Stat 2361

Done by:-
Yumna Marei

Good luck for all

Chapter 1 "1"

"Data and statistics"

تحليل تلخيص جمع فن و علم الإحصاء

* Statistics :- Science and art of collecting, Summarizing, analyzing, and interperting data. تفسير

* Data :- Facts and figures collected :- جمع الحقائق والأرقام بيانات

مصدر البيانات Data sources
- → Existing Source → مصدر موجود
- → Collected source → مصدر بجمع البيانات بنفسي

→ Example :-

1. → Existing Source → Data Bases of data :- موجود في الشركات
2. → Collected source → Observational data statzcals study.

Type Data
- Raw Data بيانات خام (غير مرتبة)
- Organized Data بيانات منظمة (مرتبة)

* Example :-

3 variable

Kid الطفل عدد	Gender الجنس	Age العمر	Favorite color اللون المفضل
1	M	4	White
2	F	5	Pink
3	M	6	Blue
4	F	4	Blue
5	F	3	Red

5 elements

→ 5 observation

شيء ملموس Data set = مجموعة الأ... all data collected for a certain study.

Type Data :- organized Data :- بيانات مرتبة

1. what the Data ? Gender, age, Favorite color.
2. what the observation? الملاحظات 5 observation.
3. what the variable ? صفات المتغيرة 3 variable = (gender, age, favorite color).
4. what the Elements? العناصر 5 elements (kind 1 ...... kind 5)

تصنيف البيانات Data classification

- نوعي (لا يوجد عمليات حسابية) Qualitative.
  - مقياس اسمي Nominal scale
  - مقياس رتبي. ordinal scale
- كمي (عمليات حسابية) Quantitative
  - مقياس الفاصل Interval scale
  - مقياس نسبي. Ratio scale

* Scales of measurement :- موازين القياس

1. → Nominal scales → names :- البيانات عبارة عن أسماء فقط
2. → ordinal scales → names and order :- (البيانات عبارة عن أسماء وترتيب (أفضلية)).
3. → Interval scales → names and order, subtraction has a meaning).
   أسماء و ترتيب و طرح
4. → Ratio scales → names and order, subtraction and ratio have a meaning).
   أسماء و ترتيب و طرح و قسمة

→ Example :-

1. → Gender → M, F → Qualitative (nominal scales).
2. → Age → 1, 2, 3, ---- → Quantitative (Ratio scales).
3. → Coler → Red, Blue ... → Qualitative (nominal scales).
4. → phone number → Qualitative (nominal scales).
5. → Salary → Quontitative (Ratio scales).
6. → Date of Birth day → Quantitative (interval scales).
7. → Address → Qualitative (nominal scales).
8. → Rating (Exemless, very good, good, Bad) → Qualatative (ordinal scales).

"3"

من ناحية الزمنية Data classification by time

- Cross-sectional Data ← بأخذ البيانات في نفس الفترة الزمنية.
  - "Same time"
- Time-series Data بأخذ البيانات في أوقات مختلفة.
  - different time periods

* Example :-

بحسب درجة الحرارة في فلسطين لمناطق مختلفة ولكن في نفس الوقت. ←

1 → Temperatur in defferent Palestinian cities on 14/9/2021 → Cross-sectional Data.

بدي أعرف سعر ذهب من تاريخ 2021/9/1 إلى تاريخ 2021/9/14 أوقات مختلفة ←

2 → Gold prices in the period 1/9/2021 — 14/9/2021 → Time-series Data.

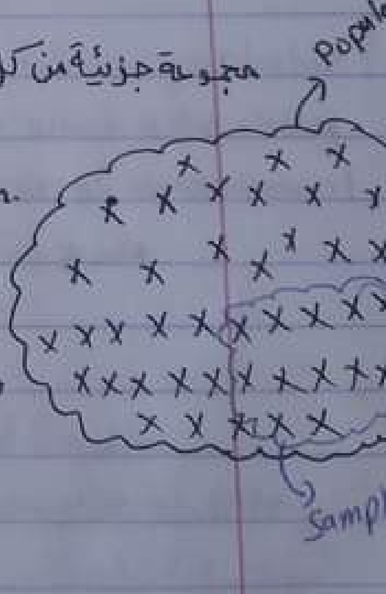
* Data Acquisition Error :-

1) → 20, 21, 22, 23, $\overset{24}{\cancel{42}}$, ---- → Quantintive (interval).

