date: 16/3/2022

Statistics : Methods used to:

① Collect data ⇒ Raw data

② Organize data

③ analyze data

④ Interpret data (تفسير معنى البيانات) why : لمساعدة صانع القرار

To get the iformations we need:

- Elements ⇒ Population

لما تجمع العناصر كلها تسمى population (set of all elements)

Census : التحقق من جميع العناصر (تعداد شامل) :- (حصر):

↳ لذلك يتم أخذ عينة (جزء من ال population):

عينة (Sample): subset of the population

Variables : طريقة أخذ الإحصائيات تتم عن طريق الأسئلة و الإجابات هي:

↳ e.g : Name / gender / number / Age / ID ...etc

Types of statistics:

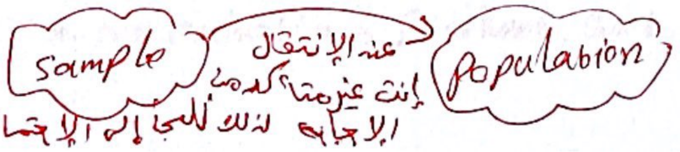
① Descriptive statistics : Methods used to summarize (organize data)

- tabular methods
- graphical methods
- Numerical mesures ; e.g: Average

⇒ طرق تلخيص ال Data

② Inferential Statistics (الإستنتاج)

: How to make conclusions about population from sample results.



probability لذلك نلجأ إلى الإحتمالية

→ Types of Variables or Data: :- (أنواع الأسئلة والبيانات المجمعة):-

① Qualitative variables: Can't assume a numerical value, it can be classified into two or more categories (Classes / Lables).

↳ Data collected: called qualitative data

e.g: gender, hair coller, marital status. الرقم الجامعي: لأنه بدل عن معلومات شخص

② Quantitative Variables: Can be measured numerically.

↳ Data collected: called quantitative data

e.g: incomes, number of children in a family, prices of homes

discrete (How many)

: Assume certin values and there usually gaps.

e.g: number of children in your family.

~~عائلة [illegible] أطفال [illegible]:~~
~~أطفال (2) الإجابة [illegible] (رقم)~~
~~[illegible] عدد الأطفال~~

:- (عدد الطلاب الي بيحضروا المحاضرة) :-
تختلف من محاضرة إلى أخرى.

Continuous (How much)

: Assume numerical value, within specific interval.

e.g: creatinine in patient's might be 1.2345615 but we can measure only 1.234 / Body mass / height / blood pressure

Parameters: كل النواتج ال طالعة من ال population

~~sample~~ statistics : كل النواتج ال طالعة من ال sample

Sources of Data:- :- (مصادر البيانات) :-

① **Primary Data:** :- (بيانات أساسية) :-

: Data must be collected.

← بيانات يقوم بإنتاجها الباحث ذات نفسه أو الـ survey (البحث) أو التجارب اللازمة لحل المشكلة.

② **Secondary Data:** :- (بيانات ثانوية) :-

: Data wich is already available on a firm or organization

↳ e.g: ← بيانات إحصائية عن معدل البطالة في عامة واشتغلت.

Methods of collecting data-: (طرق جمع البيانات) :-

1) **Cross-sectional data.**

: Data from several variables at same point of time.

: عدة متغيرات في وقت محدد.

2) **Time series data.** (أنت تقوم بتحديد الفترة)

: Data collected from a variable at several periods of time.

: متغير لعدة قراءات على فترات زمنية.

Scales of Measurements :- (موازين القياسات) :-

① **Nominal** ↗ Qualitative data **Data:** The weakest level of data, applies to data that are divided into different categories.

↳ e.g: gender / college and marital status.

② **Ordinal** ↗ Qual data **Data (Rank):** The same characterectics of nominal data and the order or Rank is meaningful.

↳ e.g: Grades ⇒ Excellent, very Good, Good, fail

production evaluation ⇒ Excellent, Good, fair and poor

Quantitative data

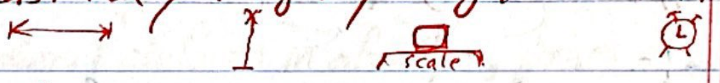
③ **Interval Data**: The same characterestic of ordinal data and the intervals between values is expressed in terms of a fixed unit of measure, Interval scale: always numiric

← بمعنى أو يعني تكون المسافة بين البيانات ثابتة بمقدار معين.

↳ e.g: Temperatures, Daily Expindibure.

Quantitative data

④ **Ratio Data**: The heighest level of measurement, it has all the properities of interval data and the Ratio between two values is meaningful. Variables: distance / heighb / weighb and Time



✻ يتطلب هذا القياس قيمة صفرية (طبيعية) للإشارة إلى عدم وجود أي شيء للمتغير عند نقطة الصفر.

✻ إذا كان الصفر لا يعني شيئاً (Nothing) : يكون Ratio

✻ إذا كان الصفر إصطلاحي يكون interval مثل درجة تجمد الماء (الصفر) لكن بالـ Fehrin hight تكون (32) ← وهذا هو صفر اصطلاحي (إبتدائي أنا قمت بوضعه)

✻ (الزيرو تبعو أنا: أنا إلي حددتو) : يكون Interval

✻ الـ Ratio scale : لا يمكن أن ينزل تحت الصفر مثل الإرتفاع، الوزن، والوقت تقاس من صفر إلى أعلى ولكن لا يمكن أن تنزل عن الصفر.

Ch.2 :-(Discriptive Statistics):-

_ How to organize Qualitative Data :-

① Tables : (الجداول):-

↳ e.g:

Class / Name / Categores	Frequency
Male	30
Female	15

s.t Frequancy = # of elements

Exp: A sample of BZU students, what drink do you prefer?

※ الـ Target هو: BZU students ، من الواقع إني عامل survey السبب: لأنه حكى بأنه أخذ sample.

Drink	Frequency	Relative Frequency (R.F)	%
Tea	200	$\frac{200}{1000} = 0.2$	20%
Coffe	250	$\frac{250}{1000} = 0.25$	25%
Juice	150	$\frac{150}{1000} = 0.15$	15%
soft Drink	300	$\frac{300}{1000} = 0.3$	30%
Energy Drink	100	$\frac{100}{1000} = 0.1$	10%
total =	$\boxed{1000}$	1	100%

_ 5 Categores

_ Sample size = # of elements / # of observation

_ Relative Freguency (R.F) → proportion ⇓ probability

$= \frac{\text{Class.Freq}}{\text{total.Freq}} = \frac{(C.F)}{(T.F)}$

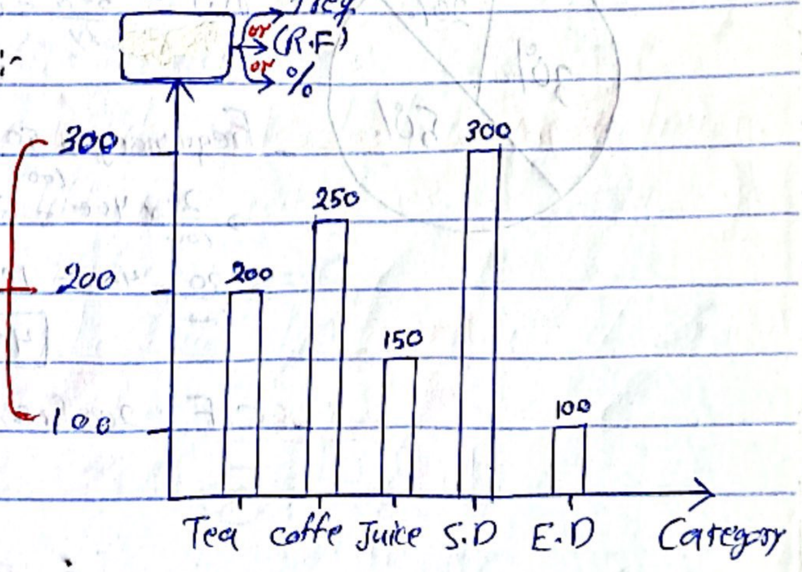
_ % ⇒ (R.F) × 100%

② Graphs : -:(الرسوم البيانية):-

_ Bar graphs , Pie chart

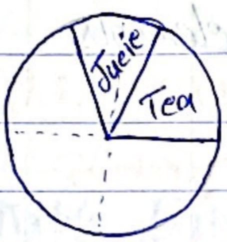
الترقيم حسب الـ frequency (زي ما بدك)

_ Sample size = 1000



Pie chart:-

✱ الحالة إلي بتطلع نسبة كل Caregore من الدائرة :



$$\text{Total. Frequency} \longrightarrow 360^\circ$$
$$\text{class. Frequency} \longrightarrow x^\circ$$

:

$$x^\circ = \left(\frac{C.F}{T.F}\right) \times 360^\circ$$

$$\boxed{x^\circ = (R.F) * 360^\circ}$$

Tea : $\frac{200}{1000} \times 360^\circ$

$= 72^\circ$

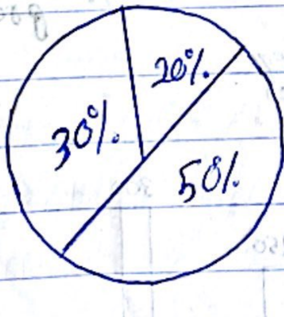
Juice : $\frac{150}{1000} \times 360^\circ$

$= 54^\circ$

Note:

✱ حاول إرسم مربع نقط في الدائرة عشان تعرف تبلش رسم :

Exp: Pie chart for 400 students.



: _ 3 categores

_ $R.F = \frac{50\%}{100\%} = 0.5, 0.2, 0.3$

_ Frequency $\Rightarrow \frac{50}{100} \times 400 = 200$

$\Rightarrow \frac{20}{100} \times 400 = 80$

$\Rightarrow \frac{30}{100} \times 400 = 120$

$\boxed{400}$ ⇒ Total frequency (Sample size)

_ C.F , 200 for 50% , 80 for 20% , 120 for 30%

_ How to organize Quantitative Data :-

① Tables:-

↳ e.g: your daily expenditure:

30, 50, 68, 100, 70, 25 ....

⊛ في حالة الأرقام يتم عمل فترات : classes/Categories Interval / Range

- هذا النوع من البيانات يوجد ثلاثة أشياء مهمة:

① Classes limits ② class width (w) ③ # of classes

$\underbrace{\text{Classes limits}}_{\text{lower (L)} \quad \text{Upper (U)}}$

⊛ حاول اختيار الطول المناسب للفترة والعدد المناسب.

⊛ نفضل ان يكون ال class width متساوي.

⊛ لا تجعل الفترات تتقاطع. (disgeint).

⊛ لازم تهتم بالبداية والنهاية بحيث البداية تكون أقل أو تساوي أصغر قيمة والنهاية تكون أكبر أو تساوي أعلى قيمة.

_ How to calculate the class width:- (The approximated one)

↳ Given the number of classes:

$$\text{width}(w) = \frac{\text{Maximum data value} - \text{Minimum data value}}{\#\text{ of classes}} = \frac{\text{Range}}{n}$$

↲ إذا كانت القسمة جواب عشري بعمل: (Round up to the next integer)

Exp: Min = 20, Max = 95, # of classes = 6

$$\Rightarrow w = \frac{95-20}{6} = \frac{75}{6} = \underset{\text{R.up}}{12.5} = \boxed{13}$$

Min + (w−1)

20 → $\boxed{32}$	6
33 − $\boxed{45}$	15
46 − 58	18
59 − 71	20
72 − 84	24
85 − 97	17

فرقهم: 13 وهو ثابت لآخره

13

⊛ إذا أردت الحصول على ال width من جدول:

W = Upper − Lower + 1

32 − 20 + 1 = 13

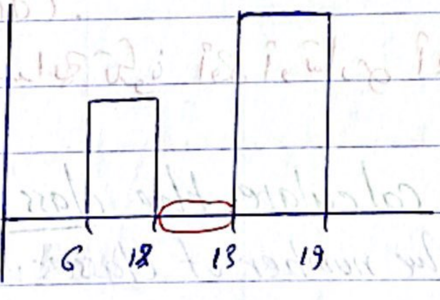
Exp: Given the following frequency distribution :-

Class	Frequency	R.F	true limit	Midpoint	Cumulative freq.	Cumul.freq. R.F
6 — 12	4	≈ 13%	5.5 — 12.5	$\frac{12+6}{2}=9$	4	4/30
13 — 19	5	≈ 17%	12.5 — 19.5	$\frac{13+19}{2}=16$	4+5 = 9	9/30
20 — 26	6	20%	19.5 — 26.5	23	4+5+6 = 15	15/30
27 — 33	10	≈ 33%	26.5 — 33.5	30	4+5+6+10 = 25	25/30
34 — 40	5	≈ 17%	33.5 — 40.5	37	30	30/30
Total =	30	100%				

↳ width : $U - L + 1 = 12 - 6 + 1 = 7$

↳ # of elements is : 30

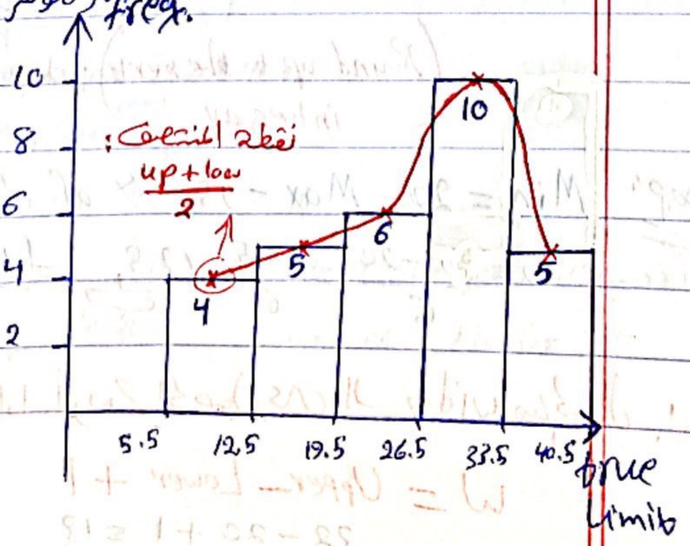
- How "true limit" comes :-



❀ الفراغ المحدد بالدائرة الحمراء ⬭ بين 13 و12 ليس من الضروري وجوده لذلك نحاول توسيع الفراغ لكي يتصل الجانبان مع بعضها البعض؛ لذلك يتم إضافة عمود يدعى "true limit" : يزيد عن up ويقلل من low

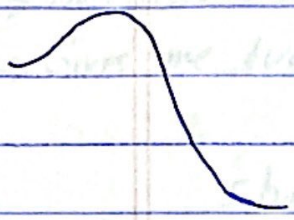
① Histograms :- : (مؤشر لشكل التوزيع) :

↳ Indicator : Shape of the distribution

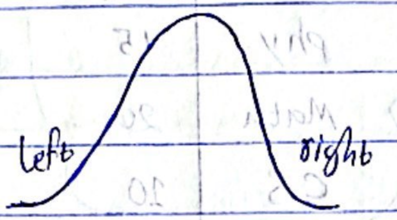


❀ تم التوصيل بين النقط لبناء الشكل لأن الخط عبارة عن منحنى.

