt{10}}\right) = 0.9$

d.f = 9.

$\frac{P_{90} - 3}{2/\sqrt{10}} = 1.383$

$P_{90} = 3.87$

* The distribution of the sample variance (S^2)

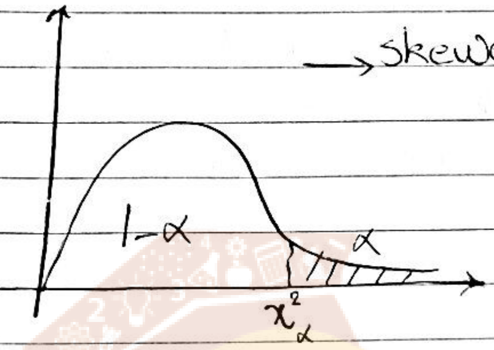
$$\underbrace{\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)}}$$

μ ل
مجهول

↳ chi-square distribution with $(n-1)$ d.f.

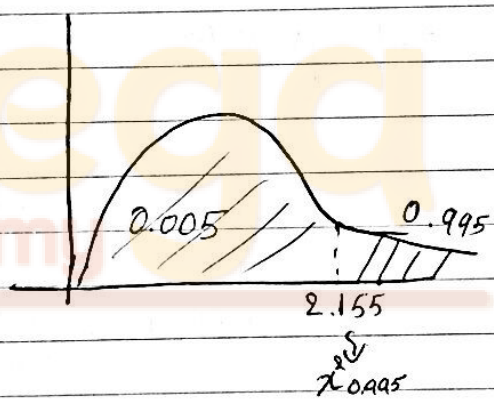
χ^2 - distribution:

→ skewed to the right.



e.g) $p(\chi^2 < 2.155) = 0.005$

$df = 10$

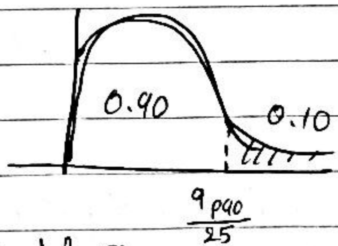


Let $X_1, X_2, \dots, X_{10} \sim n(\mu, \overset{\sigma^2}{25})$, If S^2 is the Sample variance, Find the 90^{th} percentile of S^2.

Sol.) $p(S^2 < P_{90}) = 0.90$

$p\left(\frac{(n-1)S^2}{\sigma^2} < \frac{(n-1)P_{90}}{\sigma^2}\right) = 0.90$

$p\left(\chi^2 < \frac{9\,P_{90}}{25}\right) = 0.90$ with d.f. = 9.



$\frac{9\,P_{90}}{25} = 14.6837 \rightarrow P_{90} = 40.7881$

* The distribution of the sample proportion ($\hat{p}$):-

For $n \geq 30$, then

$$\hat{p} \sim n\left(p, \frac{p(1-p)}{n}\right) \quad \text{OR} \quad \frac{\hat{p}-p}{\sqrt{\frac{p(1-p)}{n}}} \sim n(0,1)$$

e.g) Suppose that 10% of a certain production are defectives. If 400 items are drawn from the production, what is the prob. that the sample proportion will be:

(a) more than 12%
(b) btw 9% & 11%

Sol.) $p = 0.10$
$n = 400$

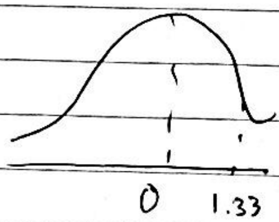
$$\hat{p} \sim n\left(p, \frac{p(1-p)}{n}\right)$$

$$\hat{p} \sim n\left(0.10, \frac{0.09}{400}\right) = \hat{p} \sim n\left(0.10, \frac{9}{40000}\right)$$