2) → 1920, 1915, $19\overset{10}{\cancel{07}}$ ----→

* population :- set of all elements to be studied → كل العناصر الممكن تندرس
  (مجتمع)

* Sample :- subset from the population → مجموعة جزئية من كل المجتمع
  (عينة)



دراسة على المجتمع * Census/population Census :- study on the whole population.

دراسة على عينة * Survey/Sample survey: study on a Sample

* دراسة على المجتمع أدق وأوضح من دراسة على عينة من المجتمع.
ولكن دراسة على عينة أسهل.

معدل أعمار الطلاب (للمجتمع) ← $\bar{X} = 25$

معدل في العينة ← $\mu = ??$ → مابقدر أعمم الرقم 25 على كل المجتمع

* Statistical inference :- الاستدلال الإحصائي

The process of making conclusions (عملية توزيع أو تعميم) on the population based on a Sample survey (على عينة).

$\bar{X} = 25 = \mu = 25$. طبعاً بقدر أعمم إذا كانت العينة البسيطة متشابهة ←

Chapter 2 :- Descriptive statistics: Tabular and Graphical presentations. "4"

Section 2.1

Summarizing Qualitative Data :- تلخيص البيانات النوعية

لا يجوز عمل عليها عمليات حسابية

* Example :-

1 stars → Hated it : كرهتوا

2 stars → Did not like it : ما حبيتوا

3 stars → like it : حبيتوا

4 stars → Really like it : حبيتوا كثير

5 stars → loved it : حبيتوا كثير كثير

* We want to rate a certain product example :- Clothing, perfume, Book, Movie, Song, Electric Device, Hotel, Food, electroizi device.

Sample :-
4s 2s 3s
5s 3s 2s
1s 4s 5s
3s 4s 2s
S = stars

* We are intersted in rating an electroizi device ? we took a Sample.

سألنا مجموعة من الأشخاص اللي بيستخدموا الأجهزة الكهربية

Q1 → what is the population ? all people who used thies device.

Q2 → How many elements do we have in the population = what is the population size ?
we don't know, The population size N is unknow.

Q3 → what is the sample size ? The sample size $= n = 12$
who many obzirfaionts/alemoud ? 12.

Q4 → How many veriables do we have ? what are many ? Are the qualitative or quantiative ?
what is the scales of measuremant ? we have 1 veriable which is the reting of the electroizi Device. It is qualitateve It has ordinal scales of mansuremant.

Q5 → Construct a frequency distrabution ? *توزيع بتكرار

يعني العدد أكم مرة تكرر في المجموعة الجزئية

Rating	Frequency dis
1	1
2	3
3	3
4	3
5	2
Total	12

→ This table is Called freq dis....

This table is a descriplive statiste

هذا الجدول مقياس وصفي.

Tabular Presentaion:- تمثيل جدولي

يجب أن يكون المجموع يساوي الرقم 12

تكرار النسبي .

Def :- Relative frequency = $\frac{\text{frequency}}{\text{Sample size}}$

Sample :-

4 2 3
5 3 2
1 4 5
3 4 2

Q6 → Construct a Relative frequency dist-- ?

Rating	Relative freq	
1	0.08	1/12 = 0.08
2	0.25	3/12 = 0.25
3	0.25	3/12 = 0.25
4	0.25	3/12 = 0.25
5	0.17	2/12 = 0.1666 ≈ 0.17
total	1.00	→ يجب مجموعهم يساوي 1

* ملاحظة :- بأخذ منزلتين بعد الفاصلة وكان ضروري أقرب لأقرب منزلة عشرية .

Def :- percent frequency = $\frac{\text{frequency}}{\text{Sample size}} \times 100$

Q7 → Construct a percent frequency dist-- ?

