Statistics. 236

Data :- كل المعلومات والحقائق والأرقام وغيرها التي نقوم بجمعها حول ظاهرة معينة.

is the science of collecting, organizing. presenting. analyzing and interpreting data.
تفسير المعلومات

* تساعدنا الإحصاء باتخاذ القرار المناسب لقضية معينة.

* population :- المجتمع الاحصائي
All elements of interest in a particular

* Sample :- any subset of the population.
نستخدم عينة بدل استخدام جميع العناصر لان ذلك يكون مكلف ويحتاج إلى وقت لذلك نستخدم العينة للسرعة شرط ان تكون عينة تمثيلية أو عشوائية.

→ Descriptive Statistics :- الإحصاء الوصفي
1- table's → Frequency Distribution.
2- Graph's → pie chart
3- Numberical methods → Average

→ Inferential statistics :- الإحصاء الاستنتاجية

Sample → population

المعلومات التي نتمكن من الحصول عليها نقوم بتعميمها

* Variables :- متغيرات تعطي نوعين من المعلومات
1- Quantitative data. :- كمية
Data assume numerical values

Ex :- Age, Income, Average

2- Qualitative data :- نوعية
it is divided into two or more Different categories.
Ex :- Gender "الجنس" → male, → Female

magor → B, → A, → F, → E

* Data Sources :- الحصول على معلومات إما من إحصائيات سابقة او اقوم بعمل دراسة إحصائية.
statistical study.

→ Descriptive stat
→ Inferential stat

* Variable → Quantative → Age, ---
→ Qualitative → Gender ---

* Scales of measurment :-
1- Nominal → qualitative + data can't be ranked
Ex :- majer :- Accounting 1
Econ 2
Busa 3
Fin 4

2- Ordinal → qualitative + data can be ranked

Ex:- letter Grades A
B
C
D
F

3- Interval scale :→ Quantitative data + Difference has a meaning

Ex:- temp 40°, 15°

$40 - 15 = 25^\circ$ has a meaning

4- Ratio scale: → quantative data + Ratio has a meaning

Ex:- Sales

* Organizing Of Data :-

→ raw data → Qualitative Data →

* Table's → Frequency Distribution.

category	# Frequency

(class) Major	# of stud	r.F
Acct	50	0.25
Busa	80	0.4
Eco	30	0.15
Fin	40	0.20
total	200	

Ex:- 15 persons were asked to taste two types of soft drinks A and B and indicate the taste of A was superioer (s) the same (M) or inferior (I) to that of B. the responses are listed below.

S, I, I, M, S, M, M, S, I, S, S, M, M, S, M.

2- Ordinal → qualitative + data can be ranked
Ex:- letter Grades A
B
C
D
F

3- Interval Scale :→ Quantitative data + Differen
has a meaning
Ex:- temp 40°, 15°
40 − 15 = 25° has a meaning

4- Ratio scale: → quantative data + Ratio has a
meaning
Ex:- Sales

* Organizing Of Data :-

→ raw data → Qualitative Data →
* Tables → Frequency Distribution.

category	# Frequency

(class) Major	# of stud	r.F
Acct	50	0.25
Busa	80	0.4
Eco	30	0.15
Fin	40	0.20
total	200	

Ex:- 15 persons were asked to taste two types of soft drinks A and B and indicate the taste of A was superioer (s) the Same (M) or inferior (I) to that of B. the responses are listed below.
S, I, I, M, S, M, M, S, I, S, S, M, M, S, M.

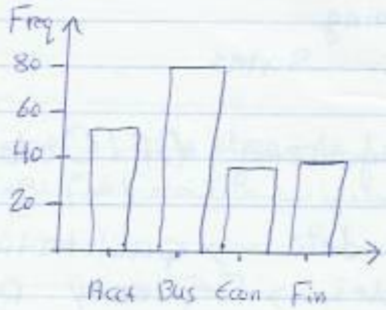
Response	# of persons
S	6
I	3
M	6

* Relative Frequency $= \dfrac{\text{Class (category) Frequency}}{\text{total Frequency}}$

* Graphs :-

1- Bar chart.

2- pie chart.



$360^\circ \rightarrow$ total

Ex:- Acct $\rightarrow$ 200 360°
50   X

Acct $= 200X = (50)(360)$
$= X = \frac{50}{200} = 90^\circ$

$(r.F)(360^\circ)$.

* Organizing and Graphing :- Quantative Data

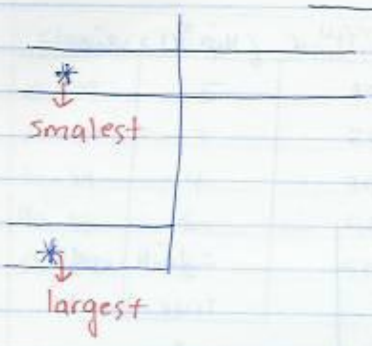
frequency Discribution.

Class Interval	Frequency

1- Class width
2- # of classes
3- Class limits.

lower limit ← 5 - 10 → upper limit

$$\text{class width} = \frac{\text{largest Data value} - \text{Smallest data Value}}{\# \text{ of classes.}}$$



L = lower limit
U = upper limit
$U = L + W - 1$

Ex:- 10 – 19
20 – 29
30 – 39

Ex :- Consider the Following data.

30, 55, 95, 89, 76, 78, 81, 45, 67, 91, 81, 86, 73, 74, 83, 60, 71, 74, 90, 51, 43, 24, 95, 89, 76, 62, 94, 95, 75, 45

construct a frequency discribution with 8 class

$$W = \frac{95 - 24}{8} \approx 9$$

Class	Frequency
20 – 28	1
29 – 37	1
38 – 46	3
47 – 55	2
56 – 64	2
65 – 73	3
74 – 82	8
83 – 91	8
92 – 100	4
	30

Qualitativ data.

* إذا كان عندنا أعشار نقوم بحسابها كتالي :-

$$U = L + W - 0.1$$

لضمان دخول الأعشار.

$W = 5$
$L = 4$
$U = 4 + 5 - 0.1$

* وإذا كانت الأعشار أكثر من منزلة كقولنا 8.55 نقوم بزيادة صفر 0.01

* Graphs

1- Histogram مدرج تكراري

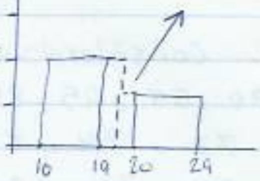
2- polygon مضلع تكراري ، متعدد الأضلاع

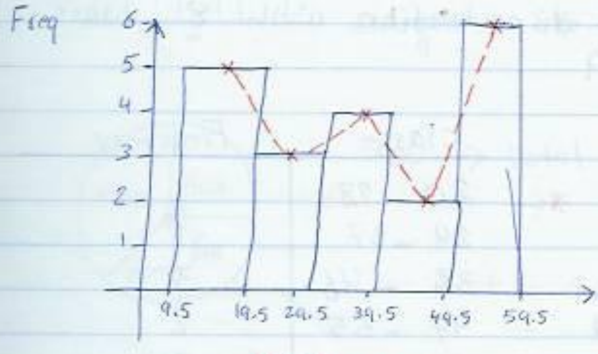
3- ogive منحنى تكراري تراكمي

class	Frequency	true limit
10 - 19	5	9.5 _ 19.5
20 - 29	3	19.5 _ 29.5
30 - 39	4	29.5 _ 39.5
40 - 49	2	39.5 _ 49.5
50 - 59	6	49.5 _ 59.5
total	20	

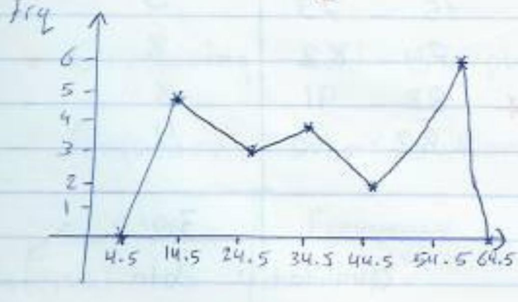
الحدود الفعلية

true limite





Histogram.



polygon

class mark
14.5
24.5
34.5
44.5
54.5

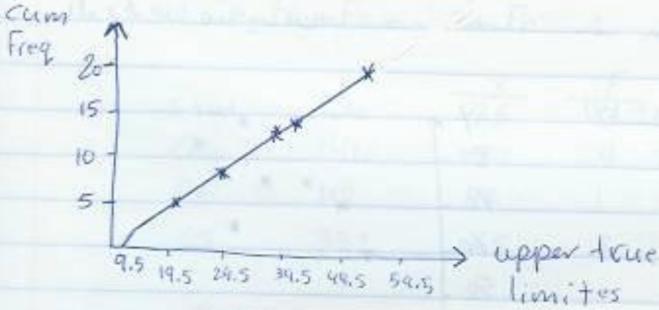
Class mark

class mid point $= \frac{L+U}{2}$

مركز الفئة

* cumulative Frequency Distribution. جدول التوزيع التراكمي التكراري

Class	Freq	upper true limits	cumulative Frequency
10 - 19	5	19.5	5
20 - 29	3	29.5	8
30 - 39	4	39.5	12
90 - 49	2	49.5	14
50 - 59	6	59.5	20
	20		



Ogive

* Stem - and - leaf display.
(ساق) (أوراق)

Ex:- Construct a stem and leaf display For the Following data :- 70, 72, 75, 64, 55, 53, 45, 98, 95, 67, 77, 83, 37, 90, 95, 65, 55, 76

stem ← 3 | 7 ← leaf

3	7
4	5
5	5 3 5
6	4 7 5
7	0 2 5 7 6
8	2
9	5 8 0 5

3	7	1
9	5	1
5	3 5 5	3
6	4 5 7	3
7	0 2 5 6 7	5
8	2	1
9	0 5 5 8	4

* Crosstabulation

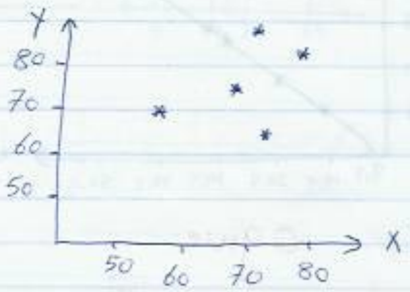
	Gender		row
Department	male	Femal	data
Acc	10	15	25
Bus	25	35	60
Eco	12	8	20
Fin	18	12	30
Col. tot	65	70	135

Colom

* Scatter Diagram :- الربط بين متغيرين

EX :-

	Acc (x)	Stat (y)
1-	75	65
2-	70	75
3.	80	80
4.	55	70
5-	70	85



m = positive

m = Negative.

* Descriptive statistics

→ Numerical Mathod

* Two measures

1- Measures of Centeral Tendency :- مقاييس النزعة المركزية

1- Mean

$\bar{X}$ ≡ sample mean   n ≡ sample size

μ ≡ population mean

[الفصل الثاني]

$\overline{X} = \frac{\Sigma X}{n}$ $n \equiv$ sample size

$\mu = \frac{\Sigma X}{N}$ $N \equiv$ population size

* weighted mean :- الوسط الحسابي الموزون

$X_1, X_2 ---- X_n$
$\downarrow \ \downarrow \ \ \ \ \ \downarrow$
$F_1 \ F_2 \ \ \ \ \ F_n$

$$\overline{w} = \frac{X_1F_1 + X_2F_2 ---- + X_nF_n}{F_1 + F_2 + --- + F_n}$$

Ex:-

		X	F
stat	236	75	3
Eng	141	70	4
PE	112	90	1
CS	332	65	3

$$\overline{w} = \frac{(75)(3) + (70)(4) + (90)(1) + (65)(3)}{3+4+1+3}$$

Ex:- Grouped data

class	F_i Freq	m_i	F_im_i
10-19	5	14.5	58
20-29	4	24.5	98
30-39	3	34.5	103.5
40-49	2	44.5	89
50-59	6	54.5	327

$$\overline{X} = \frac{\Sigma m_iF_i}{\Sigma F_i}$$

2- median الوسيط

The value in the middle of values (after arranging in an increasing order).

* to Find the median.
  1- Arrange the data
  2- Find mediam position.

* two cases :-
  a- # of items is odd n mediam position : $\frac{n+1}{2}$
  b- # of items is even m mediam position $\frac{m}{2}$, $\frac{m}{2}+1$ median ≡ average.

Ex :- Find the mediam

* 10 15 18 12 10 7 6 4 3 11 9

Arrange = 3, 4, 6, 7, 9, 10, 10, 11, 12, 15, 18

\# of items = 11 odd

median position = $\frac{11+1}{2} = 6$

median = 10

* 9, 7, 11, 12, 14, 3, 10, 15

Arrange → 3, 7, 9, 10, 11, 12, 14, 15

\# of items = 8 even

positions = $\frac{8}{2}$, $\frac{8}{2}+1 = (4, 5)$

median = $\frac{10+11}{2} = 10.5$

3- mode :- المنوال → Qualitative data

The most Fiequently occurring value.

Ex :- 10, 15, 18, 10, 12, 10

mode = 10

* يتأثر ال mean بـ extreme values (outliers) القيم الشاذة.

Ex :- wage → 400 , 500 , 600 , 1500

$\bar{X} = 750$

* إذا كان لدينا قيمة (شاذة) عالية جدا ترفع معدل الـ mean أو العكس وهذه الطريقة ليست مجدية لان ال mean يتأثر بالقيم الشاذة فلا يعطي انطباع صحيح لذلك نفضل استخدام الوسيط.

* Trimmed mean

(1) 10% — k% — (1) 10%

* نقوم بحذف القيم الشاذة وعمل الوسط للقيم المتوسطة

Ex :- $\frac{8}{71}$)

25	58	24	50	29	52	57
31	30	41	44	40	46	29
31	37	32	44	49	29	

compute 5% and 10% trimmed mean

5% →

$\cancel{24}, 25, 25, 29 ---- 52, 57, \cancel{58}$

$\left(\frac{5}{100}\right)(20) = 1$

$\bar{X} = \frac{25+25+29 ---- 52,57}{18}$

10% → $\left(\frac{10}{100}\right)(20) = 2$

$\bar{X} = \frac{25+29 ---- 52}{16}$

Data = $\bar{X}$ $\quad 65 = 0.9X + 10$

$X \to Y = \alpha X + B$

$\bar{y} = \alpha \bar{X} + B$

$y = 0.9X + 10$

$\bar{y} = (0.9)(65) + 10$

Ex :- $\overline{X} = 90$

$\frac{1}{2}x + 40 \quad = \quad \frac{1}{2}(90) + 40 = 85$

$X \rightarrow Y = (\alpha X + B)$

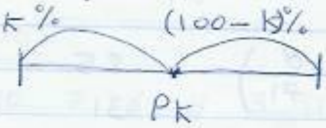
* measures of position :-

→ percentiles :- المئينات

The Kth percentile of a data set is avalue P_K such that :-

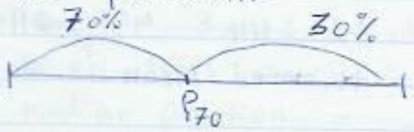
$K\%$ of the data set $\leq P_K$ and

$(100-K)\%$ of the data set $\geq P_K$

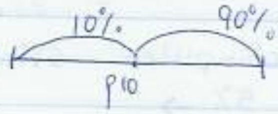


Ex :-

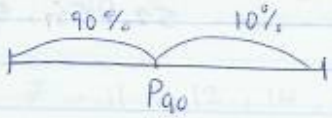
70^{th} percentile



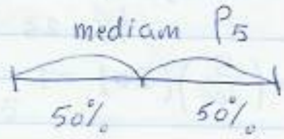
P_{10}



P_{90}



median P_5



* How to Find the K^{th} percentile

1- Arrange the data ↑

2- compute the index i

$i = \left(\frac{K}{100}\right) n \qquad n \equiv$ number of items

3- IF i is an integer (عدد صحيح) ⟶

The K^{th} percentile is the averag of value in position $i, i+1$

Ex :- 1- $n = 10 \qquad P_{50}$

$i = \left(\frac{50}{100}\right) 10 = 5 \qquad$ integer.