(a) $p(\hat{p} > 0.12) = p\left(Z > \frac{0.12 - 0.10}{3/200}\right)$

$= p(Z > 1.33)$

$= 1 - 0.9082 = 0.0918$



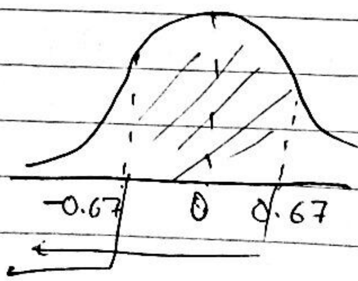
(b) $p(0.09 < \hat{p} < 0.11)$

$p\left(\frac{0.09 - 0.1}{3/200} < Z < \frac{0.11 - 0.10}{3/200}\right)$

$p(-0.67 < Z < 0.67)$

$= 0.7486 - 0.2514$

$= 0.4972$



To sum up:-

① $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim n(0,1)$ for $n \geq 30$.

② $\frac{\bar{X}-\mu}{s/\sqrt{n}} \sim t_{(n-1)}$ $\Rightarrow$ if σ^2 is unknown

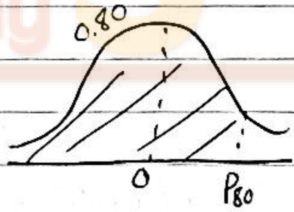
③ $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)}$ $\longrightarrow$ when μ is unknown

④ $\hat{p} \sim n\left(p, \frac{p(1-p)}{n}\right)$ OR $\frac{\hat{p}-p}{\sqrt{\frac{p(1-p)}{n}}} \sim n(0,1)$.

P 164 Q1 | $X \sim n(50,100)$
$y \sim t(15)$
$w \sim \chi^2(10)$. Find

(a) 80^{th} percentile(X) (b) 10^{th} percentile of (y) (c) 90^{th} percentile of w

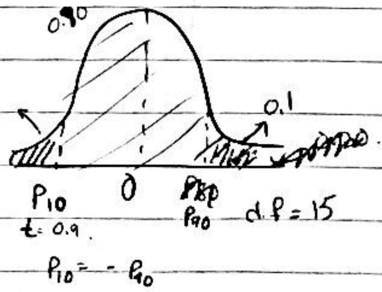
Sol.) (a) $p(X < \cancel{80}\,P_{80}) = 0.80$



$p\left(Z < \frac{P_{80}-50}{10}\right) = 0.80$

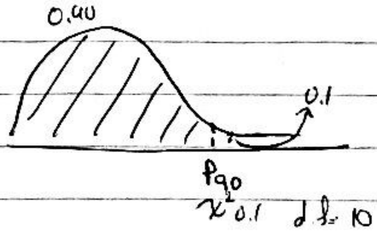
$\frac{P_{80}-50}{10} = 0.84 \rightarrow P_{80} = 58.4$

(b) $p(y < P_{10}) = 0.1$



$P_{10} = -1.341$ $P_{90} = 1.341$.

(c) $P_{90} = 15.9871$



P 164 | Q4 | $X \sim \chi^2(10)$ Find:

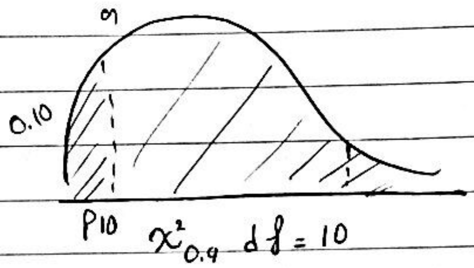
(a) the 10^{th} percentile of X

(b) the 90^{th} percentile of X

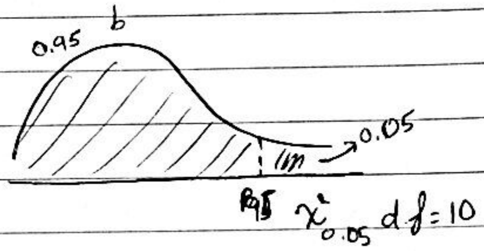
(c) the 99^{th} percentile of X

Sol.)