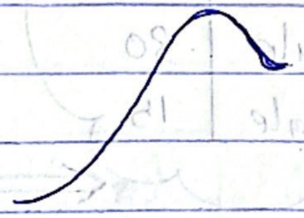
↳ Skewness :- :(مكان تواجد البيانات):



: Right skewness



: Symetric



: left skewness

(*) How many frequencies less then or equal to the upper limit : :- Cumulative_frequency

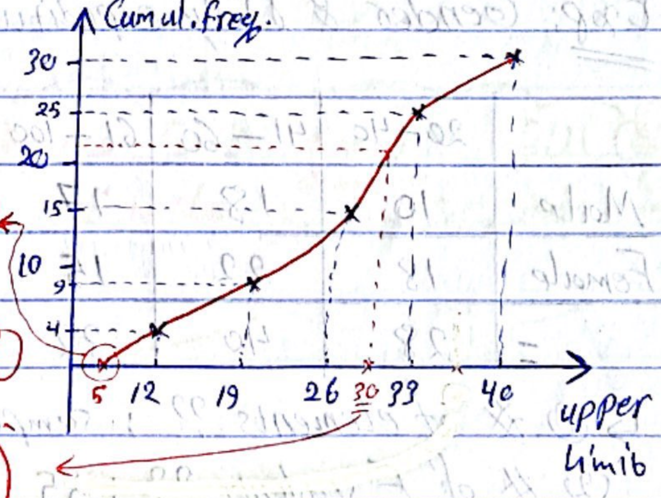
(*) ال Qulib.data ⇐ ال bar graph : متباعدة
ال Qunbi. data ⇐ ال bar graph : غير متباعدة

② Ogive :- :(Cumulative.freq.): :(التكرار التراكمي):

↳ To find a relationship between: Cumulative.freq and upper limit

⇐ ال رسمة علاقة تحدد موقع N (location) 5
أصناف من خلال ال up.limit & Cumul.freq

⇐ (*) هنا المنحنى يجب أن يكون مغلق ; يعني فئة أصغر،



(*) باستطاعتي حساب ال data لفترة أقل من
ال up.limit مثلا : نسبة الطلاب بالمعروضين؟
أقل من 30 ?? تقريباً $\frac{25}{30}$ $\frac{(\text{Cumulative freq})}{\text{n f.}}$

(2.3) is deleted

ch.(2.4) :- (Crosstabulations & Scatter diagrams) :-

↳ ① Crosstabulations :-

Gender	freq.
Male	30
Female	15

Major	freq.
phy	15
Math	20
C.S	10

Major Gender	Phy	Math	C.S	
Male	5	10	8	=23
Female	10	10	2	=22
				45

(*) الأرقام هنا عشوائية ☺

↳ هنا بشكل عام طريقة عمل ال:
Crosstabulations بين جدولين لمتغير
واحد وإيجاد علاقة بينهما ←

Exp: Gender & Daily expediture:

	20 – 40	41 – 60	61 – 100	Total =
Male	10	18	17	45
Female	18	22	15	55
=	28	40	32	= $\boxed{100}$

↳ ① # of elements ?? ; sample size = 100

② # of F students ?? : 55

③ % of M students ?? : 45%

④ % of Students with daily expenditure less than 61 ⇒ : $\frac{40+28}{100} \times 100\% = 68\%$

⑤ of male students, what is the percentage of students with expenditure less than or equal to 60 : ?? : $\frac{18+10}{45} = 0.62 \Rightarrow 62\%$

② Scatter diagram:-

↳ Indecator relationship between 2 variables

↳ Gives me two things :

Shape:

الشكل العام للنقاط المرسومة

يكون خط (Trend line).

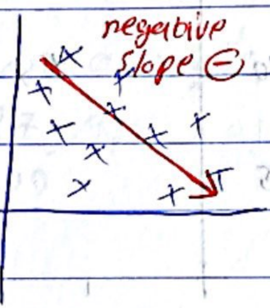
sign (الإشارة) (Relation) :-

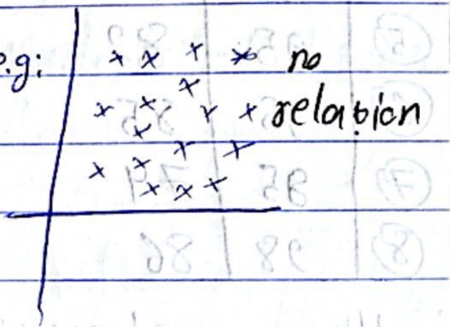
or ⊕ : طردي

or ⊖ : عكسي

No relation

e.g: 

e.g: 

e.g: 

(✱) بحاول أرسم خط من اليسار لليمين ويكون ماشي مع توجه النقاط

(✱) الفرق الـ Scatter diagrams انه بكون عنا 2 variables بكونو Quntitative والمتغير بكون independent والآخر بكون dependent :- ⇐

$x \equiv$ independent

$y \equiv$ dependent

Exp: $x \equiv$ Income

$y \equiv$ Expenditure

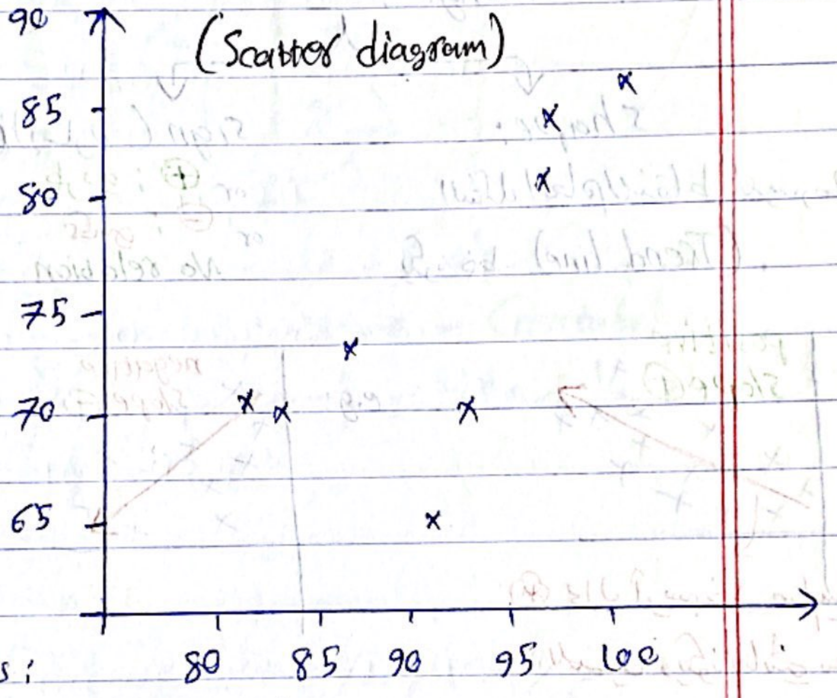
sample $(x_1, y_1), (x_2, y_2), (x_3, y_3) \ldots\ldots\ldots (x_n, y_n)$

Exp:

Tawgihi $\equiv x$, Average of univercity $\equiv y$

	x	y
①	83	72
②	92	67
③	82	70
④	93	70
⑤	95	82
⑥	95	85
⑦	95	74
⑧	98	86

$\Rightarrow$



- How many observations:
: 8

⊛ the relation is : positive

Ch. 3: (Descriptive Statistics):

- Numerical Measures:
types: ① measures of center (Location) (Central tendency)

② Measures of variation (Distripution)

Ch. (3.1)

(3.1) Measures of location:

① The mean (Average) ($\bar{x}$)

↳ $\bar{x} \equiv$ sample mean $\Rightarrow$ statistics

↳ μ (mu) $\equiv$ population mean $\Rightarrow$ paremeter

⇒ Raw data :-

اذا كان عينة sample : $x_1, x_2, \ldots, x_n$

↳ $\bar{x} = \dfrac{x_1 + x_2 + \ldots x_n}{n} = \dfrac{\sum_{i=1}^{n} x_i}{n}$, s.t $n \equiv$ sample size

↳ $\mu = \dfrac{\sum_{i=1}^{N} x_i}{N}$, s.t $N \equiv$ population Size

Exp: the following scores are a sample of Math 1411-1211 scores.

98, 60, 86, 87, 75, 87, 90.

Find the sample mean ??

$$\Rightarrow \quad \bar{x} = \frac{\sum x}{n} = \frac{583}{7} = 83.29$$

② The Median (الوسيط)

: The value in the middle of arranged data values

↳ Steps :

① Arrange the data in an increasing order ↑

② Find the median position :

case 1 : odd number of elements (n)

$$\frac{n+1}{2}$$

case 2 : even number of elements (m)

positions $\left(\frac{m}{2}, \frac{m}{2}+1\right)$

median ≡ Average of the two 2 values (positions) $\left(\frac{m}{2}, \frac{m}{2}+1\right)$

Exp: Find the medin in the past example :-

$\Rightarrow$ # of values is 7 : odd

① $\Rightarrow$ 60, 75, 86, (87), 90, 98, 100

median $\Rightarrow \frac{7+1}{2} = 4$ $\therefore$ 87 is the median

② 25, 30, 16, 9, 10, 8, 11, 24

$\Rightarrow$ # of values : 8

$\Rightarrow$ 8, 9, 10, (11), (16), 24, 25, 30

↳ positions : (4, 5)

↳ median $= \frac{11+16}{2} = 13.5$

---

③ The Mode (المنوال)

: The value that occures most frequently

↳ e.g: 98, 60, 86, $\underline{\underline{87}}$, 75, $\underline{\underline{87}}$, 90, $\underline{\underline{87}}$, 60

mode = 87 (Unimodal : منوال واحد فقط)

↳ e.g: $\underline{\underline{98}}$, (60), $\underline{\underline{98}}$, (60), 70, 65, 30

mode = 98, 60 (bimodal : أكثر من واحد)

↳ e.g: 90, 85, 76, 80, 99

mode = No mode.

⊛ في ال Qualitative data ال Mode أفضل

⊛ ال Mean, ال Median يستخدموا في ال Quantitative data لإيجاد التشتت

Exp: sample of 5 students daily expenditure :

30, 50, (40), 20, 150

↳ mean $= \frac{290}{5} = 58$

↳ median $= \frac{5+1}{2} = 3$

$\therefore$ median is $\underline{\underline{40}}$

The percentiles :-

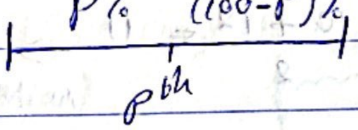
↳ related to position.

; قيمة تحدد موقع (نسبة ل data) بالنسبة لعلامات أخرى ;

_ the p^{th} percentile:

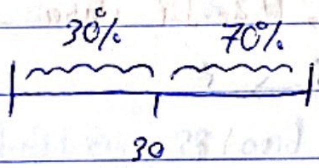
: A value, s.t $p\%$ of this data value $\leq$ the value, and $(100-P)$ of the data values $\geq$ the value.

بقية ال data

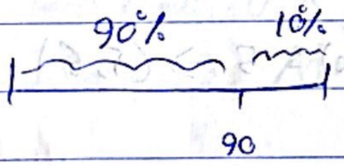


↳ e.g:

$30^{th} \Rightarrow P_{30}$

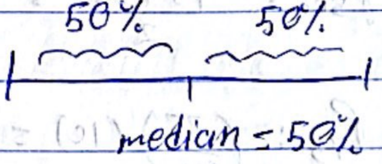


e.g: P_{90}



⇒ ال percentile إلي ترتيبه 90
هو من أحسن (أفضل) 10%.

e.g: the median rank is 50% ⇒



_ Steps to find the p^{th} percentile:-

[1] Arrange the data. (Increasing)

[2] Find the index i

$$i = \left(\frac{P}{100}\right) \times (n), \text{ s.t } n = \# \text{ of items.}$$

[3] if i is not integer; round up to the next integer

↳ e.g: $i = 4.15 \Rightarrow$ position $= 5$ ⇒ (القيمة في الموقع 5) هنا الموقع بعطيني ال p تاعي

[4] if i integer ⇒ positions : $(i, i+1)$ so percentile = Average of the 2 values in positions $(i, i+1)$.

Exp: Given the following data:-

20, 30, 12, 10, 1, 36, 25, 18, 19, 43

Find : P_{20} , P_{25} , P_{50} , P_{75} ??