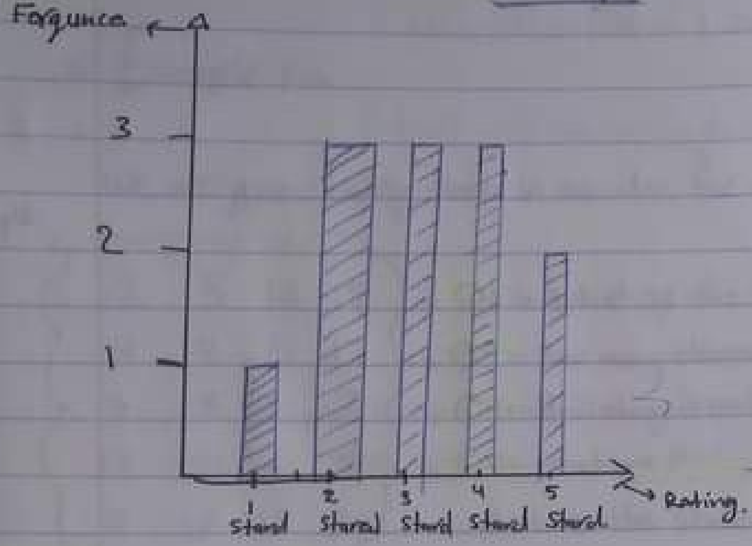
Rating	Percent freq	
1	8.33	(1/12)×100 = 8.33
2	25	(3/12)×100 = 25
3	25	(3/12)×100 = 25
4	25	(3/12)×100 = 25
5	16.67	(2/12)×100 = 16.67
Total	100	→ مجموعهم يجب ان يطلع 100

* The Tabel is Frequnse, Relative freq--, presenet freq--- to gather ?

rating	Frequnce dis-	Relative freq--	persent freq--
1	1	0.08	8.33
2	3	0.25	25
3	3	0.25	25
4	3	0.25	25
5	2	0.17	16.67
Total	12	1.00	100

شريط رسم البياني

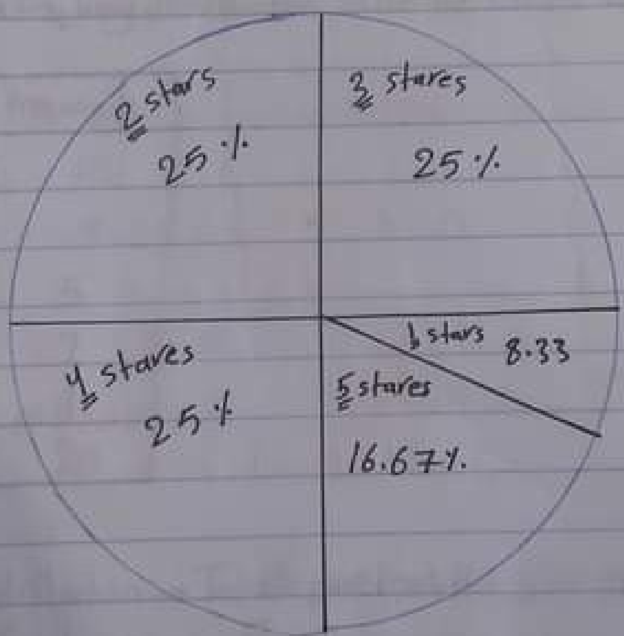
Q8 → Construct a Bar Graph :- (used :- Frequnces)



Rating	Freq
1	1
2	3
3	3
4	3
5	2
Total	12

قطاع دائري (دائماً نسبة).

Q9 → Construct a pic chat :- (used : percent freq).



Reting	Present Freq
1	8.33
2	25
3	25
4	25
5	16.67
Tota	100

Section 2.2

Summarizing Quantitative Data.

* Example :-

We are given waiting times in minutes for a certain clinic :-

Sample.

2	5	10	12
4	4	5	17
11	8	9	8
12	21	6	8
7	13	18	3

→ Q1:- What is the sample size? $n = 20$

→ Q2:- How many elements do we have? 20 Elements.