2- $n=10$ $\quad P_{65}$

$i = \left(\frac{65}{100}\right)10 = 6.5$ not integer.

4- IF i is not integer. round up نقوم بعملية التقريب ليصبح العدد صحيح

$6.05 \rightarrow 7$

$6.99 \rightarrow 7$

Ex :- consider the Following data.

16, 18, 19, 20, 20, 18, 22, 24, 7, 58, 31, 19

Find P_{25}, P_{30}, P_{60}.

7, 16, 18, 18, 19, 19, 20, 20, 22, 24, 31, 58

* $P_{25} \rightarrow$ $i = \left(\frac{25}{100}\right)12 = 3 \rightarrow$ integer.

$P_{25} \rightarrow$ Average of the two values in position

$= \frac{18+18}{2} = 18$

* $P_{30} \rightarrow$ $i = \left(\frac{30}{100}\right)12 = 3.6$ not integar.

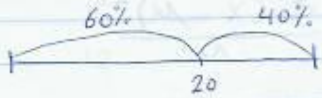
round up to 4

$P_{30} = 18$

* $P_{60} \rightarrow$ $i = \left(\frac{60}{100}\right)12 = 7.2$ not integar

round up = 8

$P_{60} = 20$



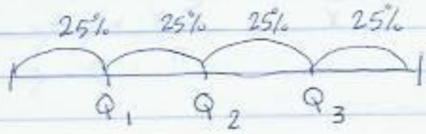
→ Quartiles :- الأرباع

$Q_1 \equiv$ First quartile $= P_{25}$

$Q \equiv$ secound quartile $= P_{50}$

mediam

$Q \equiv 3^{th}$ quartile $= P_{75}$



* things The Five numbers Summary

Smallest , Q_1 , Q_2 , Q_3 , largest

Ex :-

7 , - - - - - - , 58

7 , 18 , $\boxed{}$, $\boxed{}$, 58

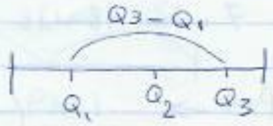
P_{50} P_{75}

* Measures of Dispersion :-

1- Range : largest data value − Smallest data value.

2- Interquartile range : IQR

$IQR = Q_3 - Q_1$



3- Standard deviation . الانحراف المعياري

$S \equiv$ Sample Standard deviation

$\sigma \equiv$ population Standard deviation

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

$\bar{x} \equiv$ Sample mean

$n \equiv$ Sample Size

$x - \bar{x} \equiv$ deviation.

$$\sigma = \sqrt{\frac{\Sigma (x - \mu)^2}{N}}$$

$\mu \equiv$ population mean

$N \equiv$ Population Size

Ex :- Consider the Following Sample data

10 , 12 , 8 , 7 , 13 , 16

Find the Standard deviation.

$$\bar{x} = \frac{\Sigma x}{n} = \frac{66}{6} = 11$$

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

x	$x-\bar{x}$	$(x-\bar{x})^2$
10	-1	1
12	1	1
8	-3	9
7	-4	16
13	2	4
16	5	25
		56

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

Ex :- Sample :

10, 7, 6, 4

x	x^2
10	100
7	49
6	36
4	16
27	201

$$S = \sqrt{\frac{201 - \frac{(27)^2}{4}}{3}}$$

* Variance $= S^2$ الانحراف المعياري
* coefficien of variation

$$CV = \frac{\sigma}{\mu} \times 100\%$$

$$= \frac{S}{\bar{X}} \times 100\%$$

	$\bar{X}$	S
Ex :- Data set 1 →	75	15
Data Set 2 →	65	12

$$CV_1 = \frac{15}{75} = 0.2$$

$$CV_2 = \frac{12}{65} = 0.18$$

* Statistic (sample) * parameter (population)

$\bar{X}$ —— Estimates → μ

S —— مؤشر → σ

* The mean and the standard deviation Applications :

Z_Score العدد ية المعيارية

$Z = \dfrac{x - \mu}{\sigma}$, $Z = \dfrac{x - \bar{x}}{s}$

population — Sample

Ex:- consider a data with $\mu = 70$, $\sigma = 5$

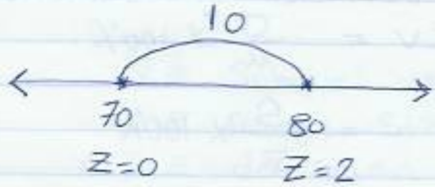
Find the Z-score → For $x = 65, 80, 70$

$\rightarrow Z_{65} = \dfrac{65 - 70}{5} = -1$

$\rightarrow Z_{80} = \dfrac{80 - 70}{5} = 2$

$\rightarrow Z_{70} = \dfrac{70 - 70}{5} = 0$

$\pm$ تدل إذا كانت القيمة اكبر أو اصغر من الوسط الحسابي



$80 = 70 + (2)(5)$
$\quad = \mu + 2\sigma$

$\rightarrow Z_{85} = 3$

$85 = \mu + 3\sigma$ $\qquad = \mu \mp 2\sigma$

$70 \mp 2\sigma$

70 ± 10 $\quad (60, 80)$ $\qquad \mu \pm k\sigma$

$\mu = (k\sigma , \mu + k\sigma)$ $\qquad$ percentage

2- chebysheve Theorem →

For any data set at least $\left(1 - \dfrac{1}{k^2}\right)$ of the data are within k standard deviation of the mean.

$k > 1$, $k = 2$ $(\mu - 2\sigma, \mu + 2\sigma)$

at least $1 - \frac{1}{(2)^2} = \frac{3}{4} = 75\%$

of the data are within 2 Standard Deviation of the mean

$k = 3$ $1 - \frac{1}{(3)^2} = 89\%$ of the data are within 3 . S.d.

Ex:- a stat 236 test has a mean $\mu = 70$ and $\sigma = 5$ Find at least what percentage of scores are in the Following rangs

a- 60 to 80

b- 62.5 to 77.5

c- 55 to 85

a) $Z_{60} = \frac{60-70}{5} = -2$

$Z_{80} = \frac{80-70}{5} = 2$

$(60, 80)$ $70 \pm 2(5)$ $\mu \mp k\sigma$

$k = 2 \rightarrow$ at least $1 - \frac{1}{4} = 75\%$ of the scores are in the range $(60, 80)$

b) $Z_{62.5} = 1.5$. $Z_{77.5} = 1.5$

$k = 1.5$ at least $= 1 - \frac{1}{(1.5)^2} = 56\%$

c) $Z_{55} = -3$ $Z_{85} = 3$

$k = 3 \rightarrow 89\%$

3- Empirical Rule

bell-shaped data المعلومات متماثلة بالنسبة للوسط

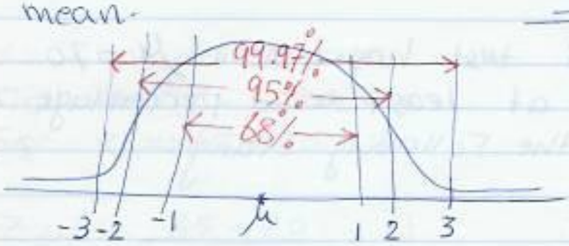
الحسابي اي على جنبية المعلومات تكون متساوية

For abell-shaped data

1- Approximately 68% of the data are within 1 S.d of the mean

2- Approximately 95% of the data are within 2 S.d of the mean

3- Approximatly all the data are within 3 S.d of the mean.



Ex:- let $\mu = 70$. $\sigma = 5$

Find the percentage of scores

1- 65 to 75
2- 60 to 80
3- 65 to 80
4- more than 65
5- more than 85

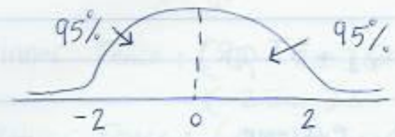
1→ $Z_{65} = -1$ $\qquad Z_{75} = 1$

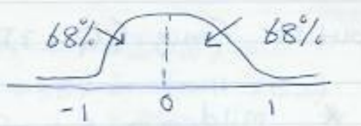
$k = 1 \xrightarrow{E.R} 68\%$

2→ $Z_{60} = -2$ $\qquad Z_{80} = 2$

$k = 2 \longrightarrow 95\%$

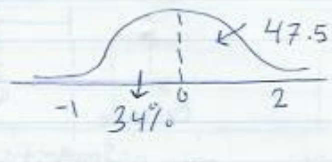
3→ $Z_{65} = -1$ $\qquad Z_{80} = 2$

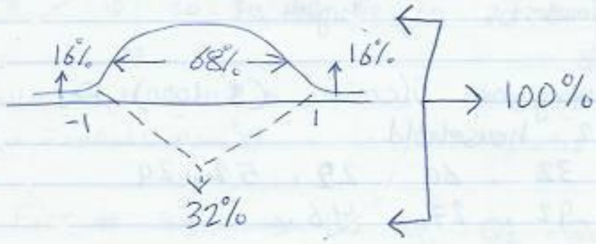




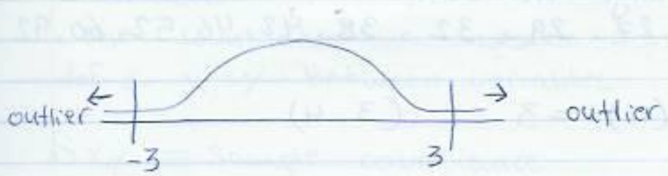
4 → $Z_{65} = -1$

$100 - 68 = 32$





* Outliers → Extreme value → very small. / very large.



$Z > 3$
$Z < -3$] outlier.

$Z > 2$
$Z < -2$] outlier

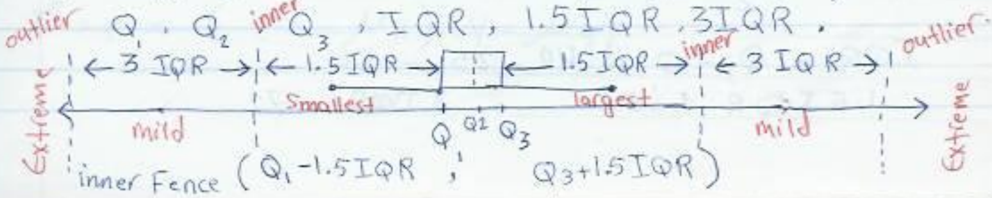
Ex:- $\mu = 55$, $\sigma = 5$

is 98 an outlier = ??

$Z_{98} = \frac{98-55}{5} = \frac{43}{5} = 8.6 > 3$ → extreme outlier.

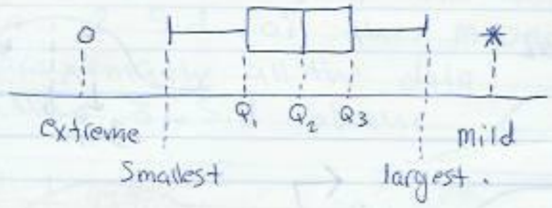
* Box - and - Whisker Plot → (Box plot)

Q_1 , Q_2 , Q_3 , IQR , 1.5IQR , 3IQR .



outlier Fence $(Q - 3IQR \, , \, Q_3 + 3IQR)$

$*$ mild   O extreme



Ex:- The Following are incomes ($ 1000) For a sample of 12 household.

23 , 17 , 32 , 60 , 22 , 52 , 29
38 , 42 , 92 , 27 , 46 ,

construct a Box plot For the data.

Solution :- Arange the data

17 , 22 , 23 , 27 , 29 , 32 , 38 , 42 , 46 , 52 , 60 , 92

$Q_1 = P_{25} = \left(\frac{25}{100}\right)(12) = 3 \quad (3, 4)$

$Q_1 = \frac{23 + 27}{2} = 25$

$Q_2 = P_{50} = \left(\frac{50}{100}\right)(12) = 6 \quad (6, 7)$

$Q2 = \frac{32+38}{2} = 35$

$Q_3 = P_{75} = \left(\frac{75}{100}\right)(12) = 9$

$Q_3 = \frac{46+52}{2} = 49$

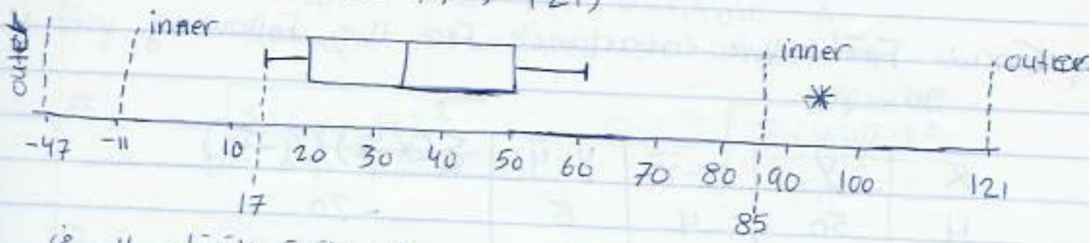
$IQR = Q_3 - Q = 49 - 25 = 24$

$1.5 IQR = 36 \quad , \quad 3 IQR = 72$

الاحد 11-9
الساعة 11
أ 120 قاعة
240 علوم

inner Fence : $(Q_1 - 1.5IQR\ ,\ Q_3 + 1.5IQR) =$
$(25 - 36\ ,\ 49 + 36) : (-11\ ,\ 85)$

outlier Fence : $(Q_1 - 3IQR\ ,\ Q_3 + 3IQR) =$
$(-47\ ,\ 121)$



هذة القيمة التي تلي -11 في الأعداد الموجودة مباشرة

هذة القيمة التي تلي 85 في الاعداد الموجودة مباشرة

* Five # sum → $(S\ ,\ Q_1\ ,\ Q_2\ ,\ Q_3\ ,\ L)$

* Measures of Association between two variable.

1- The Covariance التباين بين متغيرين

def :- X , Y between variables

$S_{x,y}$ = Sample covariance

σ_{xy} = population covariance.

$$S_{xy} = \frac{\sum (x - \bar{x})(y - \bar{y})}{n-1}$$

$\bar{x}$ = Sample mean for the variable X

$\bar{y}$ = Sample mean for the variable Y

n = Sample Size

$$\sigma_{xy} = \frac{\sum (x - \mu_x)(y - \mu_y)}{N}$$

pearson correlation cofficiant [بيرسون] معامل الارتباط

$$r_{xy} = \frac{S_{xy}}{S_x S_y}$$

S_{xy} ≡ Sample covariance.
S_x ≡ Sample S.d For variable X
S_y ≡ Sample standared deviation For variabl y

$$S_x = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n-1}} \quad , \quad S_y = \sqrt{\frac{\Sigma(y-\bar{y})^2}{n-1}}$$

$S_{xy} = 57.5$, $S_y = 11.18$ $S_x = 5.43$

$$r_{xy} = \frac{-57.5}{(5.43)(11.18)} = -0.95$$

$$-1 \leqslant r \leqslant 1$$

النقاط واقعة على خط مستقيم
وصلة موجبة (١)

→ $r = 1$ Complete positive relationship. (طردية)
→ $r = 0$ No relation shipe
→ $r = -1$ Complete negative relationship. (عكسية)

$$r_{xy} = \frac{n(\Sigma xy) - (\Sigma x)(\Sigma y)}{\sqrt{n(\Sigma x^2) - (\Sigma x)^2}\ \sqrt{n(\Sigma y^2) - (\Sigma y)^2}}$$

Ex:- The Income and Food expediture of 7 house Bolds given in the Following table

$n = 7$ $\Sigma xy = 2150$
$\Sigma x = 212$ $\Sigma x^2 = 7220$
$\Sigma y = 64$ $\Sigma y^2 = 646$

الدخل (X) Income	تابع (Y) Food Expenditure	XY	X^2	Y^2
35	9			
49	15			
21	7			
39	11			
15	5			
28	8			
25	9			

$$r_{xy} = \frac{(7)(2150) - (212)(64)}{\sqrt{(7)(7222) - (212)^2}\ \sqrt{(7)(646) - (64)^2}} = 0.96$$

Positive

* Introduction To Probability Ch_4_

Experiment , Sample space and Counting Rules

* Experiment → outcomes (observation or results) ثابتة

* Random → تتغير → تجارب عشوائية

Sample space = the set of all possible outcomes of an experiment (S)

Ex:- 1- toss a coin → Head or tail

$S = \{H, T\}$

2- Roll a die → $S = \{1, 2, 3, 4, 5, 6\}$

3- Select a prouduct For insection للتفتيش

$S = \{$ diefective , NO defective $\}$

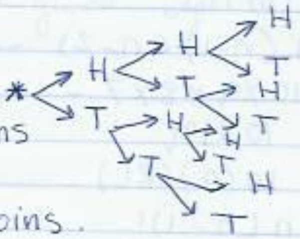
→ Counting rules :- قوانين العدّ

1- multiplication Farmula.