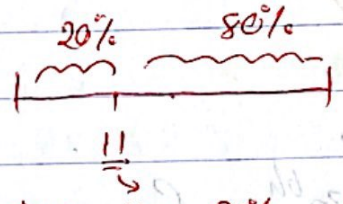
$\Rightarrow$

$\overset{1}{1}, \overset{2}{⑩}, \overset{3}{⑫}, \overset{4}{18}, \overset{5}{19}, \overset{6}{20}, \overset{7}{25}, \overset{8}{30}, \overset{9}{36}, \overset{10}{43}$

# of items = 10

P_{20} : $\left(\frac{2\cancel{0}}{10\cancel{0}}\right)(1\cancel{0}) = 2 \rightarrow$ positions (2, 3)

$\Rightarrow P_{20} = \frac{10+12}{2} = 11 \rightarrow$



تفسير: 20% من ال data أقل من 11, 80% من ال data أكبر من 11

$\Rightarrow P_{25} : \left(\frac{25}{100}\right)(10) = 2.5 \rightarrow 3$

$P_{25} = 12$

$\Rightarrow P_{50}$:(median) : $(0.5)(10) = 5 \rightarrow (5, 6)$

$P_{50} = \frac{19+20}{2} = 19.5$

$\Rightarrow P_{75} : \left(\frac{75}{100}\right)(10) = 7.5 \rightarrow 8$

$P_{75} = 30$

Special percentiles:-

① Quartiles :- (ارباع)

- the first quartile ≡ the lower quartile : $Q_1 = P_{25}$
- the Second quartile ≡ the middle quartile ≡ median : $Q_2 = P_{50}$
- the third quartile ≡ the upper quartile : $Q_3 = P_{75}$

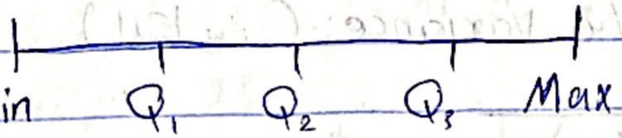
: in the ex exp the $Q_1 = P_{25} = \underline{\underline{12}}$

$Q_2 = P_{50} = \underline{\underline{19}}.5$

$Q_3 = P_{75} = \underline{\underline{30}}$

Scanned with CamScanner

→ The five number summary



↳ e.g: write the five number summary for the ex exp:

$min = 1$, $Q_1 = 12$, $Q_2 = 19.5$, $Q_3 = 30$, $Max = 43$

1, 12, 19.5, 30, 43 ⇐ تكتب هكذا

Exp: Given the five number summary for a data by:

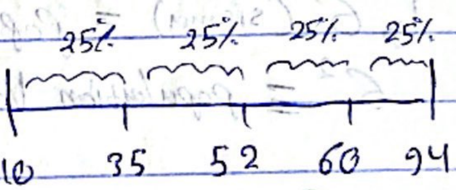
10, 35, 52, 60, 94

⇒

$min = 10$ $Q_2 = 52$ (median)

$Q_1 = 35$ $Q_3 = 60$ $Max = 94$

⇒



⇒

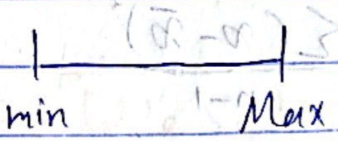
the % of data between 35 and 60 is: 50%

the % of data less than 60 is: 75%

:- Ch.(3.2) Measures of Variation (Dispersion)

① The Range:

$R = Max - min$



↳ e.g: 1, 10, ----, 43

$R = 43 - 1 = 42$

(*) ممكن تضللنا ال Max وال min لما يكونو الثنين Extremes

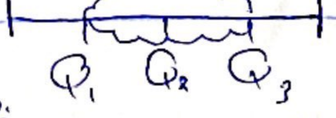
↳ e.g: 5, 40, ----, 70, 95

$R = 95 - 5 = 90$

:- لذلك بنلجأ للنوع الثاني :←

② The Inter Quantile Range :-

$IQR = Q_3 - Q_1$ ⇒



ex exp:

$Q_3 - Q_1 = 30 - 12 = 18$

③ The Variance :- (التباين)

↳ $\sqrt{\text{Variance}}$ = The standard deviation (الإنحراف المعياري)
= S.D

$(S.D)^2$ = Variance

↳ $S \equiv$ Sample S.D
$S^2 \equiv$ Sample Variance] → statistics مثال عن ال

↳ σ (sigma) $\equiv$ Population S.D
$\sigma^2 \equiv$ population Variance] → parameter مثال عن ال

_ Formulas for solving S and σ :-

↳ $S = \sqrt{\dfrac{\sum(x-\bar{x})^2}{n-1}}$, s.t $\bar{x}$ = sample mean
n = sample size
$x - \bar{x}$ = the "deviation" of the value x from the mean : $\bar{x}$

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

↳ $\sigma = \sqrt{\dfrac{\sum(x-\mu)^2}{N}}$, s.t N = population size
μ = population mean

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

Exp: Find the sample standard deviation (S.D) for the following data set: 12, 18, 17, 13, 25

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

$$\rightarrow \bar{x} = 17$$

x	$x-\bar{x}$	$(x-\bar{x})^2$
12	-5	25
18	1	1
17	0	0
13	-4	16
25	8	64
	$\boxed{0}$	$\boxed{106}$

$$\Rightarrow S = \sqrt{\frac{106}{5-1}} = \sqrt{26.5} = 5.15$$

$$\therefore \text{Variance} = S^2 = 26.5$$

✱ يزداد الـ Variation بازدياد قيمة الـ S.D.

✱ يكون جواب الـ: Range , IQR and S.D صفر إذا كانت القيم جميعها متساوية.

✱ يستعمل الـ S.D غالباً في قياس الأسهم بسبب وجود خطر النزول بسرعة والصعود بسرعة.

Exp: Exam 1 Exam 2 ; Which has more Variation

20 15

⇒ Using : Coefficent of Variation (تستخدم للمقارنة بين (2) two data set)

$$\Rightarrow \boxed{C.V = \frac{S.D}{Mean}(100\%)}$$

$$\Rightarrow \boxed{C.V = \frac{\sigma}{Mean}(100\%)} \;,\; \frac{S}{\bar{x}}(100\%)$$

(*) من الممكن أن يكون الانحراف المعياري (S.D) أكبر من الـ Mean عند القيم المتطرفة (Extreme Values) $EV^{(s)}$

{ S.D > M
2, 3, 5, 8, ..., 250 }

Exp: Consider the following sample data set : Find :

12, 15, 8, 7, 13, 19, 24

① Mean ② Median ③ IQR ④ S.D ⑤ C.V ⑥ σ

↳ 1) $\bar{x} = 14$ ↳ 2) $M = 13 \approx 14$

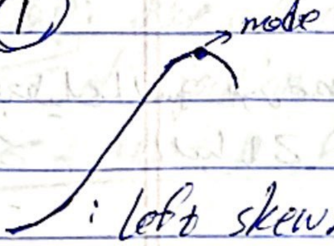
↳ 3) $IQR = Q_3 - Q_1 = 19 - 8 = 8$ ↳ 4) $\underset{(\sigma_s)}{S.D} = 6$

↳ 5) $C.V = \frac{S}{Mean} \times (100)\% = \frac{6}{14}(100)\% \approx 43\%$

حطو بالسؤال
Sample data set 1.5

↳ 6) $\underset{(\sigma_m)}{\sigma} = 5.55$

①

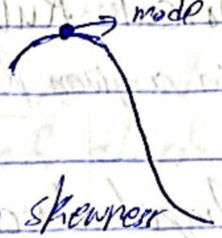


: left skewness

⊖ negative skewness

mean < median

②

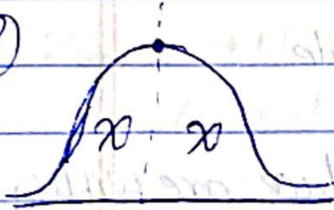


: Right skewness

⊕ Positive skewness

mean > median

③



mode = median = mean

Zero skewness.

Ch.(3.3)

- The: (Z_Score):

↳ the Z score for a value x from a data set with $\underset{(\bar{x})}{\mu}$ and S.D (σs) is:

for population ↶ for sample ↷

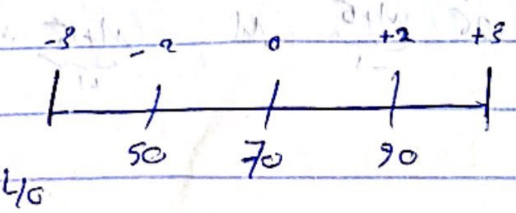
$$Z = \left(\frac{x - \overset{\text{mean}}{\mu}}{\underset{\text{standard deviation}}{\sigma}}\right), \left(\frac{x - \bar{x}}{s}\right)$$

Exp: Star 2311 ~~this~~ exam has: $\mu = 70$, $\sigma = 10$, Find the Z score for the following scores: 90, 50, 60, 78, 40

⇒

$$Z = \frac{x - \mu}{\sigma} \Rightarrow Z_{90} = \frac{90 - 70}{10} = +2$$

$Z_{50} = -2$, $Z_{70} = 0$



# of standard deviations for a value from the mean

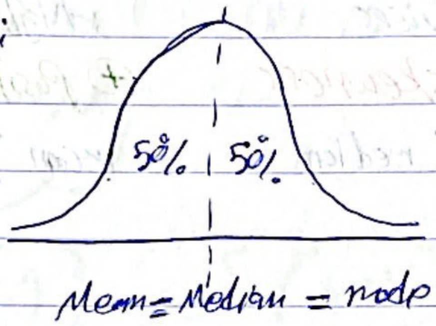
$90 \equiv 70 + 2(S.D)$ ⇐ $40 = 70 - 3(S.D)$

The Empirical Rule

▷ % whith in a given Range of data ; (بتطلعلك النسبة [illegible] البيانات) :

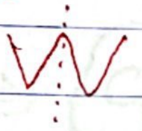
تنطبق

↳ Bell_shaped data ;



Mean = Median = mode

Bell shaped ≠ symetric عكس

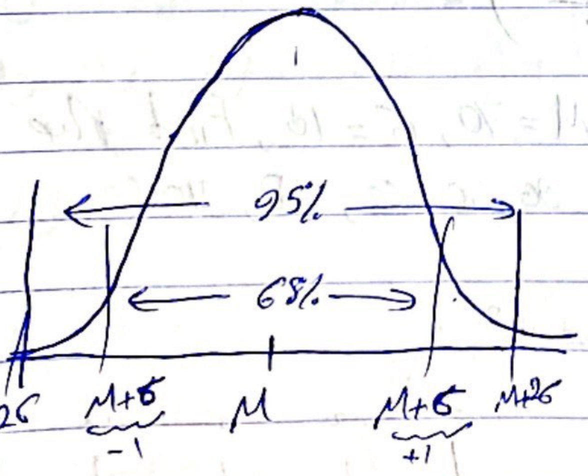


↳ $\boxed{1}$ Approximatly 68% of the data Values are within 1 S.D of the mean.

$\boxed{2}$ " " " 95% " " " " " 2 S.D of the mean

(100%) = (99.97%) Realone

$\boxed{3}$ Almost All data values are with in 3 S.D of the mean



$$Z = \frac{x-\mu}{\sigma}$$

$$\hookrightarrow \frac{\mu+\sigma-\mu}{\sigma} = 1$$

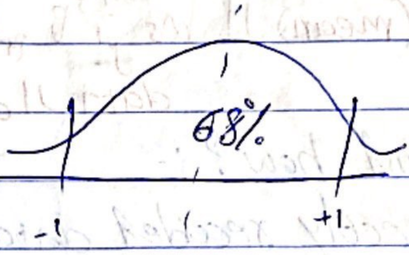
$$\hookrightarrow \frac{\mu-\sigma+\mu}{\sigma} = -1$$

Exp: The salaries of BZU employees have a bell shaped distripution with Mean = M = 1600 JOD and standard deviation σ = 400 JOD, Find the percenteages of salaries:

① Between 1200 and 2000

⇒ $Z_{1200} = \frac{1200-1600}{400} = +1$

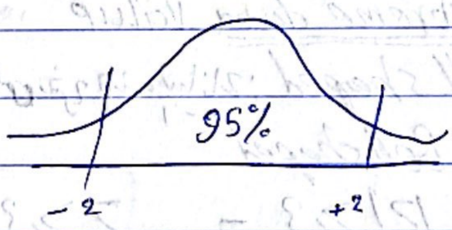
$Z_{2000} = +1$ ⇒



② Between 800 and 2400

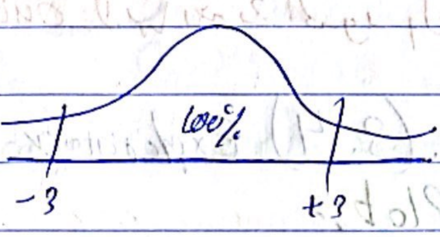
⇒ $Z_{800} = -2$, $Z_{2400} = +2$

⇒



③ Between 400 and 2800

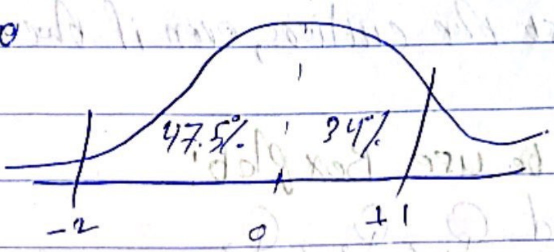
⇒



④ Between 800 and 2000

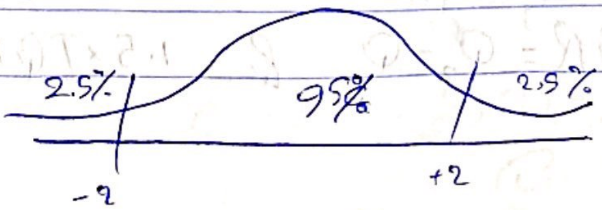
$Z_{800} = -2$, $Z_{2000} = +1$

⇒



⑤ less than 2400

⇒



97.5%

_ Oubliers:-

_ Exbreme Values:-

↳ Very small data value (بيانات قيمتها قليلة جداً)

↳ Very Large data value (بيانات قيمتها عالية جداً)

✱ يهتم بال Outliers لأنه لهم تأثير على ال (mean) (M) وال S
وتعد أرقام غير طبيعية تأثر على ال data.