→ Q3:- How many observation do we have? 20 observations.

→ Q4:- What are the variables? Are the qualitative or quatitative? what is the scales of measuremant?

Varaibles: waiting time in minute, Quantitative variables, It has the Ration scale of measurement.

Q5:- Contruct a frequeny distribution? Use the Classes ⇒ (0–4), (5–9) and so on?

Class	Frequeny dis
0–4	4
5–9	8
10–14	5
15–19	2
20–24	1
Total.	20

* Frequeny distribution → Tabula peresentaith Descriptive statistic.
* number of classes = 5
* Classes width = 5 number → [0, 1, 2, 3, 4] (0–4) الأعداد الموجودين الارقام

بالقانون
$(4-0)+1 = 5$ → classes width = (15 – 19) for all classes.

lower class limits ← → upper class limits.

$(9-5)+1 = 5$

$(14-10)+1 = 5$

⋮

قانون

Classes width = (upper class limaite – lower class limaite) + 1

Sample :-

2	5	10	12
4	4	5	17
11	8	9	8
12	21	6	8
7	13	18	3

→ Q6 → Coustract arealtive freq. Dist ?

→ Q7 → Coustract a percent freq. Dist ?

"we have 3 tabular presentation".

Class	frequency dist	Real freq. Dist = $\frac{Freq}{Total}$ =	precent freq. dist = real frq × 100
0-4	4	0.20	20
5-9	8	0.40	40
10-14	5	0.25	25
15-19	2	0.10	10
20-24	1	0.05	5
Total	20	1.00	100

→ frequency → real freq. Dist → precent freq. dist.

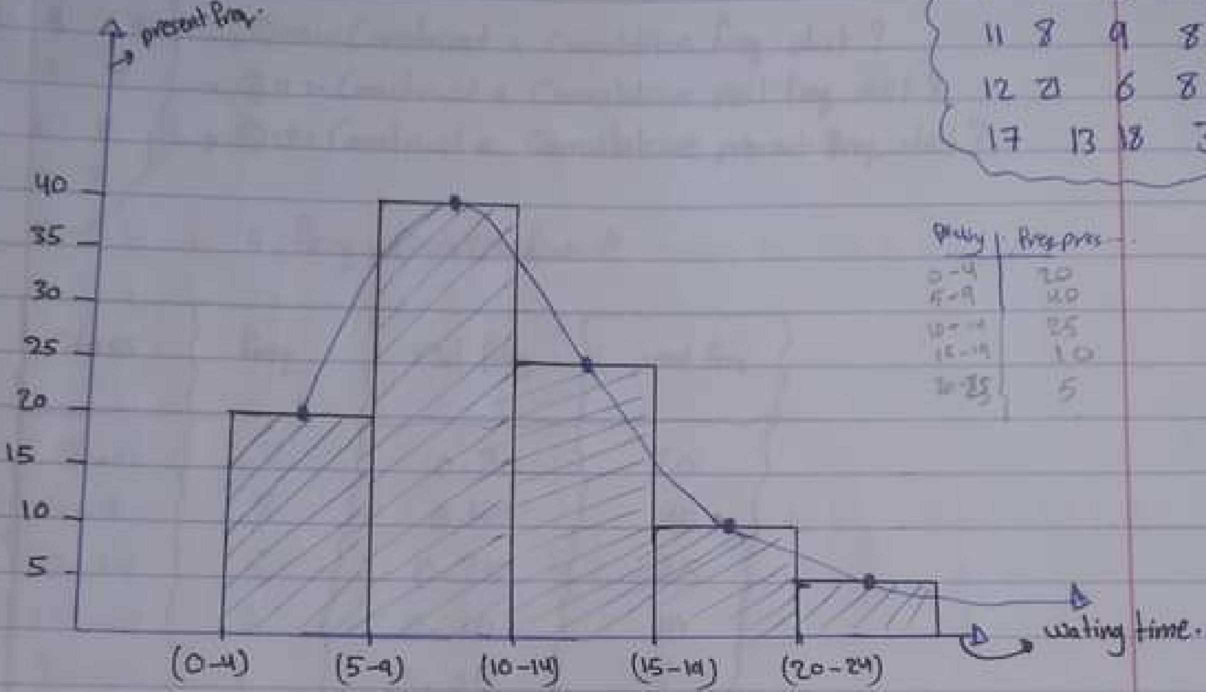
Data frequency dist
- → give intruction . I
- → Accounting to ree application (Easy to read) . II
- 1). no, of classes form 5 to 20
- 2). approx classes width = $\frac{\text{largest value} - \text{Smallest value}}{\text{no. of classes}}$ III
- → Classes width = approx . Classes width rounded up.
- → 3). Class limits = (upper class limit – lower class limit) + 1