IF an experment has K steps and step 1

has step 2 has n_2 outcome.
step k has n_k out come.
Then the total number of outcomes is
$(n_1)(n_2) --- (n_k)$.

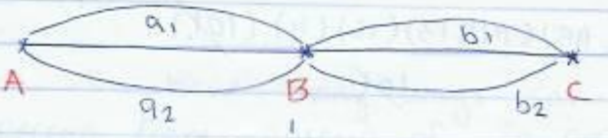
Ex:- toss two coins →
$O_2 \cdot O_2 \to 4$



$S = \{HH, HT, TH, TT\}$ 2 coins

$S = \{HHH, ----- TTT\}$ 3 coins.

(؟ يمر من A الى C ويمر من خلال B).



2 . B 3 = 6

Ex: let $A = \{1, 2, 3, 5, 7, 8, 9\}$
construct a 3 digits numbar.
a) with replacement → تكرار

$\square \quad \square \quad \square \quad = 7^3$
7 7 7

b) without replacement بدون تكرار

$\square \quad \square \quad \square$
5 6 5

العدد الأول الفردي $\{1, 3, 5, 7, 9\}$

c) an odd without replacement
5 , 6 , $\boxed{x}$
5

d) an odd with replacment

$\boxed{}$ $\boxed{}$ $\boxed{}$

7 7 5

* n Factorial n مضروب

$n! = n(n-1)(n-2) \cdots (3)(2)(1).$

$10! = 10 \times 9 \times 8 \times 7 \cdots (3)(2)(1).$

$= 10 \times 9!$

$= 10(9)(8!)$

$n! = n(n-1)!$

$= n(n-1)(n-2)!$

$0! = 1$

Ex:- $\frac{15!}{10!} = \frac{(15)(14)(13)(12)(11)\cancel{(10!)}}{\cancel{10!}}$

* Parmutation :- التبادل ← الترتيب مهم

The permutation of r objects Selected From n objects is denoted by npr.

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

$$15P_5 = \frac{15!}{5!} = \frac{15!}{(15-5)!}$$

Ex:- In how many ways can 7 persons be seated on arew of 10 chairs

$(10)(9)(8)(7)(6)(5)(4)$

$$10P_7 = \frac{10!}{3!} = \frac{(10)(9)(8)(7)(6)(5)(4)\cancel{(3!)}}{\cancel{3!}}$$

* Combination. التوافق ← الترتيب غير مهم

The number of ways that r objectes be selected From n objects without regard to order is called the number of Combination and it is given by

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

Ex:- $\binom{12}{4} = \frac{12!}{(12-4)!\,4!} = \frac{12!}{8!\,4!} = \frac{(12)(11)(10)(9)(\cancel{8!})}{\cancel{8!}\,(4)(3)(2)(1)}$

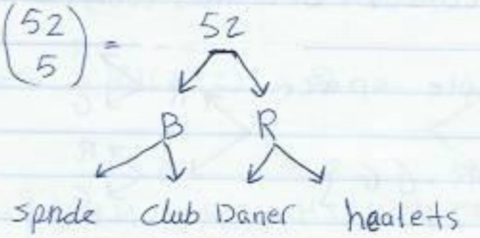
$= 495$

Ex:- In how many ways can we select by person From agroup of 20, person الترتيب مش مهم لأنه معطى اي اشخاص

$$\binom{20}{5} = \frac{20!}{(20-5)!\,5!}$$

Ex:- In How many different 5- card hands can we select From a 52 card.

$\binom{52}{5} =$



\# 5

(10, L, K, Q)

* probability :-

Sample space S

Experiment → outcomes { }

$S = \{ S_1, S_2, S_3 \text{ ---- } S_n \}$

S_i ≡ Sample point.

→ Event : any subset of the Sample space.

simple event → $A = \{S_i\}$

Compound event → more than one outcome.

Probability → $P : S \to [0.1]$

Probability → A value between 0 and 1 inclsive describing the relative possibility (chance, or liklehood) an event will occur.

A sample, compound

1. $0 \leqslant p(s_i) \leqslant 1$ $\qquad\qquad 0 \leqslant p(A) \leqslant 1$
2. $\sum p(s_i) = 1$

Ex:- A box contains a Few red and a Few green marbles. IF two marbles are randomly drawn and the colors of these marbles are observed.

1. write the sample space

$S = \{RR, RG, GR, GG\}$

2. let B ≡ both marbles are of different colors

$B = \{RG, GR\}$.

3. let $C \equiv$ at least one marble is red.

$C = \{RG, GR, RR\}$

4. let $D \equiv$ Not $\boxed{\text{more than one}}$ marble is green. at most one → 0 or 1

$S = \{RR, RG, GR\}$

→ classical Probability

let $S = \{s_1, s_2, ---, s_n\}$

Assume that $s_1, s_2, ---s_n$ are equally likely.

$p(s_1) = p(s_2), --- = p(s_n) =$

$\rightarrow p(s_i) = \dfrac{1}{\text{total number}} = \dfrac{1}{n}$

IF A is a compound event

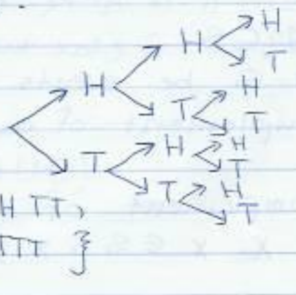
$p(A) = \dfrac{\# A}{\# S} =$

Ex:- Roll 3 coins, let $B \equiv$ at least one head

let $C =$ exactly one head, $D =$ at most one head

Find $p(B)$, $p(C)$, $p(D)$.

$\# S = (2)(2)(2) = 8$



$S = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$

$B \equiv$ at least one 1 or 2 or 3

$B = \{HHH, HHT, HTH, HTT, THH, THT, TTH\} = \dfrac{7}{8}$

$C \equiv$ exactly 1

$S = \{ TTH, HTT, THT \} = \frac{3}{8}$

$D \equiv$ at most $1 \equiv 1$ or 0

$= \{ HTT, THT, TTH, TTT \} = \frac{4}{8} = \frac{1}{2}$

* Relative Frequency Approach :-

$p(A) = \frac{R}{F} \quad 0 \leq \quad \leq 1$

Ex :-

Sample 120 students 24 bussiness major

Select one student at random

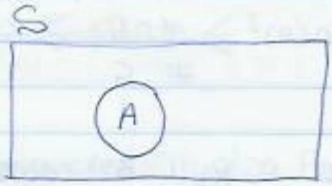
$p(\text{bussiness student}) = \frac{24}{120} = 0.2$

Subjective probability

$0 \leq p(A) \leq 1 \qquad \sum p(\,) = 1$

$S \rightarrow$ univarsal set

subsets $\rightarrow$ event



$P(S) = \frac{n(S)}{n(S)} = 1$

$p(\emptyset) = \emptyset = \frac{n(\emptyset)}{n(S)} = \frac{0}{n(S)} = 0$

event $A \subseteq S$

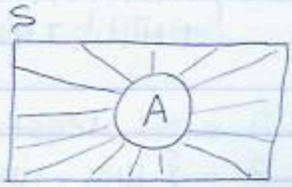
let : A. B be events of S

The complement of an event

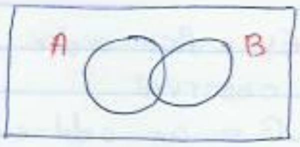
$A' \equiv$ complement of A

$A' = \{ X : x \in S, x \notin A \}$

$P(A) + P(A') = 1 \qquad A' =$



$P(A \cap B)$



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

$P(A \cup B)$

Ex:- Roll a die
let $A \equiv$ an even number is observed
$B \equiv$ a number less than 3 is observed
Find, $P(A)$, $P(B)$, $P(A \cap B)$, $P(A \cup B)$.

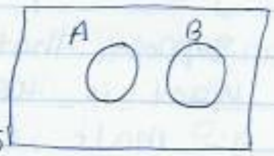
$\# S = 6 \quad S = \{1, 2, 3, 4, 5, 6\}$
$A = \{2, 4, 6\} \rightarrow P(A) = \frac{3}{6} = \frac{1}{2}$
$B = \{1, 2\} \rightarrow P(B) = \frac{2}{6} = \frac{1}{3}$

$A \cup B' \rightarrow \{2\} \rightarrow P(A \cap B) = \frac{1}{6}$
$P(A \cup B) = P(A) + P(B) - P(A \cap B)$
$= \frac{3}{6} + \frac{2}{6} - \frac{1}{6} = \frac{4}{6}$

Ex:- Roll 2 dice
$n(S) = (6)(6) = 36$
$S = \{(1,1), (1,2) \text{ ---- } (6,1) \text{ --- } (6,6)\}$
let $\rightarrow A \equiv$ Sum is equal to 8
$B \equiv$ an odd number is observed on both Face.
Find, $P(A \cup B)$
$A = \{(3,5), (5,3), (4,4), (2,6), (6,2)\}$
$B = \{(1,1), (1,3), (1,5), (3,1), (3,3), (3,5), (5,1), (5,3), (5,5)\}$
$P(B) = \frac{9}{36} \qquad P(A) = \frac{5}{36}$
$A \cap B = \{(3,5), (5,3)\} \rightarrow P(A \cap B) = \frac{2}{36}$

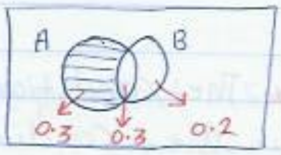
$P(A \cup B) = \frac{5}{36} + \frac{9}{36} - \frac{2}{36} = \frac{12}{36} = \frac{1}{3}$

→ Def: IF $A \cap B = \emptyset$ ($P(A \cap B) = 0$)
the A and B are mutually exclusive (disjoint) events.
IF A, B are mutually exclusive then
$P(A \cup B) = P(A) + P(B)$.



Ex:- IF $P(A \cup B)' = 0.2$
$P(A) = 0.6$, $P(B) = 0.5$

a- are A, B mutually exclusive?
$P(A \cup B)' = 0.2 \rightarrow P(A \cup B) = 0.8$
$P(A \cup B) = P(A) + P(B) - P(A \cap B)$
$0.8 = 0.6 + 0.5 - P(A \cap B) =$
$P(A \cap B) = 0.3 \neq 0$
→ A, B are not mutally Exclusive.



$A - B = \{X : X \in A, X \notin B\}$

$P(A - B) \rightarrow$ تعني إحتمال حدوث A وعدم حدوث B
$P(A) - P(A \cap B)$.

Ex:- $P(A - B) = 0.6 - 0.3 = 0.3$

Ex:- A survey of 80 studant at BZU reveated the Following regarding the gender & smaking.

	Male	Female	total
Smoke	18	7	25
no smoke	22	33	55
total	40	40	80

joint probability table →

	M	F	total
S	$\frac{18}{80} = 0.225$	$\frac{7}{80} = 0.0875$	0.3125
N	0.275	0.4125	0.6875
	0.625	0.375	1

marginal probability

Suppose that a student was selected at random what is the probabilites that :-

1- Male , $p(m) = 0.625$

2- Smoke , $p(S) = 0.3125$

3- Male and no smoke.

$p(M \cap N) = 0.275$

4- Female or no smoke

$p(F \cup n) = p(F) + p(N) - p(F \cap N)$

$= 0.375 + 0.6875 - 0.4125$

* The conditional probability →

The conditional probability of A given B

احتمال حدوث A علماً أن B واقعة.

is written as $p(A \backslash B)$ and it is defined by $p(A \backslash B) = \dfrac{P(A \cap B)}{P(B)}$

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

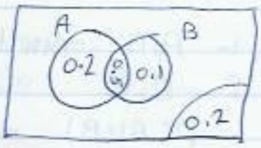
Ex :- IF $p(A) = 0.7$, $P(B) = 0.6$, $p(A \cup B) = 0.8$

Find → 1- $p(A \backslash B)$

2- $P(B \backslash A)$

3- $p(A` \backslash B)$

4- $p(A` \backslash B`)$



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

$0.8 = 0.7 + 0.6 - P(A \cap B)$

$P(A \cap B) = 0.5$

$P(A \backslash B) = \dfrac{P(A \cap B)}{P(B)} = \dfrac{0.5}{0.6}$

$P(B \backslash A) = \dfrac{P(A \cap B)}{P(A)} = \dfrac{0.5}{0.7}$

قاعدة ← $A - B = A \cap B'$

$P(A' \backslash B) = \dfrac{P(A' \cap B)}{P(B)} = \dfrac{P(B - A)}{P(B)} = \dfrac{0.1}{0.6}$

$P(A' \backslash B') = \dfrac{P(A' \cap B')}{P(B')} = \dfrac{P(A \cup B)'}{P(B')} =$

$\dfrac{1 - P(A \cup B)}{1 - P(B)} = \dfrac{0.2}{0.4} = \dfrac{1}{2}$

* Def:- two events A, B are independent if

$P(A \backslash B) = P(A)$ احتمال حدوث A بدون اي علاقة

مع حدوث B دائماً تساوي $P(A)$ لان B لا يؤثر على A لانهما

$P(B \backslash A) = P(B) = \dfrac{P(A \cap B)}{P(A)}$ مستقلان

Suppose that A, B are independant.

$P(A \backslash B) = P(A)$

$\downarrow$

$\dfrac{P(A \cap B)}{P(B)} = P(A)$

$\boxed{P(A \cap B) = P(A) \cdot P(B)}$

A, B independent $\rightarrow P(A \cap B) = P(A) \cdot P(B)$

EX:- $P(A) = 0.6 \qquad P(B) = 0.4$

Find $P(A \cup B)$ if $\rightarrow$

1- A, B mutually exclusive
2- A, B indep

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

1- m. ex $\rightarrow P(A \cap B) = 0$

$\rightarrow P(A \cup B) = P(A) + P(B)$

2- Ind: $P(A \cap B) = P(A) \cdot P(B)$

	A Aljazeera	D Abu Dhabi	C MBC	total
Male M	100	150	50	300
Female F	40	40	20	100
	140	190	70	400

a- $P(A) = \frac{140}{400}$

b- $P(A \backslash m) = \frac{P(A \cap m)}{P(m)} = \frac{100}{300}$

$P(D \backslash m)$

$P(C \backslash m)$

c- $P(A \cap F) = \frac{40}{400}$

d- $P(F \backslash C) = \frac{20}{100}$

$P(A \backslash m) = P(A) \cdot P(m) = \frac{140}{400} \cdot \frac{300}{400} =$

$P(A \backslash F) = P(A) \cdot P(F)$

$P(A \backslash m) \overset{?}{=} P(A) = \frac{100}{300}$

$\frac{100}{300} \neq \frac{140}{300}$ Independant عدم التساوي

تعني الاستقلالية.