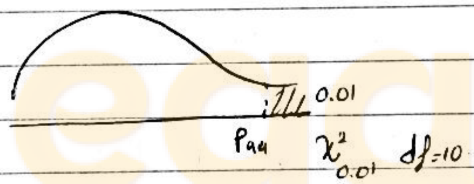
(a) $P_{10} = 4.86518$



(b) $P_{95} = 18.307$



(c) $P_{99} = 23.2093$

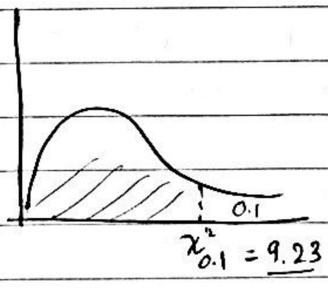


P164 | Q10 | Let $X_1, X_2, \dots X_6 \sim n(\mu, 9)$
find $p(S^2 < 16.63)$

Sol.) $p\left(\frac{(n-1)S^2}{\sigma^2} < \frac{(n-1)(16.63)}{\sigma^2}\right)$

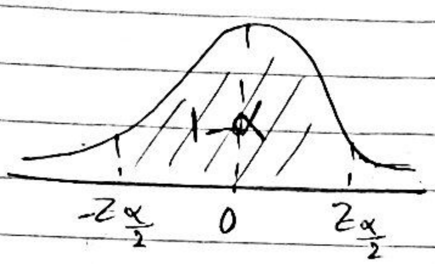
$p\left(\chi^2(5) < \frac{(6-1)(16.63)}{9}\right)$

$p(\chi^2(5) < 9.23) = \underline{0.9}$



Note :- The point estimation of μ is $\bar{x}$
- The point estimation of σ^2 is S^2
- The point estimation of p is $\hat{p}$

Confidence interval for μ:-



α: signification level

$1-\alpha$: confedence level

$Z = \dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim n(0,1)$ $\longrightarrow$

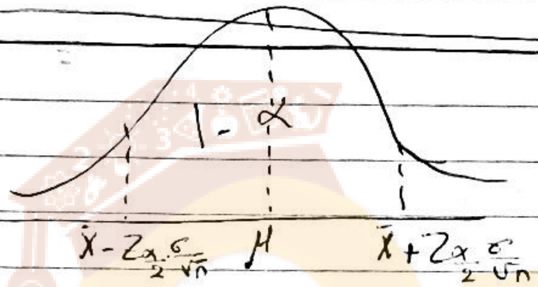
$P(-Z_{\frac{\alpha}{2}} < Z < Z_{\frac{\alpha}{2}}) = 1-\alpha$

$P(-Z_{\frac{\alpha}{2}} < \dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}} < Z_{\frac{\alpha}{2}}) = 1-\alpha$

$P(-Z_{\frac{\alpha}{2}} \cdot \sigma/\sqrt{n} < \bar{X}-\mu < Z_{\frac{\alpha}{2}} \cdot \sigma/\sqrt{n})$

$P(+Z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} > \mu-\bar{X} > Z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}) = 1-\alpha$

$P(\bar{X}+Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} > \mu > \bar{X}-Z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}) = 1-\alpha$



$\longrightarrow$ $(1-\alpha)100\%$ C.I $\quad \mu \in (\bar{X}-E, \bar{X}+E)$

E: error $= Z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$, where $\frac{\sigma}{\sqrt{n}}$ standard error

Note: If σ^2 is unknown, then the $(1-\alpha)100\%$ C.I for μ is given by:-

$(\bar{X}-t_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}}, \bar{X}+t_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}})$.

But if n is larg $(n \geq 30)$ then the $(1-\alpha)100\%$ C.I given for μ is:

$(\bar{X}-Z_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}}, \bar{X}+Z_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}})$

e.g) the salaries of teachers, in Jordan for 1990 - 2000 Are normally distributed with standard deviation 50 JD. The average Salary based on a sample of 400 teachers for 1990 - 2000 was 215 JD per month.

(a) what is the point estimate of the mean salaries & its S.E?

(b) Give a 90% C.I for the mean μ.

(c) Give a 95% C.I for the mean μ.