"9"

Sample :-

2	5	10	12
4	4	5	17
11	8	9	8
12	21	6	8
17	13	18	3

Q8 :- Constrct a Histograme ? (uses present Freq.).



Freq.	Freq. pres
0-4	20
5-9	40
10-14	25
15-19	10
20-25	5

Q9 :- Is the instogram skewed no the right ? to the left ? or symertinis . .

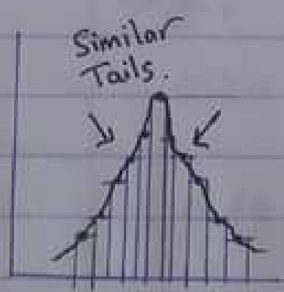
Skewed to the right.

✱ Skew ness :-

Histogram 1

The longer Tail goes to the left.

Histogram 1 = Skewed to the left



Histogram 2

Both tails are simular.

Histogram 2 = Summeluic

Histogram 3

The longer tail goes to the right.

Histogram 3 = skewed to the right.

* Cumulalrive Distributions :- التوزيعات التراكيـة

Sample :-

2	5	10	12
4	4	5	17
11	8	9	8
12	21	6	8
7	13	18	3

→ Q10 :- Constract a Cumulalrive freq. dist ?
→ Q11 :- Construct a Cumulalrive real freq. dist ?
→ Q12 : Construct a Cumulalrive precent freq. dist ?

"freq. Distribution"

Class	freq	real freq	prest freq
0 – 4	4	0.20	20
5 – 9	8	0.40	40
10 – 14	5	0.25	25
15 – 19	2	0.10	10
20 – 24	1	0.05	5
Total	20	1.00	100

"Cumulalrive Distributions".