Ex:- 30 person → woman 18 → employees
man 12 → 〃

p(w \ m.B.A)

30
w 18 ↓ m.B.A → 10
m 12 ↓ m.B.A → 8

* Bayes theorem :-

let S be a sample space

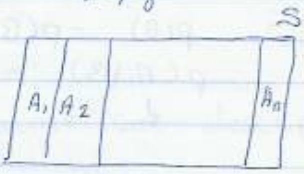
$A_1, A_2, \ldots A_n$ are events

such that : $A_1 \cup A_2 \cup \ldots \cup A_n = S$

$A_i \cap A_j = \emptyset \quad i \neq j$

let B be an event in S

$B = (A_1 \cap B) \cup (A_2 \cap B) \ldots (A_n \cap B)$



$P(B) = P(A_1 \cap B) \cup (A_2 \cap B) \cup \ldots (A_n \cap B)$

$= P(A_1 \cap B) + P(A_2 \cap B) + \ldots P(A_n \cap B)$

$P(B \setminus A_1) = \frac{P(A_1 \cap B)}{P(A_1)}$

$P(A_1 \cap B) = P(B \setminus A_1) P(A_1)$

$= P(B \setminus A_1) P(A_1) + P(B \setminus A_2) P(A_2) \ldots$

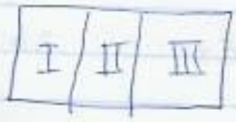
$P(B \setminus A_n) P(A_n)$

$P(B)$ = total probability.

Ex:-

	machin I	machin II	machin III
Defective	5%	2%	3%
Good			
	50%	30%	20%

Select an item at random Find the probability that the item is defective.



$D = D \,\&\, I$ or $D \,\&\, II$ or $D \,\&\, III$

$= (D \cap I) \cup (D \cap II) \cup (D \cap III)$.

$p(D) = p(D \cap I) + p(D \cap II) + p(D \cap III)$

$= p(D \backslash I)\, p(I) + p(D \backslash II)\, p(II) + p(D \backslash III)\, p(III)$

$= (5\%)(50\%) + (2\%)(30\%) + (3\%)(20\%)$.

$= 0.025 + 0.006 + 0.006$

$= 0.037$.

$p(B) = p(B \backslash A_1)\, p(A_1) + \text{----} + p(B \backslash A_n)\, p(A_n)$.

$p(A_i \backslash B) = \dfrac{p(B \backslash A_i)\, p(A_i)}{p(B)}$

$\downarrow$

$\dfrac{p(A_i \cap B)}{p(B)} = \dfrac{p(B \backslash A_i)\, p(A_i)}{p(B)}$

If an item selected at random is found to be defective what is the probability that it is from I

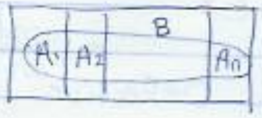
$p(I \backslash D) = \dfrac{p(D \backslash I)\, p(I)}{p(D)} = \dfrac{0.025}{0.037} = \dfrac{25}{37} =$

0.67

$p(B) = \sum p(B \backslash A_i)\, p(A_i)$

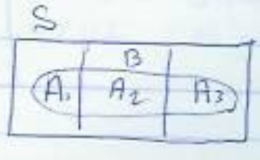
$p(A_i \backslash B) = \dfrac{p(B \backslash A)\, p(A)}{p(B)}$

$\rightarrow \dfrac{p(A \cap B)}{p(B)} = \dfrac{p(B \backslash A)\, p(A)}{p(B)}$



Ex:- $p(A_1) = 0.2$ $\quad p(A_3) = 0.3$

$p(A_2) = 0.5$



$p(B\backslash A_1) = 0.5$
$p(B\backslash A_2) = 0.4$
$p(B\backslash A_3) = 0.3$
$p(B) = p(B\backslash A_1)\,p(A_1) + p(B\backslash A_2)\,p(A_2) + p(B\backslash A_3)\,p(A_3) =$
$= (0.5)(0.2) + (0.4)(0.5) + (0.3)(0.2)$
$= 0.36$

$$p(A_2\backslash B) = \frac{p(B\backslash A_2)\,p(A_2)}{p(B)} = \frac{(0.4)(0.5)}{0.36}$$

* Discrete probability Distributions Ch. 5

Experiment → sample space-
$S = \{S_1, S_2, S_3, --- S_n\}$
$S \longrightarrow$ numberical values

* Random variables :-

A rondom variable is anumerical description of the outcome of an experiment

Ex:- toss acoin 3 times
$S = \{HHH, ----, TTT\}$ 8-out come.
define arandom variable R to be the number of heads observed
TTT → 0
HTH → 2
R. assumes the values : 0, 1, 2, 3
$R = \{0, 1, 2, 3\}$

Ex:- Roll adie twice
$S = \{(1,1) ----, (6,6)\}$
let X = sum of the two Faces
$X = \{2, 3, \ldots, 12\}$
$(1,5) \longrightarrow 6$
$(5,6) \longrightarrow 11$

Random variable

Discrete	Continuous
Assume a Finite or inFinite number oF values $\{x_1, x_2, \dots, x_n\}$	assume values that contained in an interval or a collection of intervals
معظم تطبيقاته عن ارقام معينة.	$x_1 \leq t \leq x_2$ فترات
	$\{x_1, x_2\}$ ومعظم تطبيقاته على الزمن

* Discrete Random variables :-
(D Rv)

$$X = \{x_1, x_2, \dots, x_n\}$$

Ex :- Roll die twice
X = Sume of the tow Faces.
$X = \{2, 3, 4, \dots, 12\}$

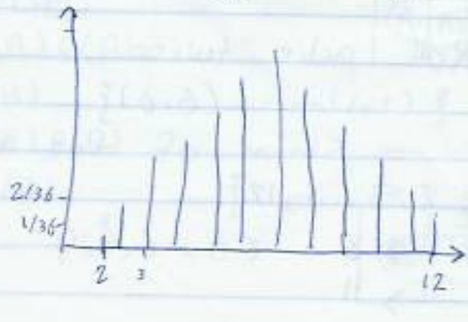
$P(x) \rightarrow$ تعني احتمال ان يكون لدي جمع الرقمين يساوي الرقم في X

X	P(x)
2	1/36
3	2/36
4	3/36
5	4/36
6	5/36
7	6/36
8	5/36
9	4/36
10	3/36
11	2/36
12	1/36

اي عبارة عن طاولة توضح العدد مع احتمالية حدوث هذه القيمة

$\leftarrow$ Probabilitiy Distribution

$$\{(2, \tfrac{1}{36}), (3, \tfrac{2}{36}) \dots\}$$



* $f(x)$ probability Function
  1- $0 \leq F(x) \leq 1$
  2- $\sum F(x) = 1$
* probability Distribution
* Expected value (mean).
* Variance. التباين

The expected value and variance
let $X = \{x_1, x_2, \dots, x_n\}$ be a DRV
with aprobability Function $F(x)$ then
The expected value of X denoted by
$E(X)$ or μ is defined as Follows
$E(X) = \sum x F(x) = \mu$

Ex: toss a coin 3 times, let X = # of heads
$X = \{0, 1, 2, 3\}$

X	0	1	2	3
$F(x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$
$xF(x)$	0	$\frac{3}{8}$	$\frac{6}{8}$	$\frac{3}{8}$
$(x-\mu)$	-1.5			
$(x-\mu)^2$	2.25	0.25	0.25	2.25

Find the expected value For X
$E(x) = \sum x F(x) = \frac{12}{8} = 1.5$

$Var(X) = \sum (x-\mu)^2 p(x)$
$Var(X) = (2.25)(\frac{1}{8}) + (0.25)(\frac{3}{8}) + (0.25)(\frac{3}{8}) + (2.25)(\frac{1}{8})$
$= \frac{1}{8}[2.25 + 0.75 + 0.75 + 2.25]$

$= \frac{6}{8} = 0.75$

$\sigma(x) = \sqrt{var(X)} = \sqrt{0.75}$

$$\begin{aligned}
var(X) &= \sum (x-\mu)^2 p(x) \\
&= \sum (x^2 - 2\mu x + \mu^2) p(x) \\
&= \sum x^2 p(x) - \sum 2\mu x\, p(x) + \sum \mu^2 p(x) \\
&= E(x^2) - 2\mu \sum x p(x) + \mu^2 \sum p(x) \\
&= E(X^2) - 2\mu.\mu + \mu^2 \\
&= E(x^2) - \mu^2 \\
&= E(x^2) - (E(x))^2
\end{aligned}$$

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

Ex:- consider the Following probability distribution.

X	P(X)	XP(X)
0	0.05	0
1	0.15	0.15
→ 2	0.1	0.2
3	0.15	0.42
4	0.25	1
5	0.3	1.5
		3.3

Find → $\boxed{1-}$ $p(x=2) = 0.1$

$\boxed{2-}$ $p(x>1) =$
$p(x=2)+p(x=3)+p(x=4)+p(x=5)$
$= 1 - p(x=0)$
$= 0.95$

$\boxed{3-}$ $p(2 \leq x \leq 5) =$
$p(x=2)+p(x=3)+p(x=4)+p(x=5)$

$\boxed{4-}$ Find $E(X)$, $Var(X)$.

X	X^2	$X^2(p(x))$
0	0	0
1	1	0.15
2	4	0.4
3	9	1.35
4	16	4
5	25	37.5
		43.4

* Var + E = دائماً موجب
$var(X) = (43.4) - (3.3)^2$
$= (43.4) - 10.89 =$
32.51

* The Binomial Distribution

Binomial Experiment :-

1. The experiment consists of n trials
2. In each trial two outcomes are possible success or Failure.
3. The probability of success p and the propability of Failure $1-p$ is the seme in each trail
4. The trails are independent.

let X be abinomial distribution with # of trails $= n$ and pr (success) $= p$

$$X = B(n, p)$$

Ex:- toss a coin 100 times $X =$ # of heads

$$X = B(100, \tfrac{1}{2})$$

let $X = B(n, p) \rightarrow$

- probability Function
- Expected value
- Variance

The probility Function is given by $\rightarrow$

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

where :- $X \equiv$ # of Success in n-trials

$n \equiv$ # of trials

$p \equiv P_r$ (success)

$1-p \equiv P_r$ (Failure)

$$\binom{n}{x} = \frac{n!}{(n-x)!\,x!}$$

Ex :- Chap. 1 , 4 أسئلة على

1 → Difective

2 → non Difective

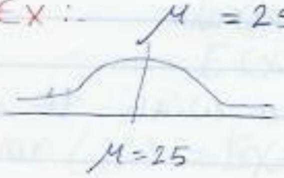
at least one

A = at least one = one or two

$\bar{A}$ = Zero defective.

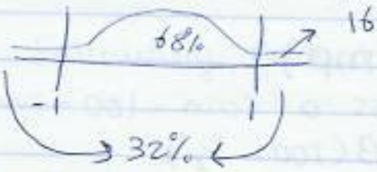
$\left(\frac{2}{3}\right)$

Ex : $\mu = 25$, $\sigma = 3$



P84

$Z = 1 \quad = P84$



$P97.5 = 31$

$P_{25} = Z = \frac{X-25}{3} =$

$1 = \frac{X-25}{3} = X = 28$

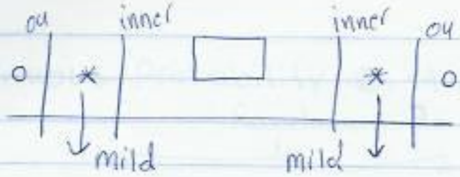
Ex :-

$\boxed{\begin{matrix} 3 & 7 \\ W & R \end{matrix}}$

a) 4 red $= \dfrac{\binom{7}{4}}{\binom{10}{4}}$

b) 2 red & 2W $= \dfrac{\binom{7}{2}\binom{3}{2}}{\binom{10}{4}}$

c) 100% = 1R 3W $= \dfrac{\binom{3}{1}\binom{7}{3}}{\binom{10}{4}}$



* Binomial Distributions:-

$X = B(n, p)$

$p = pr(\text{success})$

$1-p$

p. Function :-

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

$\binom{n}{x} = \frac{n!}{(n-x)!\,x!}$

$n \equiv$ # of trials

$X \equiv$ # of success

$p \equiv pr(\text{Function})$

$1-p \equiv pr(\text{Failure})$

let :- $X = B(n, p)$ then

* expected value $= np$

* variance $= np(1-p)$

Ex:- Exame → 20 Question → a, b, c, d

$p = 1/4 \qquad (20)(1/4) = 5$

Ex:- Toss a coin 10 times, let X = # of heads.

$X = B(10, \frac{1}{2})$

Find 1) $P(X=7) = \binom{10}{7} (\frac{1}{2})^{7} (\frac{1}{2})^{3}$

2) pr (at least 2 heads)

$P(X \geq 2) = P(X=2) + \dots + P(X=10)$

$= 1 - [P(X=0) + P(X=1)]$

$= 1 - [\binom{10}{0}(\frac{1}{2})^{0}(\frac{1}{2})^{10} + \binom{10}{1}(\frac{1}{2})^{1}(\frac{1}{2})^{9}]$

3) pr (at most 8 heads)

$P(X \leq 8) = P(X=8) + \dots + P(X=0)$

$= 1 - [P(X=9) + P(X=10)]$

4) Find the expected number of heads

$E(X) = np$

$= (10)(1/2) = 5$

5) Find the variance.

$\sigma^2 = np(1-p)$

$= (10)(1/2)(1/2) = 2.5$

$B(10, 0.7)$

$p(X=7) = \binom{10}{7}(0.7)^7(0.3)^3$

$B(10, 0.3)$

$p(X=3) = \binom{10}{3}(0.3)^3(0.7)^7$

Ex:- $\left(\frac{31}{181}\right) \rightarrow$ 5% of drivers are women

Select 10 drivers at random.

a) Binomial Exp ??

- $n = 10$
- $p(W) = 5\%$, $p(M) = 95\%$
- Indepent

$B = (10, 0.05)$

2) 2 of the drivers will be women

$P(X=2) = \binom{10}{2}(0.05)^2(0.95)^8$

3) non will be women.

$p(X=0) = (0.95)^{10}$ احتمال ان يكون M بعدد المرات المختارة.

4) Pr (at least one)

$p(X \geq 1) = p(X=1) + \dots + p(X=10)$

$= 1 - p(X=0)$

$= 1 - (0.95)^{10}$

$(0.05)(10) = 1/2$

$(n)(p) =$

* Continuous Probability Distributions Ch-6

Random

$\swarrow$ $\searrow$

Discrete	Cont.
$\{x_1, x_2, \ldots, x_n\}$	$\{a, b\}$
Binomial	1- 2-

D8- Probability Function. $\quad 0 \leq F(x) \leq 1$

$Exp \to E(x) = \sum x F(x)$. $\quad \sum F(x) = 1$

$Var(x) = \sum (x-\mu)^2 F(x)$.

$E(x) = \int_a^b x F(x) dx \quad , \quad Var = \int_a^b (x-\mu)^2 F(x)$.

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

conti $\quad P(X=x) = 0 \qquad \int_a^a = 0$

$P(x \geq 7)$, $P(x > 7)$ نفس الجواب لأن المساواة لا

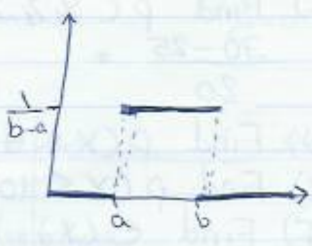
$[a, b]$, (a, b) تعني شيء.

* Uniform Distribution :-

- continuous distribution
- defined on $[a, b]$
- Uniform over $[a, b] = U[a, b]$

The probability density Function. For $U[a, b]$ is given by.