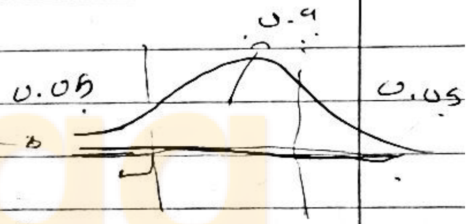
Sol.) $\bar{X} = 215$ ~~$X \sim N(215,$~~

$n = 400$

$\sigma = 50$

(a) the point estimate of μ is $\bar{X} = 215$.

$S.E(\mu) = \frac{\sigma}{\sqrt{n}} = \frac{50}{20} = 2.5$



(b) $1 - \alpha = 0.90$

$\alpha = 0.1$

$\frac{\alpha}{2} = 0.05$

$Z_{\frac{\alpha}{2}} = 1.64$

$L = \bar{X} - Z_{\frac{\alpha}{2}} S.E$

$= 215 - 1.64(2.5) = 210.9$

$U = \bar{X} + Z_{\frac{\alpha}{2}} S.E$

$= 215 + 1.64(2.5) = 219.1$

90% C.I is $(210.9, 219.1)$

* we are 90% confident that the salaries average (μ) is btw $(210.9, 219.1)$ JD

(c) $1 - \alpha = 0.95$

$\alpha = 0.05$

$\frac{\alpha}{2} = 0.025$

$Z_{\frac{\alpha}{2}} = 1.96$

$L = \bar{X} - Z_{\frac{\alpha}{2}} S.E$

$= 215 - 1.96(2.5) = 210.1$

$U = \bar{X} + Z_{\frac{\alpha}{2}} (S.E)$

$= 215 + 1.96(2.5) = 219.9$

95% C.I is $(210.1, 219.9)$

* we are 95% confident that the salaries average (μ) lies in $(210.1, 219.9)$ JD

e.g)| The mean cholestrol levels in a general population are
P205 | normally distributed.

A sample of 16 persons is taken under a test with sample mean $\bar{x} = 220$ mg/dL & standard deviation $S = 25$ mg/dL. Give a 90% C.I for the population mean μ.

Sol.) $n = 16$, $\bar{x} = 220$, $S = 25$ | $1-\alpha = 0.90$, $\alpha = 0.10$, $\frac{\alpha}{2} = 0.05$ } $df = 16-1 = 15$, $Z_{\frac{\alpha}{2}} = 1.753$

$$S.E = \frac{S}{\sqrt{n}} = \frac{25}{4} = 6.25$$

$$L = \bar{X} - Z_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}} = 220 - (1.753)(6.25) = 209.04$$

$$U = \bar{X} + Z_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}} = 220 + (1.753)(6.25) = 230.96$$

90% C.I is (209.04, 230.96)

we are 90% confident that μ lies in the interval (209.04, 230.96)

Determination of the sample size:-

$$E = Z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$$

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

e.g) A researcher wants to estimate the average weightloss of people who are on a new diet plan. I a previous study, the population standard deviation σ of weight losses is about 5kg. How larg the sample should be to estimate the mean weight loss of by 95% C.I to within 1.5 kg.

Sol.) $\sigma = 5$ kg, C.I = 95%, $E = 1.5$. | $1-\alpha = 0.95$, $\alpha = 0.05$, $\frac{\alpha}{2} = 0.025$, $Z_{\frac{\alpha}{2}} = 1.96$ |

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

$$n = \left(\frac{1.96}{1.5}\right)^2 \cdot 5^2 = 42.68 \simeq 43$$

it is required a sample of 43 persons to have a 95% C.I within 1.5 kg.