↳ why and how?:-

① Incorecbly recorded data value → (سماعة كلمة مشتبهة)

② Incorecbly Included data value → (جمعة كلمة من عينة خاطئة
مع كلاب 2 من عنزة)

③ Usual exbreme data value

↳ Bell shaped: بيانات على شكل جرس ومنتظمة

: If Bellshaped

$|z| \geqslant 3 = Z \geqslant 3$ or $Z \leqslant -3$

✱ هذه القيم تعد outliers لأنها تعدى الحدود الطبيعية لل z-score
(تسمى Extreme Values)

:- ch. (3.4) Exploratory Data Analysize

_ The Box Plot:-

↳ summarize data.

↳ Test the oubliers, even if there's no shape for the data.

_ Steps to use Box Plot:

[1] Find Q_1, Q_2, Q_3

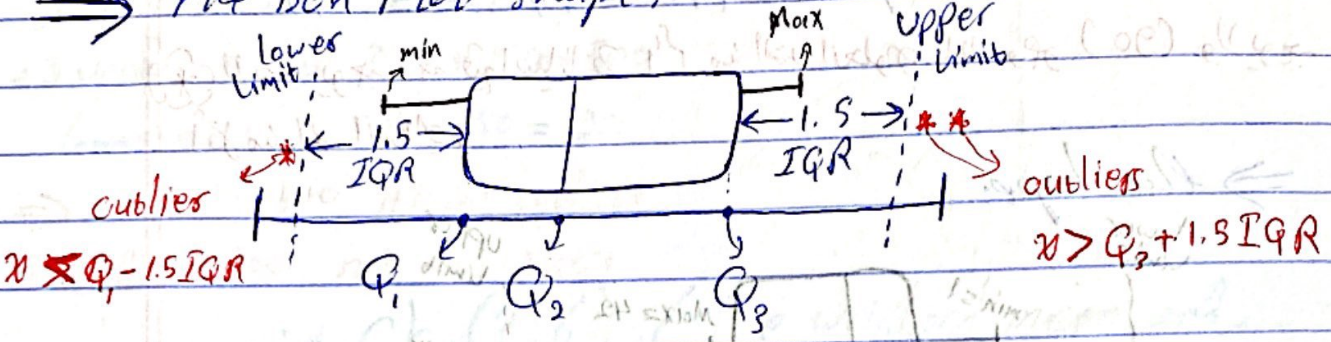
[2] Find $IQR = Q_3 - Q_1$ & $1.5 \times IQR = ??$

[3] Fence.

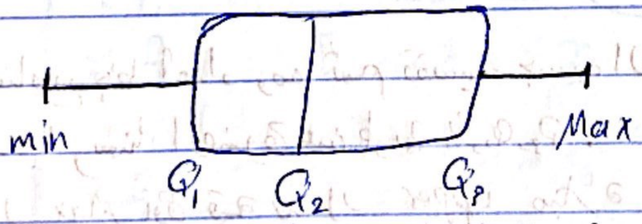
↳ Upper Limit : $Q_3 + 1.5\, IQR$

↳ Lower Limit : $Q_1 - 1.5\, IQR$

⇒ The Box Plot shape:



⇒ the needed shape:



↳ ⊢ ⊣ called (wisker) : شوارب القط

Exp: Construct a box plot for the following data: 16, 1, 20, 18, 30, 40, 42, 90

⇒ 1 2 3 4 5 6 7 8

1, 16, 18, 20, 30, 40, 42, 90

$Q_1 = P_{25} = \left(\frac{25}{100}\right)(8) = \underline{2} \rightarrow 2, 3$

$Q_1 = \frac{16+18}{2} = \boxed{17}$

$Q_2 = P_{50} = \left(\frac{50}{100}\right)(8) = \underline{4} \rightarrow 4, 5$

$Q_2 = \left(\frac{20+30}{2}\right) = \boxed{25}$

$Q_3 = P_{75} = \left(\frac{75}{100}\right)(8) = \underline{6} \rightarrow 6, 7$

$Q_3 = \frac{40+42}{2} = \boxed{41}$

$\Rightarrow IQR = Q_3 - Q_1 = 41 - 17 = 24$

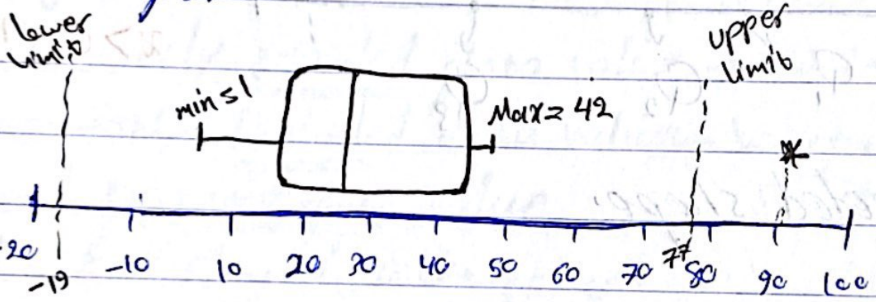
$\Rightarrow 1.5 \times IQR = 1.5\,(24) = 36$

$\Rightarrow$ Upper Limit $= Q_3 + 1.5\,IQR = 41 + 36 = 77$

$\Rightarrow$ lower Limit $= Q_1 - 1.5\,IQR = 17 - 36 = -19$

(*) لاحظ يوجد أحد قيم البيانات أكبر من الحد العلوي ألا وهو (90) ولا يوجد أقل من الحد السفلي.

$\Rightarrow$ the shape:



قمنا برسم خط أعداد، ومن ثم تقسيمه حسب القيم للبيانات الموجودة معي؛

بعدها رسمنا الصندوق عن طريق Q_1, Q_2, Q_3.

ال Max أقل قيمة بعد ال upper limit مباشرةً، وال min أعلى قيمة بعد ال lower limit مباشرةً.

ال * تعبر عن القيمة التي تعد outlier، وهو ال 90.

Exp: Check if there's an outliers or not: ?? , $Q_1 = 200$, $Q_2 = 280$

$Q_3 = 320$, Test $x = 80$, $x = 600$

$\Rightarrow$

$IQR = Q_3 - Q_1 = 320 - 200 = 120$

$\Rightarrow 1.5IQR = 180$

$\Rightarrow$ upper limit $= 320 + 180 = 500$

lower limit $= 200 - 180 = 20$

$\Rightarrow 600 > 500$ it's an outlier.

$20 \leq 80 \leq 500$ not an outlier.

Ch. (3.6): (The weighted mean and groubed data):

- Weighted mean:

Value: x	frequency	$x \cdot f$
x_1	f_1	$x_1 f_1$
x_2	f_2	$x_2 f_2$
$\vdots$	$\vdots$	$\vdots$
x_n	f_n	$x_n f_n$

$\Rightarrow$ $\bar{x}$: weighted mean $= \dfrac{\sum xf}{\sum f}$

$\bar{x}$: mean $= \dfrac{\sum x}{n}$

Exp:

x	f	$x \cdot f$
5	6	30
10	20	200
17	8	136
12	5	60
9	16	144
	$\boxed{55}$	$\boxed{570}$

$\Rightarrow$ weighted mean $= \bar{x} = \dfrac{\sum x.f}{\sum f}$

$\bar{x} = \dfrac{570}{55} = 10.37$

_ Groped data :-

↳ frequency dkstripution.

tbw:

class	freq.	midpoint (m)
10 - 18	10	⑨

الأرقام :
غير دقيقة (بالمضبوط)

$$\Rightarrow \bar{x} = \frac{\sum m.f}{\sum f}$$

s.t: $m \equiv$ (class mark)

:- Ch. (3.5 & 12.2) :(Measures of Association between 2 variables):-

↳ Z : numirical variable

↳ x : independent variable

↳ y : dependent variable

sample ⇒ observations $(x_1, y_1), (x_2, y_2)$ —

_ The Covariance (التباين بين متغيرين)

↳ $S_{x,y} \equiv$ Sample Covariance (Statistics)

↳ $\sigma_{xy} \equiv$ Population Covariance (Parameter)

_ formulas :

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

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

$\Rightarrow S_{xy} > 0 \Rightarrow \oplus$ relationship

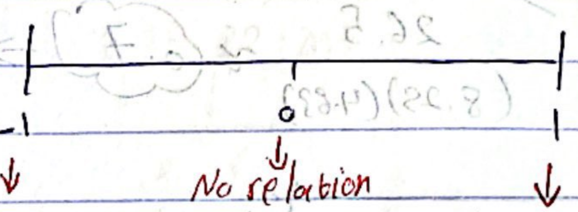
$< 0 \Rightarrow \ominus$ relationship

$= 0 \Rightarrow$ No relationship

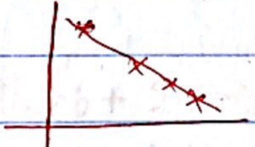
- The Corelation ≡ Strength of the relation
- The Corelation Coefficiant (معامل الارتباط):-
  - ↳ r_{xy} ≡ Sample C.C (Corelation Coefficiant)
  - ↳ ρ_{xy} ≡ Population C.C

$$\Rightarrow r_{xy} = \frac{S_{xy}]\rightarrow \text{sample covariance}}{\underbrace{S_x}_{\text{S.D for } x} \cdot \underbrace{S_y}_{\text{S.D for } y}}, \quad -1 \leq r_{xy} \leq 1$$

$$\Rightarrow S_x = \sqrt{\frac{\sum(x-\bar{x})^2}{n-1}}$$

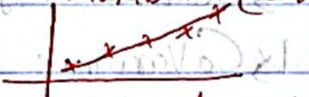


$r_{xy} = -1$; negative perfect (complete) relationship between Variables



$|r_{xy}| \to 0$ ⇒ $\boxed{\text{Weak}}$

$r_{xy} = 1$; perfect (complete) positive relationship between Variable (Strong)



$|r_{xy}| \to 1$ ⇒ $\boxed{\text{Strong}}$

Exp: (46)

$$\Rightarrow S_{xy} = \frac{\sum(x-\bar{x})(y-\bar{y})}{n-1}, \quad \bar{x} = \boxed{16}, \quad \bar{y} = \boxed{10}$$

	x	y	$x-\bar{x}$	$(x-\bar{x})^2$	$y-\bar{y}$	$(y-\bar{y})^2$	$(x-\bar{x})(y-\bar{y})$
	6	6	-10	100	-4	16	40
	11	9	-5	25	-1	1	5
	15	6	-1	1	-4	16	4
	21	17	5	25	7	49	35
	27	12	11	121	1	1	22
$\sum$	$\boxed{80}$	$\boxed{50}$	$\boxed{0}$	$\boxed{272}$	$\boxed{0}$	$\boxed{86}$	$\boxed{106}$

$$S_{xy} = \frac{\sum(x-\bar{x})(y-\bar{y})}{n-1} = \frac{106}{4} = 26.5$$

- the position R :

$$r_{xy} = \frac{S_{xy}}{S_x \cdot S_y} , \quad S_x = \sqrt{\frac{\sum(x-\bar{x})^2}{n-1}} = \sqrt{\frac{272}{4}} = 8.25$$

$$S_y = \sqrt{\frac{86}{4}} = 4.63$$

$$\Rightarrow \frac{26.5}{(8.25)(4.63)} \approx \boxed{0.7} \Rightarrow \text{Strong}$$

- Measures of Variables

x, y

$(x_1, y_1), ----, (x_n, y_n)$

↳ Scatter diagram

↳ Covariance S_{xy}

↳ Correlation r_{xy}

↳ Regression ; Equation of the trend Line

$$y = mx + b$$

$$\boxed{\hat{y} = \underbrace{b_0}_{y\text{-axis}} + b_1 x}$$

- found b_0 & b_1 using:

↳ Least squared method:

$$b_1 = \frac{\sum(x-\bar{x})(y-\bar{y})}{\sum(x-\bar{x})^2} = \frac{S_{xy}}{S_x^2}$$

$$b_0 = \bar{y} - b_1 \bar{x}$$

Exp: Consider the following table: Find: $\bar{x}$, S_x, C.V(x), CV(y), S_{xy}, r_{xy}, IQR, median.

(*) Reg mode used to find All of the above.

x	y
15	32
20	43
11	20
17	36
19	40
30	62
40	78
32	65
10	22
45	100

① mode 3 ② 1 Lin ③ Data entry : أزواج مرتبة

: 15, 32 M+

n = 6

④ Shift + 2

$1 \Rightarrow \bar{x} = 18.67$ } for (x)

$3 \Rightarrow S_x = 6.41$

⑤ Shift + 2 + →

$1 \Rightarrow \bar{y} = 38.84$ } for (y)

$3 \Rightarrow S_y = 13.89$

⑥ shift + 2 + → + →

$1 \Rightarrow A = b_0 = -1.4$

$2 \Rightarrow B = b_1 = 2.15$

$3 \Rightarrow r = r_{xy} = 0.99$

$$\Rightarrow S_{xy} = S_x^2 \cdot b_1 = (6.41)^2 \cdot 215 = 88.34$$

$$\hat{y} = b_0 + b_1 x$$

$$\hat{y} = -1.4 + 2.16 x$$

⇒ the Strength of the relation :-

$r_{xy} = 0.99$ (STRONG)

Ch. (4.1) :- (Introduction to probability) :-

Def: The probability is a numirical value that measures the chance (Likelihood) that an uncertin event will happen.