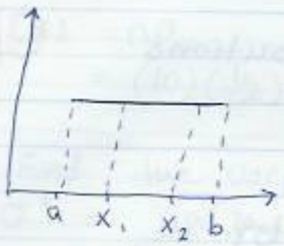
$$F(x) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & \text{otherwise.} \end{cases}$$



$$\int_a^b F(x) dx = 1$$

$$\cancel{(b-a)} \cdot \frac{1}{\cancel{(b-a)}} = 1$$

$Pr(x_1 \leq X \leq x_2)$



$$\frac{(x_2 - x_1)}{b - a}$$

$Pr(x_1 < X < x_2) = Pr(x_1 \leqslant X \leqslant x_2)$

* The expected value

IF $X = U[a,b]$ then

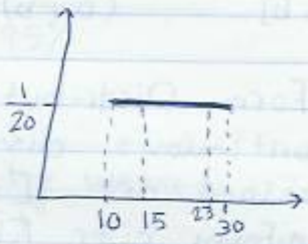
$E(X) = \frac{a+b}{2}$

$var(X) = \frac{(b-a)^2}{12}$

Ex:- let $X = U[10, 30]$

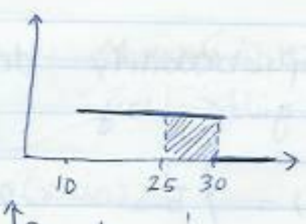
a) write and graph the P.d.F.

$$F(x) = \begin{cases} \frac{1}{30-10} & 10 \leqslant x \leqslant 30 \\ 0 & \text{otherwise} \end{cases}$$



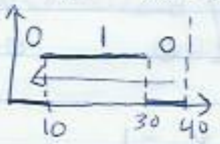
b) Find $p(15 < x < 23)$.

$= \frac{23-15}{30-10} = \frac{8}{20} = 0.4$



c) Find $p(x \geqslant 25)$

$\frac{30-25}{20} =$



d) Find $p(X \leqslant 40) = 1$

e) Find $p(X \leqslant 10) = 0$

F) Find $E(X) =$

$E(X) = \frac{10+30}{2} = 20$

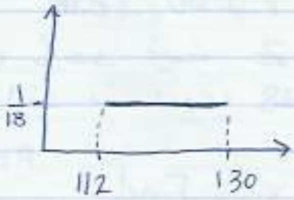
G) Find $Var(X) = \frac{(30-10)^2}{12} = \frac{400}{12}$

$Ex:- \frac{3}{202}) \rightarrow$ H M H M * نحول الساعات الى دقائق.

1 : 52 $\rightarrow$ 2 : 10

[112 , 130]

a) $F(x) = \begin{cases} \frac{1}{18} & 112 \leq x \leq 130 \\ 0 & \text{otherwise} \end{cases}$



b) no more than 5 minutes late.

$p(112 < x < 117) = \frac{5}{18}$

c) more than 10 minutes late.

$p(X > 122) = 8/18$

d) Find $E(x) = \frac{130+112}{(2)} = 121$ the avarge of plan to be late.

μ = mediam = mode.

The width is a S.d

من ١ الى -١ تشتت أقل من ٢ الى -٢

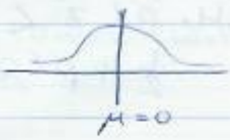
* Normal probability Distribution.

$N(\mu, \sigma)$

Standard Normal Distribution.

Normal + $\mu = 0$

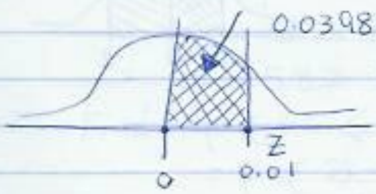
$\sigma = 1$



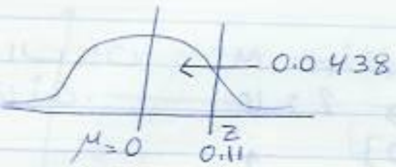
Normal $\rightarrow$ standard Normal

$N(\mu, \sigma) \rightarrow N(0.1)$

Z_Score



Z	0.00	0.01		0.08	0.09
0.0					
قيمة Z 0.1	0.0398	0.0438			
0.2					
1.5				0.4429	
3.0					



نقوم بجمع قيمة Z + القيمة التي تعطي الرقم المطلوب ثم نكتب قيمة التقاطع بينهما

مثال :- إذا اردنا قيمة 1.58 نقوم بأخذ قيمة Z = 1.5 وتقاطعها مع 0.08 لتصبح 1.58 والتي تساوي 0.0438

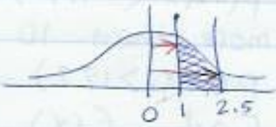
Ex:- Find.

1- $p(0 \leq Z < 1.95) = 0.4744$.

$p(1 \leq Z \leq 2.5)$.

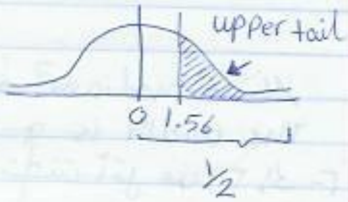
2. $p(0 < Z < 2.5) - p(0 \leq Z < 1)$

$.4938 - .3413$



3- $p(Z > 1.56)$.

$= \frac{1}{2} - p(0 < Z < 1.56)$.

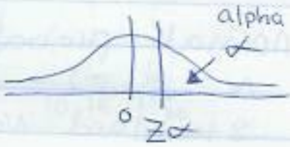


$= \frac{1}{2} - 0.4406$

$= 0.0594$

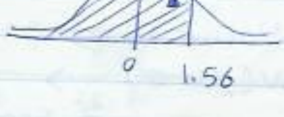
$Z_{1.56} =$

$Z_{0.0594} = 1.56$

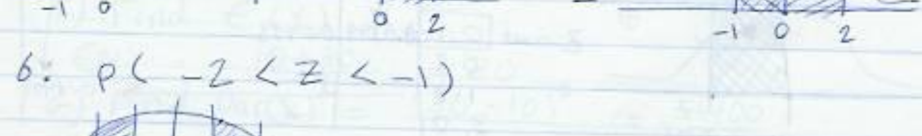


4. $p(Z < 1.56) =$

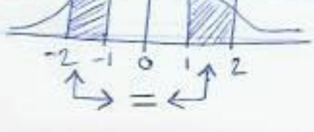
$\frac{1}{2} + p(0 < Z < 1.56)$



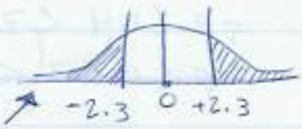
5. $p(-1 < Z < 2)$



6. $p(-2 < Z < -1)$



7. $p(Z > 2.3)$
$= \frac{1}{2} - p(0 < Z < 2.3)$

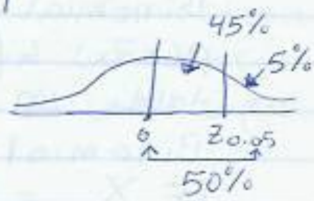


Ex:- Find Z_0 such that

1- $p(0 < Z < Z_0) = 0.3749$
   $Z_0 = 1.15$

	0.05
1.1	0.3749

2. The area in the upper tail is 0.05 $(Z_{0.05})$.
   $50\% - 5\% = 45\%$
   45% بنبحث على الرقم القريب منا



	0.04	0.05
1.6	0.4495	0.4505

1.64
1.65] 1.645

ولما نكتب قيمة 1.64 أو 1.65 اونقوم بإعطاء المتوسط .

Ex:- waiting times in a given bank are normally distributed with $\mu = 15$, $\sigma = 5$. Find the probability that a customer will wait more than 20 minutes

$\mu = 15$. $\sigma = 5$

1- Find $p(x > 20) = p\left(Z > \frac{20-15}{5}\right)$.
$= p(Z > 1)$
$= \frac{1}{2} - p(0 < Z < 1)$
$= \frac{1}{2} - 0.3413$
$= 0.1587$

2- Find $p(22 \le x < 27)$
$= p\left(\frac{22-15}{5} < Z < \frac{27-15}{5}\right)$

$= p(1.4 < Z < 2.4)$

$= p(0 < Z < 2.4) - p(0 < Z < 1.4)$

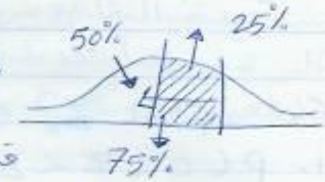
3- Find P_{75} =

نذهب إلى داخل

الجدول لإيجاد

قيمة 0.25 او

مايقارب عليها



$Z = 0.67$

$Z = \dfrac{X - 15}{\sigma}$

$0.67 = \dfrac{X - 15}{5} = \quad X = 15 + (5)(0.67) = 18.35$

* Normal Approximation to binomial

Binomial $\longrightarrow$ normal.

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

table $n = 20$

Binomial $B(n,p) \longrightarrow N(\mu, \sigma)$

IF $X = B(n,p)$ then

X can be approximated

by anormal distribution

if $np \geq 5$, $n(1-p) \geq 5$

$B(n,p) \longrightarrow N(\mu, \sigma)$

$\mu = np$

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

$N = (np\sqrt{np(1-p)})$.

$B(100, \frac{1}{2}) \longrightarrow$ Find $p(X \geq 20)$

$(X=20) + \text{----} \; X=100$.

$p(X \geq 20) = p(X > 20)$

في ال N المساواة

لا تعني شيء اما في

الـ B فالمساواة لها معنى .

Ex:- Toss a coin 50 times

$X \equiv$ # of heads observed.

Find 1) $p(X \geq 30)$

2) $p(20 \leq X \leq 32)$

3) $p(X \leq 27)$.

$X = B(50, \frac{1}{2})$

$$\left(P(X=x) = \binom{50}{x} (\tfrac{1}{2})^x (\tfrac{1}{2})^{1-x} \right)$$

$np = (50)(\frac{1}{2}) = 25$

$n(1-p) = (50)(\frac{1}{2}) = 25 \quad > 5$

$\rightarrow$ Use Normal Approximation to Binomail

With $\mu = np = 25$

$$\sigma = \sqrt{np(1-p)} = \sqrt{(50)(\tfrac{1}{2})(\tfrac{1}{2})} = \sqrt{12.5} = 3.5$$

$B(50, \frac{1}{2}) \rightarrow N(25, 3.5) \rightarrow N(0,1)$ (Z-score)

1- $P(X \geq 30) \approx P(X > 29.5)$. Correction Factor.

لدينا حالة المساواة وحتى ندخل 30 يجب ان نقللها بنص لنضمن الدخول.

$$= P\left(Z > \frac{29.5 - 25}{3.5}\right)$$

$$= P\left(Z > \frac{4.5}{3.5}\right)$$

$$= P(Z > 1.29)$$

$$= \tfrac{1}{2} - P(0 < Z < 1.29)$$

2) $P(20 < X \leq 32) \approx P(20.5 < Z < 32.5)$.

$$= P\left(\frac{20.5 - 25}{3.5} < Z < \frac{32.5 - 25}{3.5}\right)$$

3) $P(X \leq 27) \approx P(X \leq 27.5)$

4) $P(X < 27) \approx P(X < 26.5)$

5) $P(X \geq 27) \approx P(X > 26.5)$

6) $P(X > 27) \approx P(X > 27.5)$

7) $P(X = 27) \quad \sigma = 0 \approx P(26.5 < X < 27.5)$

$\binom{50}{27} (\frac{1}{2})^{27} (\frac{1}{2})^{23}$ القيمة الحقيقية.

$$= P\left(\frac{26.5 - 25}{3.5} < Z < \frac{27.5 - 25}{3.5}\right)$$

8) $P(X \geq 40) \approx P\left(Z > \frac{39.5 - 25}{3.5}\right)$

$\approx p(z > 5) \approx 0$

Ch 7

* Sampling Distributions
  - → population :- all elements of interest in a particular study.
  - → Sample :- any subset of the population.

* Inferential Statistics الاحصاء الاستنتاجية

Sample statistics $\xrightarrow{\text{Estimators}}$ population parmeters

$\bar{X}, S, r \longleftrightarrow \mu, \sigma, \rho$

* لماذا نختار Sample ولا نختار population :-

← لا يمكن حصر الجميع، لا يمكن لقاء الجميع والحصول على المعلومات المطلوبة كما انها مكلفة وتحتاج الى وقت

* Sampling :-
  - probability Sampling الجميع داخل
  - Non probability Sampling المعظم وليس الجميع داخل او هناك احتمالية دخوله.

1- Simple random Sample : A Sample Selected in such away that each element of the population has the Same Chance of being Selected.

2- Systematic (تنظيمية) Random Sample
  - arrange the elemets.
  - Starting point
  - Select an interval

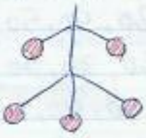
مثال على عينة رقم ٣ - ٣

1
2
$\boxed{3}$
x
x
x
x
x
$\boxed{9}$
x
x
x
x
x
$\boxed{15}$

3- Stratified random sample. العينة الطبقية العشوائية

population
↓
Subgroups.

4- Cluster Random sample
العينة العنقودية العشوائية (البسيطة).



تقسيم المناطق الى احياء ومنها تم نختار

١- المشكلة تحدد ، ٢- تحديد العينة ، ٣- تحديد الطريقة ،

→ Finite population → list, Roster.
→ infinite population → continuous process

اذا لم تكن للبيانات size population هذا يعني ان عدد الزبائن $\underline{N}$ اي غير منتهي

* Given a population with size N → Finite
Select a simple random sample of size $\underline{n}$ then number of such sample is $\binom{N}{n}$ and the probability of each sample is $\frac{1}{\binom{N}{n}}$

Ex:- Consider the following population.

$\overset{A}{20}, \overset{B}{30}, \overset{C}{40}, \overset{D}{30}, \overset{E}{50}$

Select a simple Random sample (SRS) of size 3

a) # of sample $= \binom{5}{3} = \frac{5!}{3!\,2!} = 10$

b) probability of each sample $= 1/10$

c) list all the sample.
d) Find the mean for each sample

		$\bar{x}$	$P(\bar{x})$
ABC	20,30,40	30	2/10
ABD	20,30,30	26.7	1/10
ABE	20,30,50	33.3	3/10
ACD	20,40,30	30	
ACE	20,40,50	36.7	2/10
ADE	20,30,50	33.3	
BCD	30,40,30	33.4	
BCE	30,40,50	40	
BDE	30,30,50	36.71	
CDE	40,30,50	40	2/10

Different sample → different sample means
∴ sample mean is a random variable.

probability distribution of the sample mean = sampling distribution of the mean.

R.V → 1) Shape
2) Expected value
3) Variance

$\sum \bar{x}\, p(\bar{x})$

The expected value for the sample mean is $E(\bar{x}) = \mu_{\bar{x}} = \sum \bar{x}\, p(\bar{x})$

x	p(x)
1	0.1
2	0.3
3	0.2
5	0.4

$Ex = \mu_x$

20, 30, 40, 30, 50

$\mu = \frac{20+30+40+30+50}{5} = 34$

$E(\bar{x}) = \mu$ $|\bar{x} - \mu| = 0$

↓ Sampling error.

$E(\bar{x}) = \mu_{\bar{x}} = \mu$

Give a population with mean $= \mu$ and Standard deviation $= \sigma$, then

1) $E(\bar{x}) = \mu_{\bar{x}} = \mu$

2) The standard error of the mean $\sigma_{\bar{x}}$ is given by

$$\sigma_{\bar{x}} = \begin{cases} \frac{\sigma}{\sqrt{n}} & \text{iF ① The population is inFinite. ② The population is Finite + } \frac{n}{N} \leq 0.05 \\ \sqrt{\frac{N-n}{N-1}} & \text{iF The population is Finite} \end{cases}$$

$\frac{n}{N} \leqslant 0.05 \leftarrow$ اذا كانت اكبر من 0.05 نستخدم المعادلة الثانية.