$\therefore$ 95% C.I is $(\bar{X} - 1.5, \bar{X} - 1.5)$

e.g) Suppose that jordan Bureau of Census in 2004 wants to estimate the means size μ of all jordan families by 99% C.I. It is known that the standard deviation σ for all sizes of families is 1.5. How large a sample size should the Bureau select to estimate μ within 0.02 of the population mean?

p.181

99% C.I $\quad 1-\alpha = 99 \quad n = \left(\frac{Z_{\frac{\alpha}{2}}}{E}\right)^2 \sigma^2$

$\sigma = 1.5 \quad \alpha = 0.01$

$E = 0.02 \quad \frac{\alpha}{2} = 0.005 \quad = \left(\frac{2.57}{0.02}\right)^2 (1.5)^2 \cong 37153$

$Z_{\frac{\alpha}{2}} = 2.57$

$\therefore$ 99% C.I is $(\bar{X} - 0.02, \bar{X} + 0.02)$

The C.I for the population proportion (p):-

The $(1-\alpha)100\%$ C.I for the population proportion (p) is:-

$$\left(\hat{p} - Z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}, \hat{p} + Z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right)$$

e.g It was believed in the Arab world that 50% of persons are smoking. During the year 2000, a sample of 1000 persons showed that the number of smokers is 620. Establish 95% C.I for the proportion of smokers

p.204

~~$\hat{p}$ 50%~~

$n = 1000 \quad 1-\alpha = 0.95 \quad L = \hat{p} - Z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

$\hat{p} = \frac{620}{1000} = 0.62 \quad \alpha = 0.05 \quad = 0.62 - 1.96(0.0153)$

$\frac{\alpha}{2} = 0.025 \quad = 0.590$

$Z_{\frac{\alpha}{2}} = 1.96$

$U = \overset{0.62}{\hat{p}} + 1.96(0.0153)$

$S.E = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = \sqrt{\frac{0.62(0.38)}{1000}} \quad = 0.650$

$= 0.0153 \quad$ 90% C.I is $(0.59, 0.65)$

So

more than it was thought

* Determination of the sample size :-

$$E = Z_{\frac{\alpha}{2}} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$$\frac{E}{Z_{\frac{\alpha}{2}}} = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \longrightarrow \left(\frac{E}{Z_{\frac{\alpha}{2}}}\right)^2 = \frac{\hat{p}(1-\hat{p})}{n} \longrightarrow n = \left(\frac{Z_{\frac{\alpha}{2}}}{E}\right)^2 \cdot \left(\hat{p}(1-\hat{p})\right)$$

Note: $\hat{p}$ depends on the sample size, but we dont have the sample size yet so we have 2 cases:-

Case I :- $\hat{p}$ is known from a previous study, then we use it to find n

Case II :- $\hat{p}$ is unknow from a previous study so we use $\hat{p} = \frac{1}{2}$. Thus $\therefore\ n \leq \frac{1}{4}\left(\frac{Z_{\frac{\alpha}{2}}}{E}\right)^2$.

e.g) Assume that it is required to estimate the proportion of patients suffering a bad reaction from taking a medication by 95% C.I. Determine the sample size needed if the error of estimation is about 0.01 in the following cases:-

(a) no prior information about p.

(b) previous study that showed that p is approximately 0.20

(a) $E = 0.01$ $Z_{\frac{\alpha}{2}} = 1.96$, $\hat{p} = \frac{1}{2}$

$$n = \frac{1}{4}\left(\frac{1.96}{0.01}\right)^2 \simeq 97$$

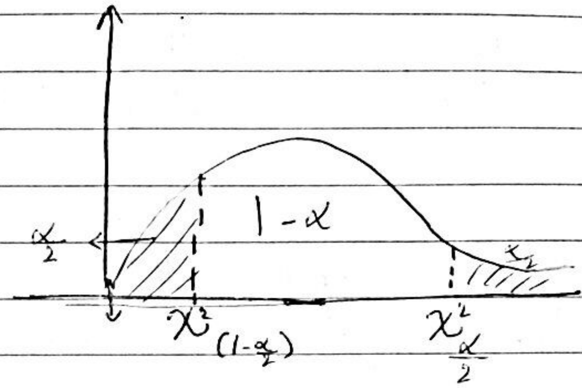
(b) $E = 0.01$ $Z_{\frac{\alpha}{2}} = 1.96$, $\hat{p} = 0.20$

$$n = 0.2\,(0.8)\left(\frac{1.96}{0.01}\right)^2 \approx 62.$$

* The confidence interval for σ^2 :-

The $(1-\alpha)100\%$ C.I for the population variance:- σ^2 is :-

$$\left(\frac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2}}} , \frac{(n-1)S^2}{\chi^2_{(1-\frac{\alpha}{2})}} \right)$$



eg
P213

Quality control engineer wishes to study the weight variation of a new product, a sample of (10) items is taken & provided $\bar{X} = 0.6$ kg , $S = 0.4$ kg . Find 90% C.I. for the variance of all items :-