- Experiment: any processes that gives outcomes. they are two types:

Random	Non Random
: يتم إجراء التجربة لأكثر من مرة وتظهر نتائج مختلفة .	: يتم إجراء التجربة لأكثر من مرة وتظهر نفس النتيجة كل مرة .

- Sample Space (S) :- (الفضاء العيني) :-

: The set of all posible outcomes of a random experiment.

$$S = \{S_1, S_2, S_3, \dots, S_n\}, \quad S_i \equiv \text{Sample Point} \equiv \text{outcome}$$

Exp: ① Toss a Coin:

$$S = \{\text{Head}, \text{Tail}\} = \{H, T\}$$

$$\# S = 2$$

② Roll a die:

$$S = \{1, 2, 3, \dots, 6\}$$

$$\# S = 6$$

③ Play football game:

$$S = \{\text{Win}, \text{loss}, \text{tie}\}$$

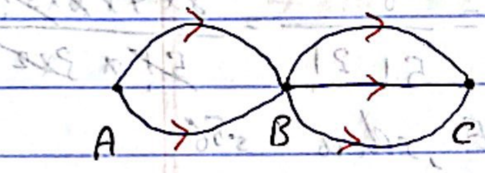
(★) All these ①+②+③ exp are One-Step-experiment

Multi_Step_experiment ≥ 2

$\# S$ = by Counting Rules:

1-) Multiplication Rule.
2-) Combination Rule.
3-) Permutation Rule.

$\boxed{1}$ Multiplication Rule:-



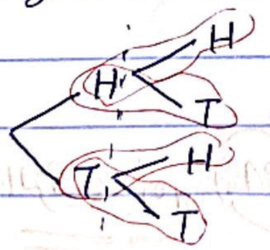
$2 \times 3 = 6$

Exp: ① Toss a Coin 2 times:

$\# S = (2)(2) = 4$ outcomes

↳ How to get the outcomes:

Tree:-



$\Rightarrow S = \{HH, HT, TH, TT\}$

② Toss a Coin 3 times

$(2)(2)(2) = 8$

③ Roll 2 dice

$(6)(6) = 36$ pairs

④ Roll a dice and Toss a Coin

$\# S = (2)(6) = 12$

select r objects from n objects s.t $r \leq n$. in two ways:

① Compination ; No order

② permutation ; order

① Compination:

The number of compinations of n objects take r at a time $= nC_r, \left(\frac{n}{r}\right)$ is given by

$$nC_r = \frac{n!}{(n-r)!\,r!} \Rightarrow 8C_3 = \frac{8!}{(8-3)!\,3!} = \frac{8!}{5!\,3!} = \frac{8\times 7\times \cancel{6}\times \cancel{5!}}{\cancel{5!}\times \cancel{3\times 2}}$$

$= 56$

Exp: List of 5 students A, B, C, D and E, select three (3) students randomly.

$\Rightarrow$ # of selections = # of compinations

$= 5C_3 = 10$

$\Rightarrow$ List All possible selections;

ABC BCD CDE

ABD BCE

ACD BED

ADE

ACE

ABE

BAC ≡ ABC ≡ ACB ≡ CBA الترتيب غير مهم

② Permutation :

The number of permutations of n objects taken r at a time : $_nP_r$ is given by $_nP_r = \frac{n!}{(n-r)!}$

$\Rightarrow {}_{10}P_7 = \frac{10!}{7!} = 604800$

Exp: How many ways can 4 students set on a line of 7 chairs

$\Rightarrow {}_7P_4 = 840$

⊛ الترتيب مهم لأنه الطالب اللي بيجي بالأول مش زي الطالب اللي بيجي بالآخر

⊛ بالنهاية لل Counting rules للحصول ع الإحتمالية (probability)

$A \equiv$ Event $\Rightarrow P(A)$: the probability that A will happen

$P(A) \equiv$ probability $\equiv$ a chance _ Likelyhood that A will happen

_ How to assign probability $P(A)$? (2 Conditions)

① $\underset{\substack{\downarrow \\ P(\phi)=0}}{0} < P(A) \le \underset{\substack{\downarrow \\ P(S)=1}}{1}$ ② $\sum P(A_i) = 1$

الحادث مستحيل

— Probability Methods (approachies)

[I] Classical Method:

; Assume All sample points are equaly Likely;
(Same probability)

$$S = \{S_1, ----, S_n\}$$

$$P(S_1) = P(S_2) = \cdots P(S_n) = \frac{\textcircled{1}}{n} \rightarrow \text{لأن الفضاء العيني واحد}$$

$\overset{\text{ادارة A}}{A}$ ⇒ A event with K sample points

$$P(A) = \frac{\textcircled{K}}{\textcircled{n}} \begin{array}{l} \rightarrow \text{عدد عناصر الحادثة} \\ \rightarrow \text{عدد العناصر الكلي} \end{array}$$

Exp: Toss a coin (Fair)

$$S = \{H, T\} \quad P(H) = P(T) = \frac{1}{2}$$

Exp: Roll a die (Fair)

$$S = \{1, 2 \cdots 6\}$$

$$P(1) = P(2) = \cdots P(6) = \frac{1}{6}$$

Exp: Roll 2 dice:

$$\#S = (6)(6) = 36$$

$$P(1,1) = P(1,2) = \cdots\cdots P(6,6) = \frac{1}{36}$$

Define $A \equiv$ sum of the two faces is: 7

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

$P(A) = \frac{6}{36} = \frac{1}{6}$

Def $B \equiv$ sum: 12

$B = \{(6,6)\}$ $\boxed{P(B) = \frac{1}{36}}$

Def $C \equiv$ sum: 1

$P(C) = \frac{0}{36} = 0 \Rightarrow$ imposible

$\boxed{2}$ Relative frequency method:

; if an experement repeted $\underline{n}$ times and event 'A' occurs '$\underline{f}$' times then:

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

Exp: In a sample of BZU: Blood Type:

A : 10 Find $P(A) = \frac{10}{60} = \frac{1}{6}$

B : 15 $P(B) = \frac{15}{60}$

AB : 8

O : 27 $P(AB) = \frac{8}{60}$

$n = 60$

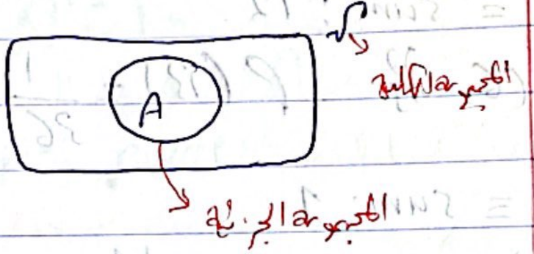
$P(O) = \frac{27}{60}$

$\boxed{3}$ Subjective method, (اعتقاد شخصي لازم الإحتمال بتاع حاجة)

$P(R) = 90\% \checkmark$

$P(R) = 60\% \checkmark$

$\underset{\text{set}}{\underset{\Downarrow}{\#(S)}} \Rightarrow \underset{\text{subset}}{\underset{\Downarrow}{P(A)}} \Rightarrow$ by Venn diagram:



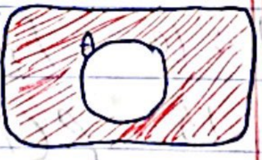
(4.3) Rules of probability

$P(A) \equiv$ The probability that A will happen

- ↳ Classical Method
- ↳ Relative frequency
- ↳ Subjective

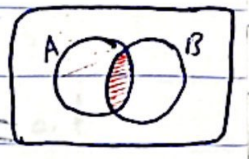
$\boxed{1}$ The probability that event A will not happen $\equiv P(A')$ is:

$$\boxed{P(A') = 1 - P(A)}$$



[2] The probability that both events will happen is:

$$P(A \cap B) \equiv P(A \text{ and } B)$$

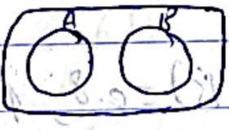


↳ If $A \cap B = \phi \Rightarrow A, B$ are disjoint

disjoint means : $P(A \overset{}{\underset{\Uparrow}{\cap}} B) = 0$

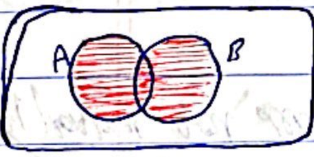
$P(\phi) = 0$

↳ If $P(A \cap B) = 0$, then A, B are called mutually exclusive ;

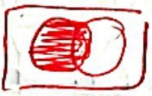


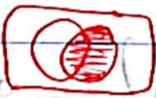
[3] The probability that (at least) one of the 2 events will happen (A or B or Both) is:

$$P(A \cup B) \equiv P(A \text{ or } B)$$
$$= P(A) + P(B) - P(A \cap B)$$



طرحت تقاطعهم لأنهم أخذوه مرتين ;

$P(A) =$


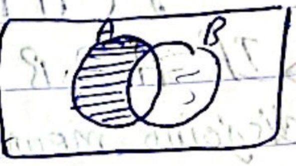
$P(B) =$


تقاطعهم أخذ مرتين
وأنا بدي إياه مرة وحدة فقط

4 $P(A-B)$:

$= P(A \cap B')$ ؛ حدوث A مع عدم حدوث B

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



عدم حدوث A أو B أو عدم حدوث الاثنين بالتالي
عدم حدوث A مع عدم حدوث B بالتالي

neither A nor B

5 $(A \cup B)' = A' \cap B'$ not A and not B

$(A \cap B)' = A' \cup B'$ Not A or not B

not (A and B)

Exp: If $P(A)=0.7$, $P(B)=0.8$, $P(A \cup B)'=0.1$

① Find $P(A \cap B)$, A, B are mutually exclusive or not ??

$\Rightarrow P(A \cup B) = 1 - 0.1 = 0.9$

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

$0.9 = 0.7 + 0.8 - P(A \cap B)$

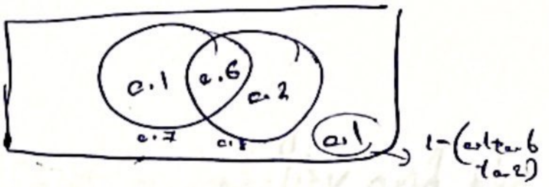
$\Rightarrow P(A \cap B) = 0.6 \neq 0$ so they are not mutually exclusive.

② Find $P(A-B) = P(A \cap B')$

$\Rightarrow P(A) - P(A \cap B) = 0.7 - 0.6 = 0.1$

③ Find $P(B-A) = P(B \cap A')$

$= P(B) - P(A \cap B) = 0.8 - 0.6 = 0.2$



④ $P(A \cup B') = P(A) + P(B') - P(A \cap B')$

$= 0.7 + (1 - 0.8) - (0.7 - 0.6)$

$0.7 + 0.2 - 0.1 = 0.8$

⑤ $P(A' \cap B') = P(A \cup B)' = 1 - 0.9 = 0.1$

Exp: Roll 2 dice :-

⇒

$\# S = (6)(6) = 36$ pairs

$S = \{(1,1), \dots, (6,6)\}$

$\Downarrow$

classical Method $P(1,1) = P(1,2) \cdots P(6,6) = \frac{1}{36}$

Define the following events:

A ≡ sum of 2 faces is: 8

B ≡ The 2 faces identical

C ≡ The absolute difference between the 2 faces is: 4

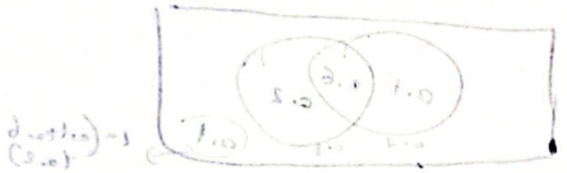
① ⇒ $A = \{(2,6), (6,2), (3,5), (5,3), (4,4)\}$

$P(A) = \frac{5}{36}$

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

$P(B) = \frac{6}{36} = \frac{1}{6}$

⇒ ③ $C = \{(6,2), (2,6), (5,1), (1,5)\} \Rightarrow P(C) = \frac{4}{36} = \frac{1}{9}$



④ ⇒ $P(A') = 1 - P(A)$

$= 1 - \frac{5}{36} = \frac{31}{36}$

⑤ ⇒ Examples of 2 mutualy exclusive events

$B \cap C = \phi$

⑥ $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

$\frac{5}{36} + \frac{6}{36} - \frac{1}{36} = \frac{10}{36}$

⑦ $P(B \cup C) = P(B) + P(C) - P(B \cap C)$

$= \frac{6}{36} + \frac{4}{36} - 0 = \frac{10}{36}$

⑧ $P(A' \cap B') = P(A \cup B)'$

$= 1 - P(A \cup B)$

$1 - \frac{10}{36} = \frac{16}{36}$

Ch:(4.4):- (Conditional probability and the independant events) :-

$P(A/B)$: the probability of an event A given an event B
(شو احتمال حدوث A عند حدوث B)

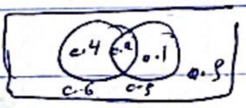
$\underset{\text{given}}{/}$

Exp: If $P(A) = 0.6$, $P(B) = 0.3$, $P(A \cap B) = 0.2$

$\Rightarrow$ Find $P(A/B) = \dfrac{P(A \cap B)}{P(B)} = \dfrac{0.2}{0.3} = \dfrac{2}{3}$