$n \rightarrow$ حجم العينة

$N \rightarrow$ populat حجم ال

Ex:- Given a population of size 5000 With $\mu = 100$, $\sigma = 20$

Find $\mu_{\bar{x}}$, $\sigma_{\bar{x}}$, iF

1) $n = 50$

2) $n = 100$

3) $n = 1000$

$\mu_{\bar{x}} = \mu = 100$

$\sigma_{\bar{x}} = \sqrt{\frac{N-n}{N-1}} \cdot \frac{\sigma}{\sqrt{n}}$ $\frac{n}{N} = \frac{50}{5000} = \frac{1}{100} < 0.05$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{20}{\sqrt{50}} = 2.82$$

2) $n = 100$ $\quad \frac{n}{N} = \frac{100}{5000} = 0.02$

$$\sigma_{\bar{x}} = \frac{20}{\sqrt{100}} = 2$$

3) $n = 1000$ $\quad \frac{n}{N} = \frac{1000}{5000} = 0.2 > 0.05$

$$\sigma_{x} = \sqrt{\frac{N-n}{N-1}} \cdot \frac{\sigma}{\sqrt{n}} = \sqrt{\frac{4000}{4999}} \cdot \frac{20}{\sqrt{1000}} =$$

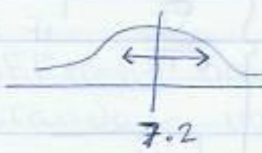
Ex:-

اسئلة على الوحدة ٦

a) Normal

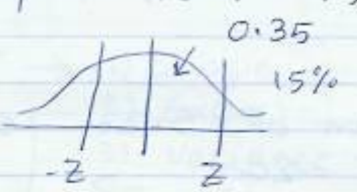
$\mu = 7.2$

$\sigma = 1.9$



$$P(6 \leq z \leq 10)$$
$$= P\left(\frac{6-7.2}{1.9} \leq Z < \frac{10-7.2}{1.9}\right)$$
$$= P(0 < Z < 1.47) + P(0 < Z < 0.63)$$

b)



$= -1.04$

Ex:-

$P = 0.8$ $\qquad B(250, 0.8)$

$n = 250$

$\left.\begin{array}{l} np \\ n(1-p) \end{array}\right\} \geq 5$

us. NA to B with $\mu = np = 200$

$\sigma = \sqrt{np(1-p)} = 6.3$

$$P(x > 210) \approx P(x > 210.5) =$$

210 - 200

6.3

*Sample Distribution of $\bar{x}$

Different Sample → different mean

Sample mean ——→ Random variable

- Expected value $E(\bar{x}) = \mu_{\bar{x}} = \mu$
- variance

$\sigma_{\bar{x}} = SE = \frac{\sigma}{\sqrt{n}}$ InF $\to \frac{n}{N} < 0.05$

$= \sqrt{\frac{N-n}{N-1}}$ $\frac{n}{N} \geq 0.05$

Shape of the distribution

IF the population is normal (From which the Sample taken)

then the Sampling distribution of $\bar{x}$ is normal regardless of the Sample size

→ Central limit theorem: CLT

IF we take SRS of size n From a population with mean $= \mu$ and Standard deviation $= \sigma$ Then the sampling distribution of $\bar{x}$ is approximately normal with mean $= \mu_{\bar{x}} = \mu$ and Standard deviation $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$ IF The sample is <u>Large</u>

large $= n \geq 30$

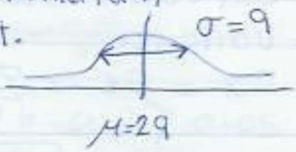
Ex:- Suppose that the student commuting time at a certain college, represented by the variable x has $\mu = 29$ and $\sigma = 9$ minutes. There are 6000 at the

College . Assume a SRS of Size 40 is selected.

1) Find the probability that the sample mean will be greater than 32 minit

$n = 40 > 30 \to \bar{x}$ has approximataly

$p(\bar{X} > 32)$ anormal dist.



$\mu_{\bar{x}} = \mu = 29$

$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{9}{\sqrt{40}} = 1.42$

$\frac{40}{6000} = 0.006 < 0.05$

$Z_{\bar{x}} = \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}}$ نستخدم المعادلة الأولى

$= P\left(Z > \frac{32-29}{1.42}\right) = 2.1$

$= \frac{1}{2} - P(0 < Z < 2.1)$

$= \frac{1}{2} - 0.4821$

2) Find the Propability that the Sample mean will be within 2 minutes (error) of the population mean.

$P(\mu - 2 < \bar{x} < \mu + 2)$

$= P\left(\frac{(\mu-2)-\mu}{\sigma_{\bar{x}}} < Z < \frac{(\mu+2)-\mu}{\sigma_{\bar{x}}}\right)$

$= P\left(\frac{-2}{\sigma_{\bar{x}}} < Z < \frac{2}{\sigma_{\bar{x}}}\right)$

$P\left(\frac{-2}{1.42} < Z < \frac{2}{1.42}\right)$

$2P\left(0 < Z < \frac{2}{1.42}\right)$

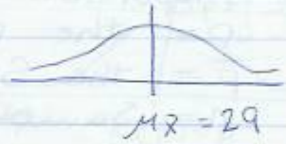
$= 2P(0 < Z < 1.41)$

← نعوضها على القانون

$2P(0 < Z < \frac{\epsilon}{\sigma_{\bar{x}}}) = (2)(0.4207) = 0.8414$

3) recalculate the probability in using a sample size $= 100$

$\sigma_{\bar{x}} = \frac{9}{\sqrt{100}} = 0.9$



$= 2P(0 < Z < \frac{2}{0.9}) =$

$= 2P(0 < Z < 2.22)$

$= 2(0.4868)$

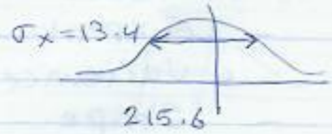
Ex:- $\left.\frac{30}{255}\right) \rightarrow \mu = \$\ 215.6 \rightarrow$ avarage

$\sigma = \$\ 85$

select SRS of size $n = 40$

a) $n = 40 > 30 \xrightarrow{CLT}$ The sampling distribution of $\bar{x}$ is approximataly normal with $\mu_{\bar{x}} = 215.6$ and standard deviation $= \frac{\sigma}{\sqrt{n}} = \frac{85}{\sqrt{40}} = 13.4$

b) $2P(0 < Z < \frac{20}{13.4})$



c) $2P(0 < Z < \frac{10}{13.4})$

$P(\bar{x} > 220) \quad \frac{220 - 215.6}{13.4}$

نسبة

* Sampling Distribution of the proportion

total : 6000
$E = 1000$
proportion → $E = \frac{1000}{6000} =$

Proportion ≡ percentage
p ≡ the population proportion.
$\bar{p}$ ≡ the Sample proportion.
n ≡ Sample Size
N ≡ population size
X ≡ # of elements of interest.
$p = \frac{X}{N}$, $\bar{p} = \frac{X}{n}$

$\bar{p}$ is apoint estimate For p
↓ Statistic ↓ parameter.

$\bar{p}$ → statictic
p → parameter.
$|\bar{p} - p|$ sampling error.

different Sample → different proportions
$\bar{p}$ is a random variable.
- Expect value
- Variance
- Shape.

IF we select a SRS of Size n. From apopulation with a proportion p, then The sample proportion $\bar{p}$ will be a arandom variable with $E(\bar{p}) = \mu_{\bar{p}} = p$ and standard error oF deviation.

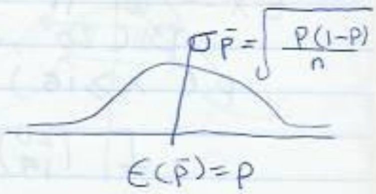
$$\sigma_{\bar{p}} = \sqrt{\frac{p(1-p)}{n}}$$

* Centeral limit Theorem:

The sampling distribution of $\bar{p}$ is approximately normal with $E(\bar{p}) = p$ and $\sigma_{\bar{p}} = \sqrt{\frac{p(1-p)}{n}}$ if the sample size is

Large : Large: $np \geq 5$

$n(1-p) \geq 5$



Ex:- $\frac{3}{7}$) $70\% = p$

$p = 0.7$ proportion of the population.

a) $n = 100$

1) more Than 0.58

$P(\bar{p} > 0.58)$

$np = (0.7)(100)$

$n(1-p) = (0.3)(100)$ > 5

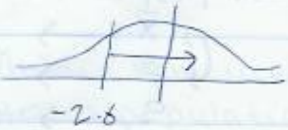
$\rightarrow \bar{p}$ has a pproximately anormal distribution with $E(\bar{p}) = 0.7$

$\sigma_{\bar{p}} = \sqrt{\frac{(0.7)(0.3)}{100}} = 0.045$

a) $P(\bar{p} > 0.58) = P(Z > \frac{0.58 - 0.7}{0.045})$

$= P(Z > -2.6)$

$= \frac{1}{2} + P(0 < Z < 2.6)$



$Z_{\bar{p}} = \frac{\bar{p} - p}{\sigma_{\bar{p}}}$

b) between 0.6 and 0.75

$P(0.6 < \bar{p} < 0.75)$

$= P\left(\frac{0.6 - 0.7}{0.045} < Z < \frac{0.75 - .7}{0.045}\right)$

c) within 0.08 of the population.

proportion
$= 2P\left(0 < Z < \frac{0.08}{0.045}\right)$

d) $P(\bar{p} < 0.6)$

Ex:-
$B(\overset{n}{20}, \overset{p}{0.4})$

$P(X \geq 15) = P(X = 15) + \dots + P(X = 20)$

$\binom{20}{15}(0.4)^{15}(0.6)^{5} + \dots$

$np, n(1-p) \geq 5$

$\underset{8}{(20)(0.4)}, \underset{12}{(20)(0.6)} \geq 5$

normal app. to Bin.

$P(X \geq 15) \simeq P(X > 14.5)$

$= P\left(Z > \frac{14.5 - 8}{\sqrt{(20)(0.4)(0.6)}}\right)$

$P(X \geq 15) \qquad X = 15 \longrightarrow \frac{15}{20} = 0.67$

$P(\bar{p} > 0.67) = P\left(Z > \frac{0.67 - 0.4}{\sqrt{\frac{(0.4)(0.6)}{20}}}\right)$

* properties of Estimators.

$\bar{X} \longrightarrow \mu$ (ميو)

$\bar{p} \longrightarrow p$

$S \longrightarrow \sigma$ (سجمة)

let $\bar{\alpha} \longrightarrow \alpha$ (ألفا)

Statistic $\longrightarrow$ parameter.

$|\bar{\alpha} - \alpha| \rightarrow$ sample error.

1- Unbaisdness غير متحيز

$\bar{\alpha}$ is Unbaised estimator For α if

$E(\bar{\alpha}) = \alpha$

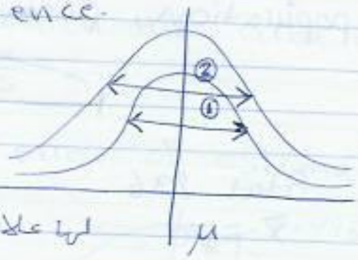
$E(\bar{x}) = \mu$

$E(\bar{p}) = p$ $\quad \bar{X}, \bar{P}$ Unbaised

2. Efficiency الفاعلية

The statistic with smaller standard deviation is more efficience.

3. Consistency الثبات



$\bar{x}$ $\searrow$ $\quad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$ $\quad \bar{x}_1$, $\bar{x}_2$

as $n \uparrow$ $\quad SE \downarrow$

لها علاقة بزيادة العينة ومقدار الخطأ

* Interval Estimation

Ch. 8

Inferential Statistics.

- Estimation
  - point (نقطة) اي نقطة معامل نقطة. Ch.7
  - Interval (فترة) Ch.8
- Hypothesis testing Ch.9

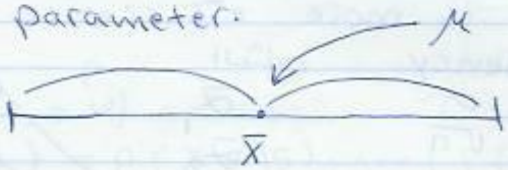
constructe an enterval (a, b) $\downarrow \mu$

point $\quad$ interval

Estimation: A procedure by which numerical value or values are assigned to a population parameter based on the information collected from a sample

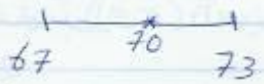
value + values $\equiv$ Estemators.

* Interval Estemation :- In terval estimation, an interval is constructed around a point estimate and it stated that this interval is likely to contain the corresponding population parameter.



Stat 236

$\bar{X} = 70$



Each interval is constructed with regard to a <u>GIVEN</u> confidence level.

$1-\alpha$ 90%, 95%, 99%

اذا لم يذكر في الامتحان معامل الثقة (قيمته) فإننا نستخدم 95% كمتوسط.

$1-\alpha = 95\%$
↓ 5%

$1-\alpha = 96\%$
↓ 4%

Confidence Interval

1) μ → large sample, → small

2) p → large, → small

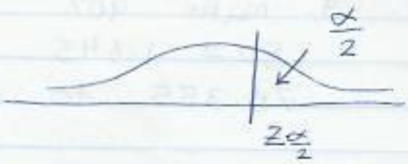
Confidence Interval For μ large Sample case $(n \geq 30)$

The $(1-\alpha)\,100\%$ confidence interval For μ is given by

$$\bar{X} \pm Z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \quad , \quad \bar{X} \pm Z_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}}$$

σ is known σ is unknown

$\bar{X} \equiv$ Sample mean
$n \equiv$ Sample Size
$\sigma \equiv$ population S.d
$S \equiv$ Sample S.d
$Z_{\frac{\alpha}{2}} \equiv$ The value of Z For which the area is The upper tail is $\frac{\alpha}{S}$

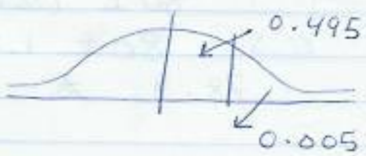


$E = Z_{\frac{\alpha}{2}} \frac{\alpha}{\sqrt{n}}$ maximum error of estimate.

Ex:-

$99\% \rightarrow \alpha = 1\% \quad = \frac{\alpha}{2} = \frac{1}{2}\%$

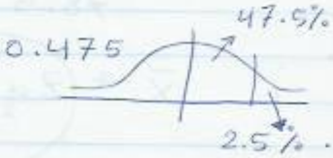
$Z_{\frac{\alpha}{2}} = Z_{0.005} =$



$2.575 \rightarrow 0.4949 \quad 0.4951$

$95\% \rightarrow \alpha = 5\% = \frac{5}{2} = 2.5\% = .025$

99%	2.575	
95%	1.96	
90%	1.645	



Ex:- Given a population With $\sigma = 6$ A SRS is selected gave $\bar{X} = 80$

1. Make a 99% confidence interval $n = 36$
   $80 \pm (2.575)(\frac{6}{6})$
   80 ± 2.575
   77.425 to 82.575 true من 99% احتمالية
   means ان يكون ال populat في هذه الفترة.

2. Make a 95% confidence interval, $n = 36$
   80 ± 1.96
   78.04 to 81.96

3. make 90%

80 ± 1.645

78.355 to 81.645

4. width of the interval $\leftrightarrow$ confidence level

((( )))

77.455 78.04 78.355 81.645 81.96 82.575

5. make 99% confidence , n = 81 level

$80 \pm (2.575)\left(\frac{6}{9}\right)$

80 ± 1.7

78.3 to 81.7

6. Make a 99% . n = 100

$80 \pm (2.575)(0.6)$

80 ± 1.5

78.5 to 81.5

$\bar{X} \pm Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$ maximum error of estimate = ϵ

$$\epsilon = Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \Rightarrow \sqrt{n} = \frac{Z_{\frac{\alpha}{2}}\sigma}{\epsilon} \Rightarrow n = \left(\frac{Z_{\alpha/2}\sigma}{\epsilon}\right)^2$$

هذا قانون إيجاد العينة (الحجم).