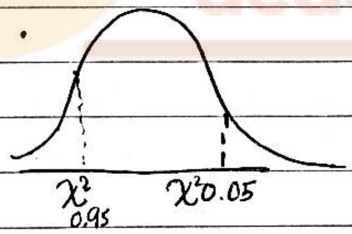
Sol.) $n = 10$ $\quad d.f = 9$

$\bar{X} = 0.6$, $S = 0.4$

$1 - \alpha = 0.90$

$\alpha = 0.10$

$\frac{\alpha}{2} = 0.05$



$\chi^2_{0.05} = 16.919$

$\chi^2_{0.95} = 3.325$

$$L = \frac{(n-1)s^2}{\chi^2_{0.05}} = \frac{9(0.4)^2}{16.919} = 0.09$$

$$U = \frac{(n-1)s^2}{\chi^2_{.95}} = \frac{9(0.4)^2}{3.325} = 0.43$$

then, 90% C.I of σ^2 is:- $(0.09 , 0.43)$.

(119)

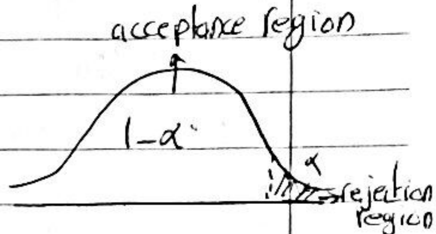
(1)

* Test of hypotheses:-

* we have 2 types of hypotheses:-

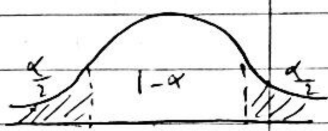
① null hypothesis (H_0):
A standard about a population parameter (μ or σ^2 or p.. etc)
this is assumed right until it is declared false

② Alternative hypothesis (H_1):-
Hypothesis that is true when H_0 is false.



e.g) Test the hypotheses:-

(a) $H_0 : \mu = 200$ Vs $H_1 : \mu < 200$
(b) $H_0 : \mu = 200$ Vs $H_1 : \mu > 200$ } 1-tailed test
(c) $H_0 : \mu = 200$ Vs $H_1 : \mu \neq 200$ ⟶ 2-tailed test



* Types of erors:-

	H_0 (true)	H_1 (true)
Reject H_0	Type (I) eror رفض نظرية صحيحة	✓
Accept H_0	✓	Type II eror قبول نظرية خاطئة

Note: probability of Type I eror is:
$\alpha = p(\text{reject } H_0 / H_0 \text{ is true})$
α is called the significent level
probability of type II eror is:
$\beta = p(\text{Accept } H_0 / H_1 \text{ is true})$
$1-\beta$: power of the test.

Note: test statistic is a quantity measured to decide whether to accept or reject H_0.
If the test statistic lies in the acceptance region, then we accept H_0, otherwise, we reject H_0

* Inference About the sample mean (μ):

1. If σ is known, then the test statistic is:

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$$

2. (a) If σ is unknown & n is small, then the test statistic is:

$$t = \frac{\bar{X} - \mu}{S/\sqrt{n}}$$

(b) If σ is unknown & n is larg ($n \geq 30$) then the test statistc is

$$Z = \frac{\bar{X} - \mu}{S/\sqrt{n}}$$

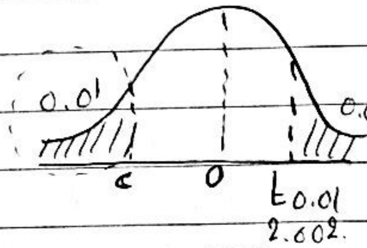
e.g p.205 The mean cholestrol level in a general population are normally distributed. A sample of 16 persons is taken under a test with sample mean $\bar{X} = 220$ & standard deviation $S = 25$