$\Rightarrow P(B/A) = \dfrac{P(B \cap A)}{P(A)} = \dfrac{0.2}{0.6} = \dfrac{1}{3}$

$\Rightarrow P(B'/A) = \dfrac{P(B' \cap A)}{P(A)} = \dfrac{P(A) - P(A \cap B)}{P(A)}$



$= \dfrac{0.6 - 0.2}{0.6} = \dfrac{0.4}{0.6} = \dfrac{2}{3}$

$\Rightarrow P(A'/B') = \dfrac{P(A' \cap B')}{P(B')} = \dfrac{P(A \cup B)'}{P(B')} = \dfrac{0.3}{0.7} = \dfrac{3}{7}$

Independent Events :- حدوث أحد الحادثين لا يؤثر على حدوث الآخر

Def: A, B are independent if

$$\boxed{P(A/B) = P(A),\quad P(B/A) = P(B)}$$

↳ independed تعريف

Exp: If $P(A) = 0.5$, $P(B) = 0.4$, $P(A \cup B) = 0.6$

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

$$0.6 = 0.5 + 0.4 - P(A \cap B)$$

$P(A \cap B) = 0.3 \neq 0$ ∴ Not mutually exclusive

$P(A/B) \overset{??}{=} P(A)$ $\Rightarrow P(A/B) = \frac{P(A \cap B)}{P(B)} = \frac{0.3}{0.4} = 0.75$

$0.75 \neq P(A)$ ∴

A, B are not independent

$\Downarrow$

dependent

Suppose that A, B are independent

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

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

Same for $P(B/A)$

Exp: $P(A)=0.4$, $P(B)=0.5$, $P(A\cap B)=0.2$, A, B independent

$\Rightarrow P(A\cap B)=P(A)\cdot P(B)$ (or) $P(A/B)=P(A)$

$0.2 = 0.4 \cdot 0.5$ $\quad \frac{0.2}{0.5} = 0.4$

Exp: If $P(A)=0.6$, $P(B)=0.3$ Find $P(A\cup B)$:

① A, B are mutualy exclusive: $P(A\cup B) = \cancel{P}P(A)+P(B) - 0$
$= 0.9$

② A, B independent: ~~$P(A/B)=P(A)$~~
$P(A\cup B) = P(A)+P(B) - P(A\cap B)$

Since A, B independent: $P(A\cap B) = P(A)\cdot P(B)$
$P(A\cup B) = P(A)+P(B)-P(A)\cdot P(B) = 0.6+0.3-0.6(0.3)$
$= 0.72$

③ Dependent :: No answer.

Exp: Health insurance:-

	© Yes	Ⓓ No	
Ⓐ 18-34	$\frac{750}{2000}$	$\frac{170}{2000}$	$\frac{920}{2000}$
Age Ⓑ 35 or older	$\frac{950}{2000}$	$\frac{130}{2000}$	$\frac{1080}{2000}$
	$\frac{1700}{2000}$	$\frac{300}{2000}$	$\boxed{\frac{2000}{2000}}$

⇒ Cross tabulation (Probability table) to make probability using frequency

① Joint

② $P(35 \text{ or Elder}) \Rightarrow P(18\text{-}34)$

③ Mutually exclusive events:

$\Rightarrow (C, D) \quad \Rightarrow (A, B)$

④ ⓐ $P(No) = P(D) = \frac{300}{2000} = 0.15$

ⓑ $P(No / 18\text{-}34) \Rightarrow P(D / A) = \frac{P(D \cap A)}{P(A)} = \frac{\frac{170}{2000}}{\frac{920}{2000}} = \frac{170}{920}$

ⓒ $P(D/B) = \frac{P(D \cap B)}{P(D)} = \frac{\frac{130}{2000}}{\frac{1080}{2000}} = \frac{130}{1080}$

ⓓ $P(A/D) = \frac{P(A \cap D)}{P(D)} = \frac{\frac{170}{2000}}{\frac{300}{2000}} = \frac{170}{300}$

⑤ The probability that the individual is 18-34 Ⓐ or has no insurance Ⓓ;

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

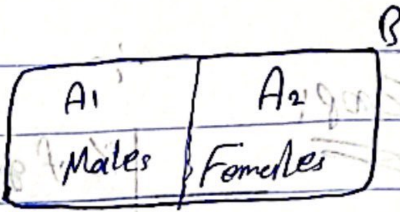
$$= \frac{920}{2000} + \frac{300}{2000} - \frac{170}{2000}$$

⑥ Does health insurance and age independent ??

خُد مثلاً A & C أو B & D أو أجزاء الأعمدة مع أجزاء الصفوف للتأكد!

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

$$\frac{750}{2000} \neq \frac{920}{2000} \cdot \frac{1700}{2000} \Rightarrow \text{Dependent}$$



- Bayes Theorem:-

$A_1 \cup A_2 = S$, $A_1 \cap A_2 = \phi$

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

$P(B) = \underline{P(A_1 \cap B)} + \underline{P(A_2 \cap B)}$

⇒ Total probability:-

$= \boxed{P(B/A_1)\,P(A_1)} + \boxed{P(B/A_2)\,P(A_2)}$

$$\boxed{P(B) = P(B/A_1)P(A_1) + P(B/A_2)P(A_2)}$$

Bayes Theorem:

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

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

Exp:

	% of good parts	% of Bad parts
supplier 1	98%	2%
supplier 2	95%	5%

$P(S_1) = 0.65$

$P(S_2) = 0.35$

① If an item selected at random what is the probability that the Selected item is Bad

B = Bad and from S_1 or Bad and from S_2

$B = (B \cap S_1) \cup (B \cap S_2)$

$P(B) = P(B \cap S_1) + P(B \cap S_2)$

$= P(B/S_1)P(S_1) + P(B/S_2)P(S_2)$

$= 2\%(0.65) + 5\%(0.35)$

$= 0.0305$

② If an item found to be Bad what is the probability that it's from supplier 1

$\Rightarrow P(S_1/B) = \frac{P(S_1 \cap B)}{P(B)} = \frac{P(B/S_1)P(S_1)}{P(B)}$

$= \frac{0.02(0.65)}{0.0305}$

Ch(5) :- (Discrete Random Variables) :-

- How to relate outcoms with numirical value :-

↳ using Random variable ; (وصف رقمي لل outcoms عن تجارب عشوائية)
(x / y / R ...)

Exp: Toss a coin 3 times :-

$\# S = (2)^3 = 8$, $S = \{HHH, \ldots, TTT\}$

- count the number of heads in each observation

# of tails = y

(y) Tails		Head (x)
0	← HHH →	3
2	← HTT →	1
1	← HTH →	2
3	← TTT →	0

⇓ x = # of heads

Exp: Roll 2 dice :-

$\#(S) = (6)(6) = 36$, $S = \{(1,1), \ldots, (6,6)\}$

- Define x = sum of the faces shown up :-

↳ $x = 2, 3, \ldots, 12$ ⇒ Discrete Random variable

- Define y = The difference between the shown faces :-

↳ $y = 0, \pm 1, \pm 2, \pm 3, \pm 4, \pm 5$ ⇒ Discrete Rand. Var.

(٣ حالات ولادة)

Exp: Consider the experement of 3 sequential births

$\# S = (2)^3 = 8$, $S = \{BBB, \dots, GGG\}$

- Define X = Number of girls :-

↳ X can assume the next values:

$X = 0, 1, 2, 3 \Rightarrow$ Disc. Rand. Var.

- Construct a Table (freq. distribution)

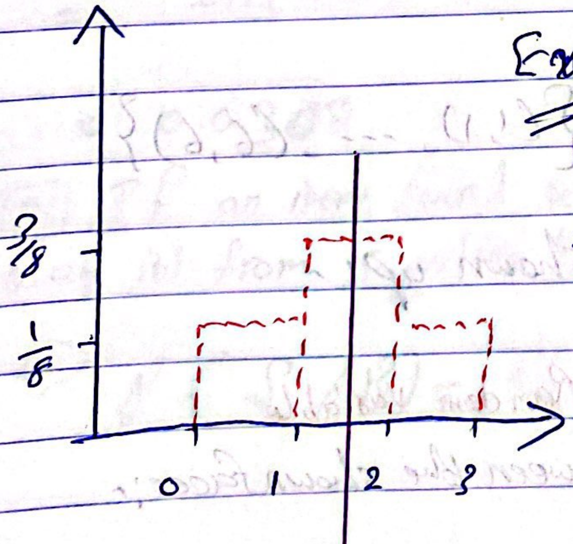
x	$P(X=x)$
0	1/8
1	3/8
2	3/8
3	1/8

$P(X=x) = f(x) =$ the probability function

$\Rightarrow$ probability distribution: (شروط)

① $0 \leq f(x) \leq 1$

② $\sum f(x) = 1$



Symetry ??

Exp: Roll a dice :-

$S = \{1, 2, \dots 6\}$

$f(x) = \frac{1}{6}$ since $\Rightarrow X = \{x_1, \dots x_n\}$

$f(x_i) = \frac{1}{n}$

X is a Uniform Disc. Rand. Var.

Exp: Consider a Disc. Ran. Var. x with probability function:-

$f(x) = \frac{x+1}{10}$, $x = 1, 2, 4$, Does $f(x)$ a valid probability function ??

$\Rightarrow$ $f(1) = \frac{1+1}{10} = \frac{2}{10}$, $f(2) = \frac{3}{10}$, $f(4) = \frac{5}{10}$

① $0 \leq f(x) \leq 1$?? $\boxed{\checkmark}$

② $\sum f(x) = 1$?? $f(1) + f(2) + f(4) = \frac{2}{10} + \frac{3}{10} + \frac{5}{10} = \frac{10}{10} = 1$ $\boxed{\checkmark}$

⊛ أي Disc. Rnd. Var. له تمثيل بثلاثة أشياء مهمة: الجدول / الرسمة / الإقتران.

- Let x be a Disc. Rnd. Var. with probability function $f(x)$ then:-

① The expected value of $f(x)$, expectations is:

$$E(x) = \sum x_i \, f(x_i) \quad (\text{weighted mean})$$

(expected)

② The variance of x , $Var(x)$:

$$V(x) = \sum (x - E(x))^2 f(x)$$

$$\sigma(x) = \sqrt{Var(x)}$$

Exp: Find $E(x)$, $f(x)$ for the example of sequantial birth

⇒ $xf(x)$	$x-E(x)$	$(x-E(x))^2$	$(x-E(x))^2 f(x)$
0	$0-1.5=-3/2$	9/4	9/32
3/8	$1-1.5=-\frac{1}{2}$	1/4	3/32
6/8	$2-1.5=1/2$	1/4	3/32
3/8	$3-1.5=3/2$	9/4	9/32
$\boxed{12/8}$	$\boxed{0}$		$\boxed{\frac{24}{32}=\frac{3}{4}}$

$$E(x)=\sum xf(x) = \frac{12}{8} = 1.5 \quad , \quad \sigma(x)=\sqrt{\sum (x-E(x))^2 f(x)} = \frac{\sqrt{3}}{2}$$

Exp: Given the following distripution:-

x	$f(x)$
-2	0.05
-1	0.15
0	a
1	0.15
3	0.2
4	0.1
5	0.1

① Find $f(0)$ (Find a)

$a = 1 - \underbrace{0.75} \rightarrow \sum f(x)$

$= 0.25$

② Find $P(-1 < x \leq 3)$

$\hookrightarrow f(0)+f(1)+f(3)$

$= 0.25 + 0.15 + 0.2$

$= 0.6$

③ $P(x \geq 0) = f(0) + f(1) \cdots + f(5)$

$= 1 - P(x < 0)$ ⇒ أسرع بدل ما أجمعهم كلهم بطرح واحدة والباقيين

$= 1 - (f(-1) + f(-2))$

$= 1 - (0.15 + 0.05) = 1 - 0.2 = 0.8$

④ Find the E(x) and Var(x)

أول خطوة ① لإيجاد الـ E(x) بجيبها: shift + 2 + 2

:- باستخدام الآلة الحاسبة :-

① تأكد من أنها مظبوطة SD mode On,

② أدخل القيم كالتالي:

$Var(x) = \boxed{\text{بعيشك انسى طلعت}}$ شينة 😅

x	f(x)
-2	0.05
⋮	⋮
5	0.1

$-2 + shift + , + 0.05 + (M+)$

يصبح معك الشاشة:

$\boxed{0.05}$

③ تكرر نفس الخطوة لجميع الأرقام والمفروض بأن المجموع يظهر عند الشاشة 1

④ اضغط shift ثم رقم ① واختر: Σ(x) اللي رقمها: 2

$\Rightarrow E(x) = 1.4$

Ch.(5.4): (Binomial Distribution):

— Binomial Experement:- Success — Falure (No success)

↳ ① the experement consists of **n trials** ($n \geq 1,0$).

↳ ② In each trial **2 outcoms** are possible ; success or failure

↳ ③ The probability of success $\equiv p$

and the probability of failure $= 1 - p\,(q)$ } → are constants on each triale.

↳ ④ The trials are **Independent**

Exp: 4 sequential births:-

- $n = 4$
- Boy or Girl
- $P(B) = P(G) = \frac{1}{2}$
- Independent

— Consider a binomial experiment with number of trials $= n$ and $P(\text{Success}) = p$, Define the Rndom variable:

$X \equiv$ # of Success in n trials then

$\Rightarrow X = 0, 1, 2, \ldots, n$

0 ← لم تنجح ولا تجربة

n ← تنجح جميع التجارب

يعني لو بدو النتيجة بنات: إما لا يوجد بنات 0

إما بنت واحدة 1

إما بنتين 2

إما كلهم بنات n

x is a Disc. Rnd. Var. and is called a Binomial Random Var.