$Z_{\frac{\alpha}{2}} \rightarrow$ Given.

$\epsilon \rightarrow$ Given.

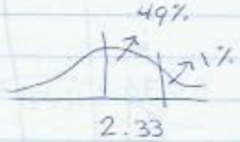
Pilot Study . عينة تجريبية $\sigma = \frac{\text{Range}}{4}$

EX:- $\frac{36}{297}$) $\sigma = 6$ minutes.

a) How large asample $n = ?$
0.98 probability
to within 2 minutes or less

$\sigma = 6$ $\qquad Z_{\alpha/2} = 2\% = 1\% \longrightarrow 98\%$

$E = 2$ $\qquad Z_{\frac{\alpha}{2}} = 2.33$



$$n = \left(\frac{(2.33)(6)}{2}\right)^2 = 49$$

b) $n = 49$, $\bar{x} = 32$

98% confidence interval

$$32 \pm (2.33)\frac{6}{\sqrt{49}}$$

$$32 \pm (2.33)(6/7)$$

30 to 34

* Interval Estimation For μ :

$n \geq 30$ (CLT)

σ Known $\qquad \bar{x} \mp Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$

σ unknown $\qquad \bar{x} \pm Z_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}}$

Small Sample case $n < 30$

population is Normal

population is not Normal

Suppose that the population is normal

σ is Known $\qquad \bar{x} \pm Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$

σ is not Known $\qquad \bar{x} \pm t_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}}$ t-distribution (t-table)

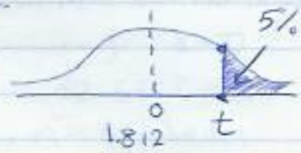
Z ① $n \geq 30$

② $n < 30$ + σ known + Normal

t 1) $n < 30$ + σ unknown + Normal

upper-Tail area

degrees of Freedom	0.1	0.05	0.025	0.01	.005
1					
2					
⋮					
10		1.812			
30					
⋮					
120					
⋮					
∞					



$dF = n-1$

$\bar{X} \equiv$ Sample mean

$S \equiv$ Sample S.d

$n \equiv$ Sample Size

$t_{\frac{\alpha}{2}} \equiv$ The value of t (From t-table) For which The area in the upper tail is $\frac{\alpha}{2}$ with $dF = n-1$

Ex:- For a t-distribution with $dF=12$ Find The area that is in each region

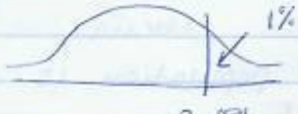
ⓐ to the left of 1.782

100% − 5% = 95% = left ←

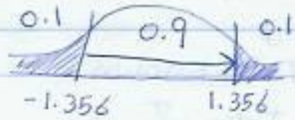
ⓑ to the right of −1.356

ⓒ right 2.681

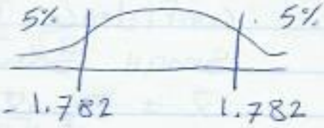
ⓐ



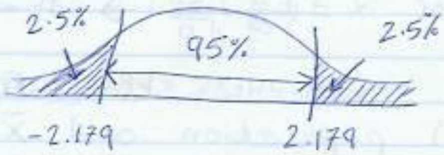
ⓑ



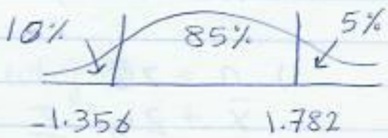
ⓓ left -1.782



ⓔ between -2.179 to 2.179



ⓕ -1.356 to 1.782



Ex: A random sample of delivery times for couries service produced the following results (in hours).

7.5 3 4.2 6 10.1 5 6.3

ⓐ Construct a 95% confidence interval for the average delivery time for this courier services?

$\bar{X} = \frac{\Sigma X}{n} = 6.1$

$S = \sqrt{\frac{\Sigma(X-\bar{x})^2}{n-1}} = 2.33$

→ Assume that the population is normal

Interval $\bar{x} \pm t_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}}$

$95\% \rightarrow \alpha = 5\% \longrightarrow \frac{\alpha}{2} = 2.5\%$

$t_{0.025} = 2.447$ $\quad dF = 7-1 = 6$

Int: $6.1 \pm (2.447)(\frac{2.33}{\sqrt{7}}) =$

6.1 ± 2.15

3.95 to 8.25

Confidence Interval
Small Sample case

$\bar{x} \pm Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$ or $\bar{x} \pm t_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}} \rightarrow dF = n-1$

Ex:- A random sample is selected From a $N(\mu, \underset{\sigma}{8})$ population and $\bar{x} = 21.2$ is obtained construct (أنشئ) a 95% confidance interval For μ in the Following cases.

1) $n = 25$ normal + σ known + $n < 30$

$\bar{x} \pm Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} = 21.2 \pm 1.96 \frac{8}{5}$

3.136

18.2 to 24.2

2) $n = 100$ $21.2 \pm 1.96 \frac{8}{10}$

1.568

19.632 to 22.768

3) $21.2 \pm 1.96 \frac{8}{20}$

0.784

20.416 to 21.984

Ex:- BZU , Al Najah $\longrightarrow$ GPA

	BZU	Alnajah
Sample Size	64	81
$\bar{x}$	76	78
S	10	12

Construct 95% confidence interval for

μ_{BZU} , $\mu_{Alnajah}$ $\qquad \bar{x} \pm Z_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}}$

BZU : $76 \pm 1.96 \left(\frac{10}{8}\right)$

2.45

73.55 to 78.45

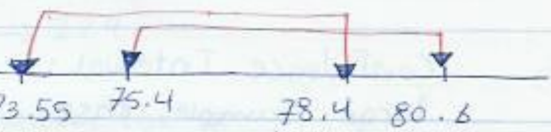
Alnajah : $78 \pm 1.96 \left(\frac{12}{9}\right)$

2.6

75.4 to 80.6

What is your conclusion? why ??

اذا كان هناك تقاطع هذا
يعني صعوبة الوصول إلى
نتيجة لان قد تأتي حالة
تكون فيها النجاح أقل من بيرزيت وبالتالي اذا اردنا المقارنة
يجب ان لا يكون لدينا تقاطع .



Ex: $\frac{25}{293}$ → 20 , 20 , 28 , 6 , 11 , 17 , 23
16 , 22 , 18 , 10 , 22 , 29 , 19 , 32

b) assume that the population is normal

a) compute 95% confidence interval.

$\bar{x} = \frac{\Sigma x}{n} = \bar{x} = 19.5$ $\qquad$ [2] $\quad$ [3]

$S = 7.1$ $\qquad$ $x\sigma n$ $\quad$ $x\sigma n-1$

population $\quad$ Sample

$19.5 \pm 1.96 \frac{7.1}{\sqrt{15}}$

$19.5 \pm \boxed{3.6}$

15.9 to 23.1

قيمة الخطأ هو 3.6 وهنا ان
السؤال يطلب ان يكون الخطأ
فقط 2 نقوم بزيادة الـ sample (عددها).

c) $E = \pm 2$ $\qquad Z_{\frac{\alpha}{2}} = 1.96$

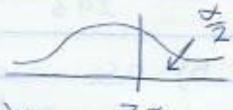
قيمة الخطأ

* Confidence Interval For aproportion:

$$\mu \rightarrow \bar{x} \pm Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$$
$$\bar{x} \pm t_{\frac{\alpha}{2}} \, s/\sqrt{n}$$

$P \rightarrow$ propulation proportion.
$\bar{P} \rightarrow$ Sample proportion.
$\bar{P} \rightarrow$ is apointe estimate For P
C.L.T n large $\rightarrow \bar{P}$ have N.D
$np \geq 5$ with $E(\bar{p}) = P$
$n(1-p) \geq 5$ $\quad \sigma_{\bar{p}} = \sqrt{\frac{P(1-P)}{n}}$

Confidence Interval:
large Sample case $np \geq 5 - n(1-p) \geq 5$
The $(1-\alpha)\,100\%$. Confidence interval For P is given by $\bar{p} \pm Z_{\frac{\alpha}{2}} \sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$
$\bar{P} \equiv$ Sample proportion
$n \equiv$ Sample Size
$Z_{\frac{\alpha}{2}} \equiv$



$\in \equiv$ maximum error of estimate $\quad Z_{\frac{\alpha}{2}}$

$$= Z_{\frac{\alpha}{2}} \sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$$

$$\frac{n}{N} > 0.05 \rightarrow F.P.C.F = \sqrt{\frac{N-n}{N-1}}$$

$$\bar{p} \pm Z_{\frac{\alpha}{2}} \boxed{\sqrt{\frac{N-n}{N-1}}} \sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$$

Finite $\leftarrow$ نضعها في القانون إذا كان حجم الpopulation

Ex: A Sample of 400 observation taken From apopulation produced asample proportion of 0.63 make a 95% confidence

interval For P
→ Infinite

Interval $\bar{p} \pm Z_{\frac{\alpha}{2}} \sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$ = 95% احتمال قدرة
لل ... population True prop...
موجودة في مدة النسبة (بين الرقمين)

$$0.63 \pm (1.96)\sqrt{\frac{(0.63)(0.37)}{400}}$$

$$0.63 \pm 0.05$$

$$P \rightarrow 0.58 \text{ to } 0.68$$

Ex:- #2 → Sample → 400
life The product → 224

ⓐ point estimate.
$\bar{p}$ is apoint estimate for p

$$\bar{p} = \frac{224}{400} = 0.56$$

ⓑ 95% confidence Interval

$$0.56 \pm 1.96\sqrt{\frac{(0.56)(0.44)}{400}}$$

$$0.56 \pm 0.05$$

$$0.51 \text{ to } 0.61$$

→ دراسة سابقة .
→ pilot study
→ $P = \frac{1}{2}$

$$n = \frac{Z_{\frac{\alpha}{2}}^2 \, P(1-P)}{\epsilon^2} = \frac{(Z_{\frac{\alpha}{2}})^2}{(\epsilon)^2} P(1-P).$$

Ex:- n = 150 , 105

a) 95% For p

$$\bar{p} = \frac{105}{150} = 0.7$$

Interval : $0.7 \pm 1.96\sqrt{\frac{(0.7)(0.3)}{150}}$

$$0.7 \pm 0.073$$

الساعة 10
يوم الجمعة
محاضرة

b) 0.05 . 99% 0.7

$$n = \left(\frac{Z_{\frac{\alpha}{2}}}{E}\right)^2 p(1-p) \qquad Z_{\frac{\alpha}{2}} = 2.575$$

$$E = 0.05$$

$$p = 0.7$$

$$0.5$$

$= 557$

$= 663$

* Hypothesis Testing. **Ch.9**

Inf. Stat

- Estimation
  - point
  - Interval
- Hyp. tes
  - claim $\leftrightarrow$ population parameter.
    - Reject
    - Accebt.

* Hypothesis :- A statement about population parameter for the purpose of testing

* Hypothesis testing : A procedure based on sample information and probability theory to detemine whether the hypothesis is reasonable statement.

* Null Hypothesis :- A Stetment or "Claim" about apopulation parameter. that is assumed to be true until it is declared False.

عكسها H_o → H_a

μ_o

اصد يدعي $\mu = 85$ ← و يبقى صحيح مالم ينكره احد

اذا كانت H_o صحيحة هذا يعني ان H_a خطأ والعكس صحيح.

H_a → research Hypothesis.

$H_o \rightarrow$ Regect $\rightarrow H_o$, Accept $\rightarrow H_o$

Null Hypothesis	Researcher: Accept H_o	Researcher: Regect H_o
H_o is true	correct conclusion	Type I error.
H_o is Fales	Type II error.	correct conclusion.

Type I error : Regect H_o when it is true.
Type II error : Accept H_o when it is False.

P_r (type I error) $= \alpha \rightarrow$ Given.
P_r (type II error) $= B \rightarrow$??
Given $\alpha = 1\%, 5\%, 10\%.$
power of the test $= 1 - B$

* 5 Steps :-

1- State the null and the alternative hypothesis; H_o, H_a

2- Select a level of significant.
$\alpha = ??$ $\quad$ pr (type I error).

3- Idintify a statistic test
Select a distribution to use Z, t

Statistic test $\Rightarrow$ Z test

$$Z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}$$

4. Formulate a rejection rule.
large sample case for μ. $(n \geq 30)$.
let μ_0 be a given value for μ in the null Hypothesis, we have one of the Following cases. (α isgiven)

* Hypothesis.

1- $H_o \rightarrow \mu = \mu_0$, $H_a \rightarrow \mu \neq \mu_0$

2- $H_0 \rightarrow \mu \geq \mu_0$

$H_a \rightarrow \mu < \mu_0$

3- $H_0 \rightarrow \mu \leq \mu_0$

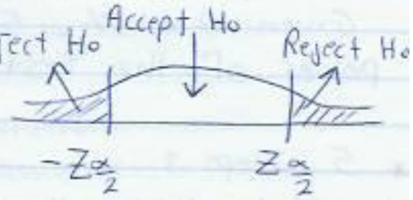
$H_a \rightarrow \mu > \mu_0$

* Statistic test $\rightarrow Z = \dfrac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}$

$S \leftarrow \sigma$

* Rejection Rule

reject $H_0 \rightarrow$ IF $Z > Z_{\frac{\alpha}{2}}$ or $Z < -Z_{\frac{\alpha}{2}}$

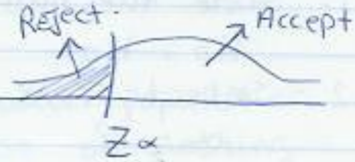


→ $Z_{\frac{\alpha}{2}}$, $-Z_{\frac{\alpha}{2}}$

Critical values.

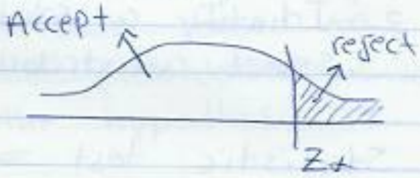
→ reject $H_0 \rightarrow$ IF $Z < -Z_\alpha$

one tailed test



→ reject $H_0 \rightarrow$ IF $Z > Z_\alpha$

one tailed test.



Ex:- at least \$ 35000

$H_0 \rightarrow \mu \geq 35000$, $\bar{X} = 33124$

Sample Size $= n = 150$, $S = \$\ 5400$

$\alpha = 1\%$

1- $H_a : \mu < 35000$

2- $n = 150 \rightarrow Z$ - test

$Z = \dfrac{\bar{X} - \mu_0}{S / \sqrt{n}} = \dfrac{33124 - 35000}{5400 / \sqrt{150}} = -4.25$

3- reject H_0 IF $Z < -Z_\alpha$

$\alpha = 1\% \rightarrow Z_\alpha = 2.33$

4- since $Z = -4.25 < -Z_\alpha = -2.33$
reject H_0

Yes, They should not Opean.

Ex:- ① $\begin{cases} H_0 = \mu = 50 \\ H_a = \mu \neq 50 \end{cases}$, Sample size $= n = 64$
, $\bar{X} = 53$
$\alpha = 0.01$, $S = 12$

p-value
Confidence Interval.