Test a 1% significance level the the mean colestrol is less than 230

$n = 16 \rightarrow df = 15$, $\bar{X} = 220$ $S = 25$

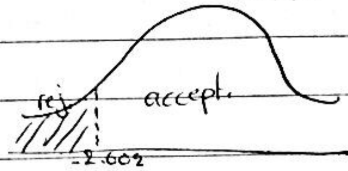
$\alpha = 0.01$

$H_0: \mu = 230$, $H_1: \mu < 230$



test statistic: $t = \frac{\bar{X} - \mu}{S/\sqrt{n}}$

$= \frac{220 - 230}{25/4} = -1.6$



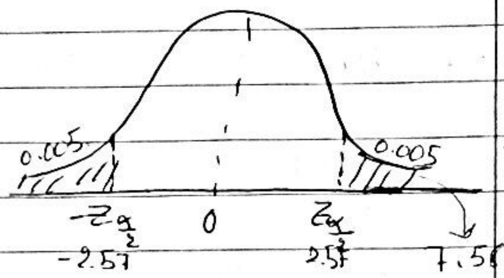
$t_{0.01} = 2.602$. $\rightarrow$ critical value is $c = -2.602$.

Since -1.6 lies in the acceptance region we accept H_0

e.g) It was believed in the arab world that 50% of persons are smoking. During the year 2000, a sample of 1000 persons showed that the no. of smokers is 620, can you conclude that the proportion of smokers is different ~~than~~ from 50%? use $\alpha = 0.01$.

$P = 0.50$

$n = 1000 \rightarrow \hat{p} = \frac{620}{1000} = 0.62$



$H_0: p = 0.05$, $H_1: p \neq 0.50$

$\alpha = 0.01 \rightarrow \frac{\alpha}{2} = 0.005$

$-Z_{\frac{\alpha}{2}} = -2.57$

$Z_{\frac{\alpha}{2}} = 2.57$

test statistic is: $Z = \dfrac{\hat{p} - P_0}{\sqrt{\frac{P_0(1-P_0)}{n}}}$

$Z = \dfrac{0.62 - 0.50}{\sqrt{\frac{0.5(0.5)}{1000}}} = 7.58$ $\rightsquigarrow$ Since test statistic is in the rejection region we reject Ho.

* Test for the population variance σ^2:-

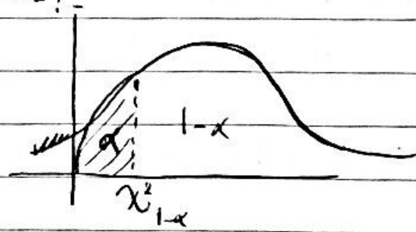
$H_0: \sigma^2 = \sigma_0^2$

H_1: (a) $\sigma^2 < \sigma_0^2$

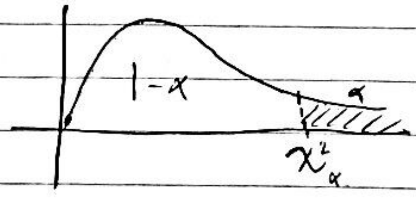
(b) $\sigma^2 > \sigma_0^2$

(c) $\sigma^2 \neq \sigma_0^2$

(a)



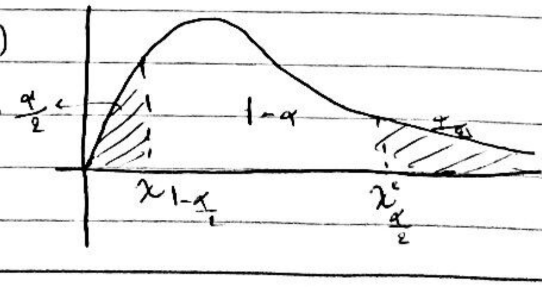
(b)



test statistic

$\chi^2 = \dfrac{(n-1)S^2}{\sigma_0^2}$

(c)



eg) Quality control engineer wishes to study the weight variation of a new product. A sample of 10 items is taken and provided