↳ $x = B(n ; p)$

x ↓ # of success $x (0 \to n)$

n ↓ عدد المرات الي قمت بها التجربة

p ↳ نسبة نجاح التجربة (النتيجة)

Let $x = B(n, p)$, then:-

↳ ① The probability function for x is given by:-

$$f(x) = (nC_x)(p)^{x}(1-p)^{n-x}$$

n = # of trials. $f(x)$ = the probability of x success in n trials.

$p = P(\text{Success})$.

x = # of success. $x = 0, 1, 2, \ldots, n$

↳ ② The expected value of x:-

$$E(x) = np$$

n ↙ عدد المحاولات p ↳ Probability of success

↳ ③ The variance of x $Var(x) = np(1-p)$

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

♂

Exp: 62% of BZU ~~are~~ students are female students

(female & male لأنه (Binomial) هي بيعرف ال p) *

↳ ① In a sample of 1500 students; what is the expected number of female (إناث) students and what is the variance

* طالما بسأل عن ال female (إناث) إذا ال success هو ال female.

$$\hookrightarrow X = B(n, p)$$
$$= B(1500, 0.62)$$

عشان أجيب ال Var(x) بدي E(x)

$$\hookrightarrow E(x) = np = 1500(0.62) = \underbrace{930}$$

↙ توقع للعينة 1500 أن يكون عدد الإناث هو 930

$$\hookrightarrow Var(x) = np(1-p) = 930(1-0.62)$$
$$= 353.4$$

$$\sigma(x) = \sqrt{353.4}$$

↳ ② In a sample of 800 students what is the expected number of male students.

$$\hookrightarrow X = B(800, \overset{1-0.62}{\overset{\uparrow}{0.38}}), \quad \hookrightarrow E(x) = np = 800(0.38) = 304$$

$$\hookrightarrow \underbrace{f(x)} = nC_x (p)^x (1-p)^{n-x}$$

احتمال ال x ↙ $\quad {}_{1500}C_{800}(0.38)^{800}(1-0.38)^{1500-800}$

success

↳ ③ In a sample of 10 student's what is the probability of exactly 4 female (إناث) student's

$$\Rightarrow f(4) = {}_{10}C_4 \, (0.62)^4 \, (1-0.62)^6$$

↳ ④ In a sample of 8 student's what is the probability of

↳ * at least 3 female student's

$$P(x \geq 3) = f(3) + f(4) + \cdots f(8)$$

$$= 1 - P(x < 3)$$

$$= 1 - [f(0) + f(1) + f(2)]$$

↳ * no more than 2 female student's

$$P(x \leq 2) = f(0) + f(1) + f(2)$$

↳ ⑤ In a sample of 12 student's, what is the probability of at least 2 male student's

$$x = B(12, 0.38)$$

$$P(x \geq 2) = 1 - P(x < 2)$$

$$= 1 - [f(0) + f(1)]$$

:- Ch. (5.5) :- (The poisson Distripution) :-

\# of occurences in a given ⟵ period

:- (عدد مرات حدوث حدث ما في فترة زمنية محددة)

Disc. Rnd. Var. with probability function :-

$$\boxed{f(x) = \frac{\mu^{x} e^{-\mu}}{x!}}$$

:- x = # of occurences

$f(x)$ = propability of x occunences

$x = 0, 1, 2, \ldots, n$

μ = expected number of occurencees per period (avg)

Exp: 93 page (218), 10 passengers per minute

↳ ① No arrival in 1 min

$f(0)$ with $\mu = 10$

$$\Rightarrow f(0) = \frac{10^{0} e^{-10}}{0!} = 0.000045 \approx 0$$

↳ ② three or fewer in 1 min

$$P(x \leq 3) = f(0) + f(1) + f(2) + f(3)$$

$$= \frac{10^{0} e^{-10}}{0!} + \frac{10^{1} e^{-10}}{1!} + \ldots$$

↳ ③ No arrival in 15 second period

$$\begin{array}{l} 10 \text{ pass} \longrightarrow 1 \text{ min } (60 \text{ sec}) \\ ? \longleftarrow 15 \text{ sec} \\ \mu = 2.5 \end{array}$$

$$\Rightarrow f(0) = \frac{10^{0} e^{-2.5}}{0!} = 0.082$$

↳ ④ at least one per 15 sec

$P(x \geq 1)$ with $\mu = 2.5$

$= f(1) + f(2) + \cdots f(10)$

$= 1 - f(0)$

$= 1 - \frac{10^{0} e^{-2.5}}{0!} = 1 - 0.082$

↳ ⑤ What is the expected # of passengers in half hour period:-

$$1\text{ min} \longrightarrow 10\text{ pass}$$

$$30\text{ min} \longrightarrow x\text{ pass}$$

$$x = 300\text{ pass}$$

↳ ⑥ $P(2 \leq x \leq 6) = f(2) + f(3) + \cdots + f(6)$

$P(2 < x < 6) = f(3) + f(4) + f(5)$

Ch. (6) : (Continuous Random Variable) :-

Interval : $\overset{}{\underset{a}{|}}\!\!\text{———}\!\!\underset{b}{|}$

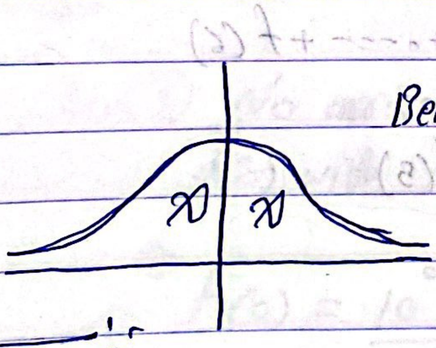
x = Continuous , $P(a \leq x \leq b)$ = Area under the curve from the top

$$= \int_a^b f(x)\,dx$$

$$P(a \leq x \leq b) = P(a < x < b)$$

$$\therefore P(x = a/b) = 0$$

- The normal distribution :- (التوزيع الطبيعي)

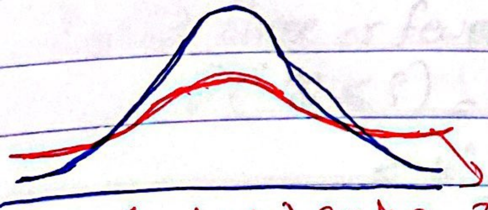


Bellshaped – Symetric

mean = median = mode

x = Normal

$= N(\mu, \sigma)$



زادت قمته باتجاه العلو (زادت نسبتها)

mean : يحدد من

&

S.D

↳ $f(x)$; probability function :

$$= \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$$P(5 < x < 10)$$

$$\int_5^{10}$$

– Standard normal distribution :-

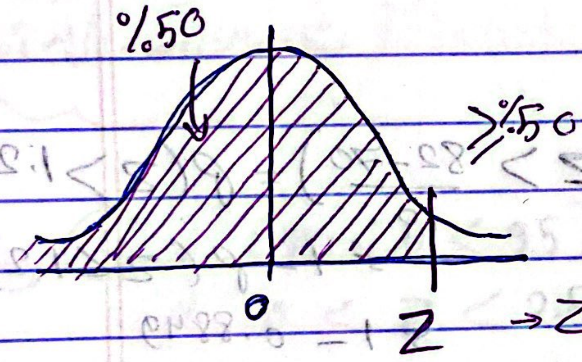
Normal distribution with $\mu = 0$ and $\sigma = 1$

$$X = N(\mu, \sigma) \xrightarrow{Z = \frac{x-\mu}{\sigma}} Z = N(0,1)$$

عشان أسهل
احل بجوله

$\Downarrow$

Z - Table



$\rightarrow Z \geq 0$

– Cumulative Probability for the standard normal distribution :-

من الجدول الخاص بقيم Z

$1.94 = \underbrace{1.9}_{\text{من العمود}} + \underbrace{0.04}_{\text{من الصف}} \Rightarrow P(Z < 1.94) = 0.9738$

$P(Z < 2.00) = 0.9772$

$P(Z < 2.15) = 0.9842$

$P(Z > 2.15) = 1 - P(Z < 2.15) = 1 - 0.9842$

Exp: $x = N(\mu, \sigma)$, $Z = \frac{x-\mu}{\sigma}$ (ch. 6.2)

Stat 2311 Midterm exam are normaly distriputed with mean of 70 and S.D 10.

$\Rightarrow$

$x = N(70, 10)$

↳ ① Find $P(x > 82)$

$$\Rightarrow P(x > 82) = P\left(Z > \frac{82-70}{10}\right) = P(Z > 1.2)$$

$$= 1 - P(Z < 1.2)$$

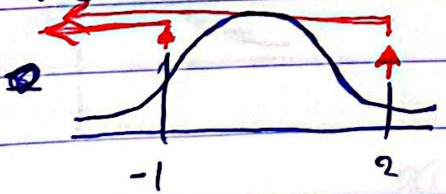
$$= 1 - 0.8849$$

$$= 0.1151$$

↳ ② $P(60 < x < 90)$ → مساواة أو بدونا مساواة نفس الشيء

$\Downarrow$

$P(-1 < Z < 2)$



$= P(Z < 2) - P(Z < -1)$

فش بالسالب بالجدول

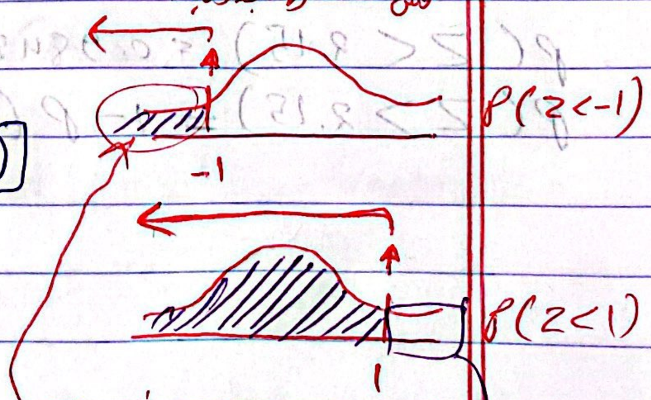
$= P(Z < 2) - P(Z > 1)$

$= P(Z < 2) - [1 - P(Z < 1)]$

$= P(Z < 2) + P(Z < 1) - 1$

$= 0.9772 + 0.8413 - 1$

$= 0.8183$

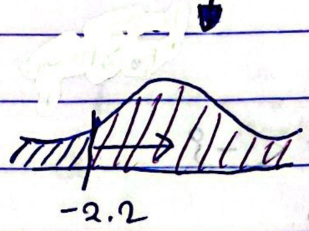


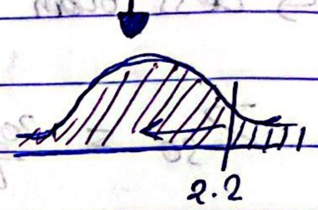
كتا 2 للقطعة [illegible]

$1 - P(Z < 1)$

$P(Z < 1) - 1$

↳ ③ $P(x > 48) = P(Z > -2.2) = P(Z \leq 2.2)$





$= 0.9861$

✻
الجدول Z بشتغل
ع إشارة أصغر من <
=
عشان هيك منحول أي إشارة أكبر

↳ ④ $P(70 < x < 95)$

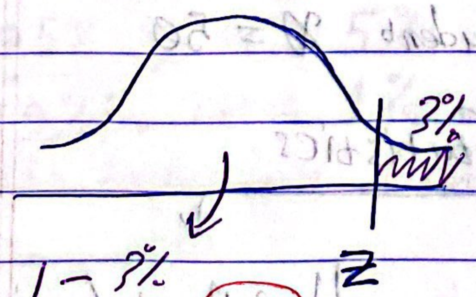
$P(0 < Z < 2.5) = P(Z < 2.5) - P(Z < 0)$

$= 0.9938 - 0.5$

$= 0.4938$

↳ ⑤ Find a score x such that 3% of the score are greater than z to this value.

⇒ $P(X > x) = 3\%$



$1 - 3\%$

$1 - 0.03 = \boxed{0.97}$

بروح ع الجدول
بشوف الأقرب إلها
وباخد الصف والعمود

⇒ $Z = 1.88$

$Z = \frac{x - \mu}{\sigma}$

$x = 88.8$

↳ Less than 30

$Z_{30} = \frac{30-70}{10} = -4 < -3$

↓

Extreme (outlier)

Ch.(8.1):(Inferential Statistics):-

(إحصائيات إستنتاجية)

sample → statistics

population → parameter

We need → sample
→ propability

- Target population:-

↳ Unknown parameter:-

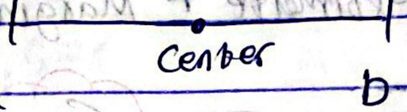
BZU daily expendrure for student $\bar{x} = 50$ ← فرض

Inferential Statistics

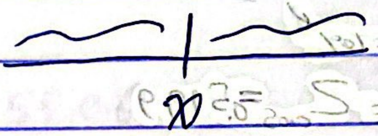
Estimation (تقدير) → نقطة point (single) | فترة Interval

Hypothesis (فرضيات) → نختبرها tests

- Estimation:
  - ↳ Interval:
    ; Range of values



  - ↳ point (single):
    point estimate ± Margin of error (E)



$\bar{x} \pm E$

- Each interval should be construeted with a given probability. this probability called: (Confidence Level)

$(100 - \alpha)$ 100% , α = type of error

- Common levels:

90% , α = 10%
95% , α = 5%
$\boxed{99\%}$, α = 1%
↳ heighest Level of confidence.