2. $n = 64$.

$$Z = \frac{\bar{X} - \mu}{S/\sqrt{n}} = Z = \frac{53-50}{1.5} = +2$$

3. rejection region :- reject H_0 if
$Z > Z_{\frac{\alpha}{2}}$ or $Z < -Z_{\frac{\alpha}{2}}$

$\alpha = 1\% = Z_{\frac{\alpha}{2}} = 2.575$.

$$-2.575 < 2 < 2.575$$

Accept H_0

p-value $\equiv pr(Z > Z_t)$
$\equiv pr(Z > 2) = \frac{1}{2} - pr(0 < Z < 2)$
$= \frac{1}{2} - 0.4772 = 0.0228$.

rejection rule :- reject H_0 if p-value $< \alpha \left(\frac{\alpha}{2}\right)$
p-value $p = 0.0228 > 0.005$
Accept H_0

* تلخيص الخطوات :-

1- H_0, H_a
2- Z
3- rejection rule. $(pr(Z > Z_t))$.
p-value $< \alpha$

Construct a 99% confidence interval

$\bar{x} = Z_{\frac{\alpha}{2}} \, S/\sqrt{n}$

$53 \pm (2.575)(1.5)$

53 ± 3.9

t (49.1 to 56.9)

$\mu_0 = 50$

Accept H_0.

Assume $n < 30$

IF the population is normal → sample → normal

→ IF σ is known → Z-test. $Z = \dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$ →

rejection rule.

reject H_0 if $Z < -Z_\alpha$

→ IF σ is unknown → t-test $t = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}}$

reject H_0 if $t < -t_\alpha$

$dF = n-1$

$n < 30$ + Normal + σ unknown → t

$n > 30$ + Normal + σ known → Z

H_0 :- $\mu = \mu_0$ tow tailed

H_a :- $\mu \neq \mu_0$

reject rule.

reject H_0 if $t < -t_{\frac{\alpha}{2}}$ or $t > t_{\frac{\alpha}{2}}$

Ex:- $\frac{34}{341}$ → $H_0 = \mu = 20$

$H_a = \mu \neq 20$

Sample = 18, 20, 16, 19, 17, 18

$n = 6$, $\bar{x} = 18.$, $S = 1.4$

$n < 30$ t-test.

Assume the population is normal.

normal + $n = 6$, < 30 + σ unknown = t-test

Statestic test :- $t = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}} = -3.8$

reject H_0 if $t > t_{\frac{\alpha}{2}}$ or $t < -t_{\frac{\alpha}{2}}$ $dF = 5$

$t_{\frac{\alpha}{2}} = 2.571$ Since $t = -3.8 < -t_{\frac{\alpha}{2}} = -2.571$ reject H_0

* p-value = $p(t > zt)$

$p(t > 3.8)$

0.01		0.005
3.365	3.8	4.032

reject H_0 if p-value $< \frac{\alpha}{2}$ وتكون بطريقة استنتاجية لذلك يفضل استخدامها في حالة Z .

Ex:- Is the temperature required to damage a computer on the average less than 110 degree ?.
Because of the price of computer, 20 computer were tested to see what minimum temp. will damage the computer $\bar{x} = 109$ degree. $S = 3$ dagree , use $\alpha = 5\%$. what is your conclusion.

دائماً فيها ← $H_0: \mu \geq 110$

مساواة $H_a: \mu < 110$ الباحث يريد أن يثبت هذه النظرية إذاً هذا هو Ha ويعني انه تحكم العكس .

$n = 20$ + Normal + σ unknown.

$t = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}} = -1.49$

reject H_0 :- IF $t < -t_\alpha$

$t_\alpha = 1.72$ $dF = 19$

Conclusion :- Since $t = -1.49 > t_\alpha = 1.72$

Accept H_0

* Tests About the population proportion.

$$\mu \begin{cases} L.S.C \\ S.S.C \end{cases}$$
H_0 :- $\mu = \mu_0$
H_a :- $\mu \neq \mu_0$ } proportion.

$\bar{p}$:- Sample proportion.
p :- population proportion.

$\bar{p} \to p$ $\qquad \bar{p} \pm Z_{\alpha/2}\, \sigma_{\bar{p}}$

$$\sigma_{\bar{p}} = \sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$$

let p_0 be given value for p in H_0

① H_0 :- $p = p_0$ ② H_0 :- $p \geqslant p_0$ ③ H_0 :- $p \leqslant p_0$
H_a :- $p \neq p_0$ $\quad H_a$:- $p < p_0$ $\quad H_a$:- $p > p_0$

two tailed test $\qquad$ one tailed $\qquad$ onetailed

$np \geqslant 5$, $n(1-p) \geqslant 5 \xrightarrow{C.I.T}$ Sampling dist. of $\bar{p}$ is approximately normal $\to$
Statistic test $\to$ Z - test.

$$Z = \frac{\bar{p} - p_0}{\sigma_{p_0}} \qquad \sigma_{p_0} = \sqrt{\frac{p_0(1-p_0)}{n}}$$

rejection rule :-
② reject H_0 if
$Z < -Z_\alpha$

Ex: $\frac{44}{347}$) $\rightarrow$ H_0: $P = 0.2$
H_a: $P \neq 0.2$

Sample: $n = 400$, $\bar{p} = 0.175$

a) $\alpha = 0.05$ Rejection Rule

reject H_0 if $Z > Z_{\frac{\alpha}{2}}$ or $Z < -Z_{\frac{\alpha}{2}}$

$\alpha = 5\% \rightarrow Z_{\frac{\alpha}{2}} = Z_{0.025} = 1.96$

reject H_0 if $Z < -1.96$ or $Z > 1.96$.

b) Compute Z

$$Z = \frac{\bar{p} - p_0}{\sigma_{p_0}} = \frac{0.175 - 0.2}{0.02} = -1.25$$

c) Find P-value =

$Pr(Z > 1.25)$

$= \frac{1}{2} - Pr(0 < Z < 1.25)$

$= \frac{1}{2} - 0.3944$

$= 0.1056$

d) Concl.

Since $-1.96 < 1.25 < 1.96$

الرقم 1.25 محصور بين المنطقتين لما ان

1.96 أو 1.96- ← reject وتأتي بينهما

اذاً هي Accept.

accept H_0.

reject H_0 if P-value $< \frac{\alpha}{2}$

$0.1056 > 0.025$

Accept H_0.

Ex: $\frac{55}{347}$) $\rightarrow$ at least 20%

$n = 596$, 83

H_0 :- $p \geq 20\%$
H_a :- $P < 20\%$
$Z_{0.05} = 1.96$

$\bar{p} = \frac{83}{596} = 0.139$

تكمل السؤال كالبقية

* حل اسئلة خارجية :-

1) H_0 :- $\mu = 264$ $\quad n = 10$
H_a :- $\mu > 264$ $\quad \bar{x} = 270$
$S = 15$

Assume the population is normal
$n = 10$ + normal + σ unknown = t

2) H_0 :- $\mu \geq 90.000$
H_a :- $\mu < 90.000$
reject H_0

3) $P = 0.7$, $\bar{p} = \frac{160}{200} = 0.8$
H_0 :- $P \leq 0.7$
H_a :- $P > 0.7$
$n = 200$, 160

$$Z = \frac{\bar{p} - P_0}{\sigma_{p_0}} = \frac{0.8 - 0.7}{\sqrt{\frac{(0.7)(0.3)}{200}}} = 3.1$$

reject H_0 if $Z > Z_\alpha$
$\alpha = 1\% \rightarrow Z_\alpha = 2.33$
Since :- $Z = 3.1 > 2.33$
reject H_0

4) H_0 :- $P = 0.15$ $\quad \bar{p} = 0.17$
H_a :- $P > 0.15$ $\quad n = 1000$

$$Z = \frac{\bar{P} - P_0}{\sigma_{P_0}} = \frac{0.17 - 0.15}{\sqrt{\frac{(0.15)(1-0.15)}{1000}}}$$

$\alpha = 10\%$

تكملة حل رقم (1)

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{270 - 264}{15/\sqrt{10}} = 1.265$$

reject H_0 if $t > t_\alpha$

$\alpha = 0.1 \rightarrow t_\alpha =$ $dF = 9$

$t_\alpha = 1.383$

Since $t = 1.265 < t_\alpha = 1.383$

Accept H_0, no

* Inferences about two population. Ch. 10

consider two populations.

population 1 with mean μ_1 and S.d σ_1

population 2 with mean μ_2 and S.d σ_2

Select two Independant samples from the two population

Sample one from population 1: $\bar{X}_1$, S_1

sample two from population 2: $\bar{X}_2$, S_2

The difference between the two mean.

$\mu_1 - \mu_2$ is aparameter.

$\bar{X}_1 - \bar{X}_2$ is the point estimation for $\mu_1 - \mu_2$

statistic.

$\bar{X}_1 - \bar{X}_2$ is arandom variable.

$$E(\bar{X}_1 - \bar{X}_2) = \mu_{\bar{x}_1 - \bar{x}_2} = \mu_1 - \mu_2$$

$\sigma_{\bar{x}_1 - \bar{x}_2} \equiv$ Standard error.

$$\sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

C.L.T $\rightarrow$ $n_1, n_2 \geq 30$

* Confidence Interval For $\mu_1 - \mu_2$

The $(1-\alpha)100\%$ confidence interval For $\mu_1 - \mu_2$ (large Sample case)

$$(\bar{x}_1 - \bar{x}_2) \pm Z_{\alpha/2}\, \sigma_{\bar{x}_1 - \bar{x}_2}$$

Ex:- Two independent Samples taken From two population gave the Following results

Sample 1	Sample 2
$n = 50$	$n_2 = 35$
$\bar{x}_1 = 13.6$	$\bar{x}_2 = 11.6$
$S_1 = 2.2$	$S_2 = 3$

ⓐ What is the point estimate of the different between the two population means.

$\bar{x}_1 - \bar{x}_2 = 13.6 - 11.6 = 2$ is a point estimate For $\mu_1 - \mu_2$

ⓑ a 90% confidence interval (construct)

ⓒ 95% confidence interval

$$\sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}} = \sqrt{\frac{(2.2)^2}{50} + \frac{(3)^2}{35}}$$

$$\sqrt{\frac{4.84}{50} + \frac{9}{35}} = \sqrt{0.0968 + 0.257} =$$

0.6

interval :-

$\bar{x}_1 - \bar{x}_2 \pm Z_{\alpha/2}\, S_{x_1 - x_2}$ $\quad 2 \pm (1.645)(0.6)$

2 ± 1 1 to 3

$\mu_1 - \mu_2 \in (1, 3)$

$\mu_1 > \mu_2$

$\mu_2 - \mu_1 \in (-3, -1)$

$\mu_2 - \mu_1 < 0$

$\mu_2 < \mu_1$

* Small Samples case.

n_1 or $n_2 < 30$

$n_1 < 30$ and/or $n_2 < 30$

- Assume that the two population are normal
- Assume that $\sigma_1^2 = \sigma_2^2 = \sigma^2$
- (σ^2) IF σ^2 is known. then

$(\bar{X}_1 - \bar{X}_2) \pm Z_{\frac{\alpha}{2}} \sigma_{\bar{X}_1 - \bar{X}_2}$

$$\sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_1^2}{n_2}} = \sqrt{\sigma^2 \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

IF σ^2 is unknown.

$S_p \equiv$ pooled estimater For σ

$$S^2_p = \frac{(n_1 - 1) S_1^2 + (n_2 - 1) S_2^2}{n_1 + n_2 - 2}$$

$$S_{\bar{x}_1 - \bar{x}_2} \equiv \sqrt{Sp^2 \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

* Hypothesis testing about difference between two means :-

Hypothesis

let μ_0 be agive value For μ is H_0

① H_0 :- $\mu_1 - \mu_2 = \mu_0$

H_a :- $\mu_1 - \mu_2 \neq \mu_0$

② H_0 :- $\mu_1 - \mu_2 \geq \mu_0$

H_a :- $\mu_1 - \mu_2 < \mu_0$

③ $H_0 :- \mu_1 - \mu_2 \leq \mu_0$
$H_a :- \mu_1 - \mu_2 > \mu_0$

* large sample case
statistics test μ_0

$$Z = \frac{(\bar{X}_1 - \bar{X}_2)(\mu_1 = \mu_2)}{\sigma^2(\bar{X}_1 - \bar{X}_2)}.$$

Ex:- consider the Following hypothesis test

H_0 :- $\mu_1 - \mu_2 \leq 0$

H_a :- $\mu_1 - \mu_2 > 0$

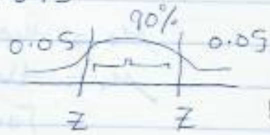
the results For two independent samples taken From the two population are

Sample -1-	Sample -2-
$n_1 = 40$	$n_2 = 50$
$\bar{x}_1 = 25.2$	$\bar{x}_2 = 22.8$
$S_1 = 5.2$	$S_2 = 6$

① use $\alpha = 0.05$ what is your conclusion

② $n_1, n_2 \geq 30 \rightarrow$ Z-test.

③ reject H_0 if $Z > Z_\alpha$ $\quad Z_\alpha = 1.645$



④ $$Z = \frac{(\bar{x}_1 - \bar{x}_2) - \mu_0}{\sigma \bar{x}_1 - \bar{x}_2}$$

$$= \frac{(25.2 - 22.6) - 0}{S\bar{x}_1 - \bar{x}_2} =$$

$$S\,\bar{x}_1 - \bar{x}_2 = \sqrt{\frac{S_1\bar{x}^2}{n_1} + \frac{S_2^2}{n_2}} = \sqrt{\frac{(5.2)^2}{40} + \frac{(6)^2}{50}} \approx 1.2$$

$$= \frac{2.4}{1.2} = 2$$

⑤ Conclusion : $Z > Z_\alpha$ reject H_0

$\mu_1 - \mu_2 > 0$

$\mu_1 > \mu_2$

Find the p-value.

p-value $\equiv P_r(Z > 2)$

$\equiv \frac{1}{2} - 0.4772$

$= 0.0228$

Reject H_0 if p-value $< \alpha$

$0.0228 < 0.05 \rightarrow$ reject H_0

Sample Size :- Testing Hypothesis.

$\alpha \equiv pr\,(\text{type I error})$

$B \equiv pr\,(\text{type II error})$

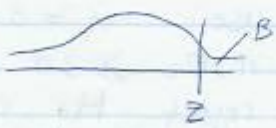
$1-B \equiv$ power of the test

$$n \equiv \frac{(Z_\alpha + Z_B)^2 \sigma^2}{(\mu_0 - \mu_a)^2}$$

$n =$ Sample Size

$\sigma =$ Standard deviation For the population.

$Z_\alpha =$ 

, $Z_B =$ 

$\mu_0 =$ the given value For μ in H_0

$\mu_a =$ the value For the population used. For type-II error.

$H_0 : \mu \geq 10$

$H_a : \mu < 10$

Ex :- Consider the Following hypothesis.

$\frac{64}{375}$) $H_0 :- \mu \geq 10$

$H_a :- \mu < 10$

$n = 120$, $\sigma = 0.05$

$\sigma = 5$, $\mu_a = 9$

$B = 0.2912 \rightarrow 0.1$

حتى نستطيع

تقليل هذا الرقم يجب العمل على زيادة حجم العينة .

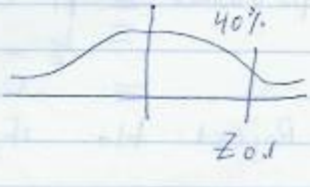
$\sigma = 5$

$Z_\alpha = 1.645$

$Z_B = 1.28$

$\mu_a = 9$

$\mu_0 = 10$



$$n = \frac{(1.645 + 1.28)^2 (5)^2}{(1)^2} =$$

$$n = 214$$

$$n = \frac{(Z_\alpha + Z_B)^2 \sigma^2}{(\mu_0 - \mu_a)^2} =$$

$$n = R\ (Z_\alpha + Z_B)^2 = \quad \sqrt{\frac{n}{R}} = Z_\alpha + Z_B$$

$$Z_B = \sqrt{\frac{n}{R}} - Z_\alpha \quad = \quad \alpha \text{ is given.}$$