$\bar{X} = 0.6$ kgs , $S = 0.4$ kgs

(a) Test, $H_0: \sigma^2 = 0.5$ Vs $H_1: \sigma^2 > 0.5$

Use $\alpha = 0.025$

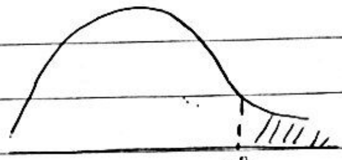
(b) Test $H_0: \sigma^2 = 0.74$ Vs $H_1: \sigma^2 \neq 0.74$.

Use $\alpha = 0.10$

Sol.) (a) $\bar{X} = 0.6$ $\alpha = 0.025$

$S = 0.4$

$n = 10 \longrightarrow df = 9$



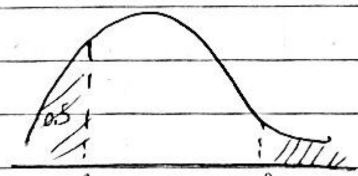
$\chi^2_{0.025} = 19.0228$

test statistic : $\chi^2 = \dfrac{S^2(n-1)}{\sigma^2} = \dfrac{(0.4)^2(9)}{0.5} = 2.88$

we accept H_0

(b) $\alpha = 0.10$ $\frac{\alpha}{2} = 0.05$

$df = 9$



test statistic : $\chi^2 = \dfrac{S^2(n-1)}{\sigma^2}$

$= \dfrac{0.4(9)}{\cancel{0.05}(0.74)^2} = 2.62$

we reject H_0.

Summarys:- Chap 5+6

1 sampling distributions:-

1. $\mu, \sigma, \bar{x}, n \longrightarrow \bar{x} \sim n(\mu, \frac{\sigma}{n})$
2. $\mu, S, n \longrightarrow \frac{\bar{x} - \mu}{S/\sqrt{n}} \sim t_{(df)}$
3. $\sigma, s, n \longrightarrow \frac{s(n-1)}{\sigma^2} \sim \chi^2 df$
4. $p, \hat{p}, n \longrightarrow \hat{p} \sim n(p, \frac{p(1-p)}{n})$.

Confidence interval:- CI

if σ is known $\longrightarrow (\bar{x} - Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}, \bar{x} + Z_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}})$.

if σ is unknown $\xrightarrow{n<30} (\bar{x} - t_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}}, \bar{x} + t_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}})$

" " $\xrightarrow{n \geq 3} (\bar{x} - Z_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}}, \bar{x} + Z_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}})$

propotion $(p, \hat{p}) \longrightarrow (\hat{p} - \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \cdot Z_{\frac{\alpha}{2}}, \hat{p} + \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \cdot Z_{\frac{\alpha}{2}})$.

for σ / μ is unknown $\longrightarrow \left(\frac{(n-1)s^2}{\chi^2_{\frac{\alpha}{2}}}, \frac{(n-1)s^2}{\chi^2_{1-\frac{\alpha}{2}}} \right)$.

sample size:-

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

$$n = \left(\frac{Z_{\alpha/2}}{E} \right)^2 \cdot (\hat{p}(1-\hat{p})) \longrightarrow$$
1. ~~$\hat{p}$ is known~~
2. $\hat{p}$ unknown $\longrightarrow \hat{p} = 0.5$.

Testing hypotheses:-

$1-\alpha$: Acceptance region
α : rejection region

Test

Test statistics:-

σ is known: $Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$

σ is unknown: $Z = \frac{\bar{x} - \mu}{S/\sqrt{n}}$

proportion: $Z = \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}}$

χ^2 (σ): $\chi^2 = \frac{S^2(n-1)}{\sigma^2}$