(*) For μ:
↳ ① σ - (Known case):
تحولها لل probability + ال SD
معروفين و نشتغل ع جدول (Z)

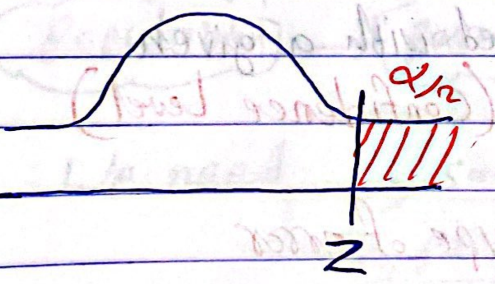
↳ ② σ - (Unknewn case):
نشتغل على جدول (t)

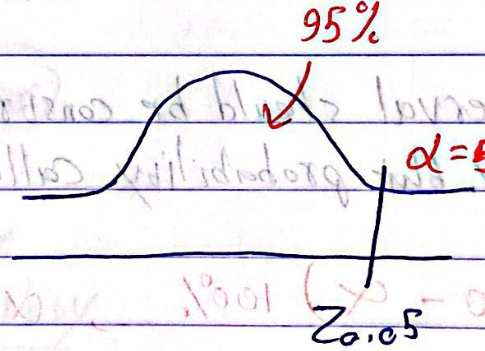
- The $(100-\alpha)100\%$ confidence interval for μ (σ- Unknown case) is :

Point estimate ± Margin of error

$$\equiv \bar{x} \pm Z_{\frac{\alpha}{2}} \boxed{\frac{\sigma}{\sqrt{n}}} \rightarrow \text{Standard error of the mean}$$

- Find Z for 90, 95, 99





Exp: A random sample of size 100 (n) provided a sample mean of 60 ($\bar{x}$), assume that population standard deviation is 30 (σ- Known) :- Construct 90%, 95%, 99% Confidence Interval for μ

$$\Rightarrow \bar{x} \pm Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \Rightarrow 60 \pm Z_{\frac{\alpha}{2}} \frac{30}{\sqrt{100}}$$

$$60 \pm Z_{\frac{\alpha}{2}}(3)$$

Confidence level	α	$\alpha/2$	$Z_{\alpha/2}$
90%	10%	0.05	1.645
95%	5%	0.025	1.960
99%	1%	0.005	2.576

① ⇒ since 90% $Z_{\frac{\alpha}{2}}(0.05) = \underline{1.645}$

$60 \pm 3(1.645)$

$60 \pm 4.94 \Rightarrow \underbrace{(55.06, 64.96)}$

There's a probability (chance) of 90% = 0.9 that the population mean μ is between 55.06 to 64.96

↳ ② 95%

$Z_{\frac{\alpha}{2}} = 1.960$, $60 \pm 3(1.960)$

$60 \pm 5.88 \Rightarrow (54.12, 65.88)$

↳ ③ 99% $\Rightarrow 60 \pm 3(2.576)$

$60 \pm 7.73 \Rightarrow (52.27, 67.73)$

⊛ لو أجعل ال confidence أعلى بـ أقل خطأ:

أقوم بزيادة ال sample لأربعة أضعاف: $n * 4 = \underline{400}$

- Know for σ-Unknown case:

$$\bar{x} \pm \underbrace{t_{\frac{\alpha}{2}}}_{\text{من جدول بتطلع عليه}} \frac{S}{\sqrt{n}} \quad , \quad \underbrace{df}_{\text{degrees of freedom}} = n-1$$

ال t-Distribution

(*) لما تزداد قيمة df بازدياد $\underset{s}{(n)}$ تقترب قيمة t من Z $(t \approx Z)$

Exp: A random sample of size $\overset{n}{25}$, provided a sample mean of $\overset{\bar{x}}{100}$ and a standard deviation of $\overset{s}{20}$.

- Construct 95% Confidence interval for μ

* ونستخدم ال (t-dis.) لما نتعامل مع (sample). وإذا ال population
نستخدم Z

$$\hookrightarrow t_{\frac{5\%}{2}} = t_{0.025} \Rightarrow 100 \pm t_{0.025} \frac{20}{\sqrt{25}}$$

$$\Rightarrow 100 \pm \underbrace{(2.064)}(4) = 100 \pm 8.26$$

$$= (91.74 , 108.26)$$

بروح ع الجدول عمود
ال $\boxed{0.025}$
والصف بتطلع عليه مو
ال $df = n-1$
$= 25-1 = \boxed{24}$

(*) Standard error of the mean: $\boxed{\frac{S}{\sqrt{n}}}$

-:- Ch.(9) -:- (Hypothesis Testing) :-

; Claim about a value of a parameter.

↳ The null hypothesis:

↳ ; claim about μ (H_0)

بأي شخص محاكس معه فرضية بطلع (محاكسة).

↳ The alternating hypothesis;

↳ :(H_a): Researcher hypothesis.

- Conclusion :-

↳ Regect H_0 : the test is significant when H_0 regected

(or)

↳ Do not Regect H_0 ;

-:-

- H_0 : claim about μ (Given value for null hypothesis)
- H_a : Alternating hypothesis comes from the resercher :-

↳ Resercher
→ Reject H_0 ✓
→ Do Not Regect H_0 ✗

Steps:-

↳ $\boxed{1}$ State H_0 and H_a

↳ $\boxed{2}$ select a level of significance

select α (type of error I (نوع أول))
Probability: Pr(type I error)

↳ $\boxed{3}$ Identify a test statistics:-

$$Z\text{-test}: \; Z = \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$

↳ $\boxed{4}$ formulate a rejection rule:-

↳ Critical values approch
↳ P-value approch
} ⇒ بنحسب بس ال Z-test
(Calculated Z)

↳ $\boxed{5}$ Conclusion:-

Reject H_0 ✓

Do Not Reject H_0 ✗

Hypothesis	Statistic Test	Regection Rule	Critical Value(s)
1. $H_0: \mu = \mu_0$ $H_a: \mu \neq \mu_0$	$Z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$	Regect H_0 if $Z > Z_{\frac{\alpha}{2}}$ or $Z < -Z_{\frac{\alpha}{2}}$ (2) two tailed test	$\pm Z_{\frac{\alpha}{2}} = \pm 1.96$
2. $H_0: \mu \geq \mu_0$ $H_a: \mu < \mu_0$	$Z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$	Regect H_0 if $Z < -Z_\alpha$ Left (lower) taield test (one t test)	$-Z_\alpha = -1.645$
3. $H_0: \mu \leq \mu_0$ $H_a: \mu > \mu_0$	$Z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$	Regect H_0 if $Z > Z_\alpha$ Right (upper) tailed test (one tail	$Z_\alpha = 1.645$

$\alpha = 0.05$

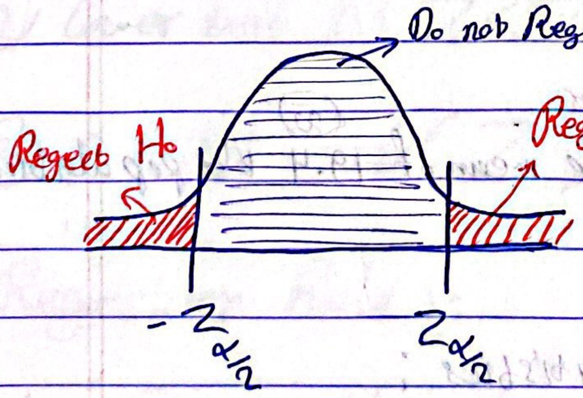
✱ H_a بتحدد الـ tailed إثنين ولا واحد : $H_0 \Rightarrow \mu = 10$

$H_a \Rightarrow \mu > 10$ → الحالة الشغ :

$H_0 : \mu = \mu_0$

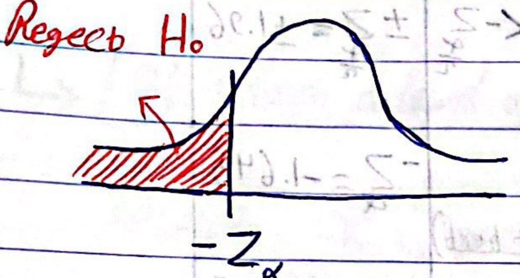
① $H_a : \mu \neq \mu_a$ → 1. Hypothesis : ↳ 2- tailed test

↳ $\pm Z_{\frac{\alpha}{2}}$ critical values



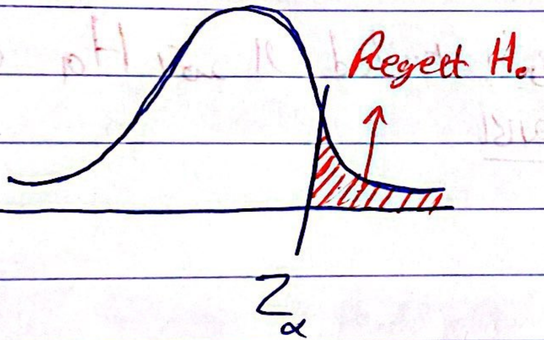
———— :-

② Regect H_0 if $Z < -Z_\alpha$



↳ One-tailed test (Lower)
↳ Critical = $-Z_\alpha$

③ Regect H_0 if $Z > Z_\alpha$



↳ One-tailed test (upper)
↳ Critical = Z_α

———— :-

Exp: $H_0 : \mu \geq \boxed{20} \rightarrow \mu_0$
$H_a : \mu < 20$ ∴ Lower

sample of $\overset{n}{50}$ provided a sample mean of $\overset{\bar{x}}{19.4}$ the population standard deviation is $\overset{\sigma}{2}$

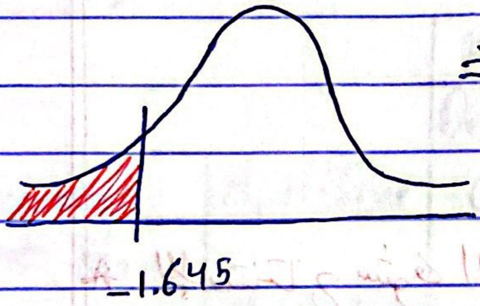
↳ ① Compute value of test statistics:

$$Z = \frac{\bar{x} - \mu_0}{\sigma} \sqrt{n} = \frac{19.4 - 20}{2} \sqrt{50} = -\cancel{3}2.12$$

— since we are in the case 2 : 1 critical value:-

↳ Regect H_0 if $Z < -Z_\alpha$

$\alpha : 5\% \Rightarrow -Z_{0.05} = -1.645$



$\Rightarrow Z = -2.12 < -Z_\alpha = -1.645$

∴ Conclusion : Regect H_0

— P-value approch :- (another rule for regection)

① upper tail $P(Z > Z(\text{test}))$

calculated

② Lower tail $P(Z < Z(\text{test}))$

③ 2 tailed $2P(Z > |Z|)$

— Regection Rule :-

↳ Regect H_0 if P-value $< \alpha$

↳ Do not Regect if P-value $\geq \alpha$

$Z = -2.12$ lower tail

↳ calculated

↓

$P(Z < -2.12)$

$= P(Z \underset{\Updownarrow}{>} 2.12)$

$= 1 - P(Z < 2.12)$

$1 - 0.9830 = 0.017$

* الإستنتاج بشوف ال p-value :

$p\text{-value} < \alpha = 0.05$

$0.017 < 0.05$

∴ Regect H_0

∴

↳ if α changed to be 1%

$p\text{-value} = 0.017 > 0.01$

∴ Do not Regect H_0

$P(\text{Regect a true } H_0) = \text{level of significan } (\alpha)$

$P(\text{type II error}) = \beta$

; the power of the test : $1-\beta$

* كل ما كانت β أقل بكون الشغل أحسن.

		Researcher	
		Do not Regect H_0	Reject H_0
Conclusion about H_0	H_0 is: True	Correct Conclusion	Type I Error
	H_0 is: False	Type II Error	Correct Conclusion

* α إذا لم تحدده السؤال تكون 0.05.

Exp: $H_0: \mu = 15$

$H_a: \mu < 15$ (lower) , $n=30$, $\bar{x} = 13.5$, $\sigma = 3.4$

$$\Rightarrow Z = \frac{\bar{x} - \mu_0}{\sigma}\sqrt{n} = \frac{13.5-15}{3.4}\sqrt{30} = -2.42$$

$$\Rightarrow P\text{-Value} = P(Z < -2.42)$$

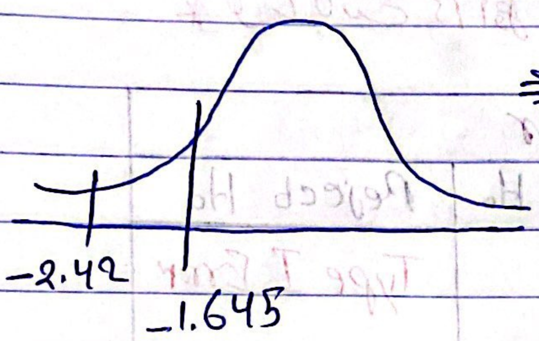
$$= 1 - P(Z < 2.42)$$

$$= 1 - 0.9922 = 0.0078$$

using cribical value approch:-

Regect H_0 if $Z < -Z_\alpha$

$\alpha = 0.05 \Rightarrow -Z_{0.05} = -1.645$



$\Rightarrow$ since $Z = -2.42 < -1.645$

Regect H_0

↳ if α changed to be 1% $\alpha = 0.01$

$P\text{-value} < \alpha$

$0.0078 < \alpha$

$\therefore$ Regect H_0

Exp: $H_0: M \geq 75$, $n = 64$, $\bar{x} = 70$, $\sigma = 16$

(a.) since $H_0: M \geq 75$, $H_a: M$ should be $< 75 \Rightarrow H_a: M < 75$

$$Z = \frac{70 - 75}{16}\sqrt{64} = -2.5$$

- Regect H_0 if $Z < -Z_\alpha$ & $\alpha = 1\%$

$-Z_{0.01} = -2.326$, since $-2.5 < -2.326$

Regect H_0

the Instructor claim is not true.