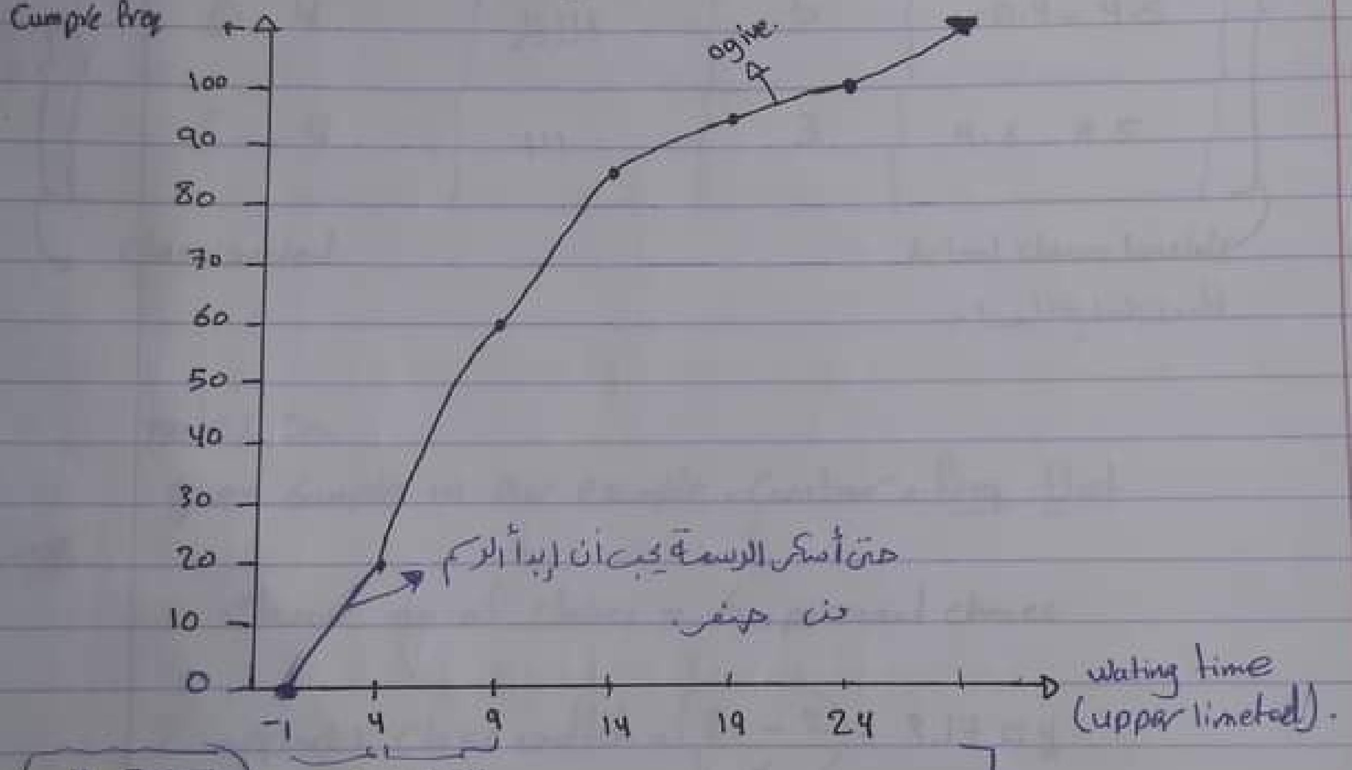
wating time	Cum freq	Cum real. freq $\frac{\text{Cum freq}}{20}$ =	Cum presnt freq.
time ≤ 4	4	0.20	20
time ≤ 9	4+8 = 12	0.60	60
time ≤ 14	4+8+5 → 17	0.85	85
time ≤ 19	4+8+5+2 → 19	0.95	95
time ≤ 24	4+8+5+2+1 → 20	1.00	100

→ Q10
→ Q11
→ Q12

Sample:-

2	5	10	12
4	4	5	17
11	8	9	8
12	21	6	8
7	13	18	3

Q13 → Constaict an Ogive? (used: cumpersant freq).



4−5 = −1

الفرق بينهم 5

في بعض الاحيان يجب ان أزيد على أرقام waiting time نص

فتصبح هذا الارقام هكذا :

[−0.5, 4.5, 9.5, 14.5, 19.5, 24.5]

وهم Actual upper limeted

note :-

Example :-

0–4	0–4	5–9	5–9	0–4	0–4	5–9	0–4
3.1	2.4	7.3	6.2	4.1	4.2	4.7	4.5

lowe class limeted – upper class limeted	Count	Freq	Actuall lower – Actuall upper
0 – 4	|||||	5	-0.4 – 4.5
5 – 9	|||	3	4.6 – 9.5

Class limeted.

Actual classes limeted

الحدود الفعلية للفئات .

notes :-

given Sample in our example . Constear a Freq . Dist

III

1- Choose no of classes = 6 → personal chooes.

2- apporex classe widtth = $\frac{21-2}{6} = 3.17 \approx \underline{4}$

حسب الاختلاف
الجبرية - الترميزية

Classe widtth = $\underline{4}$

3- clases =

$$\begin{array}{rcl} & 2-5 & \\ 4+ & 6-9 & +4 \\ +4 & 10-13 & +4 \\ 4+ & 14-17 & +4 \\ +4 & 18-21 & +4 \end{array}$$

* حسب أي طريقة من النماذج محددلي أحل أي سؤال .

*تلخيص شابتر (2.ع)

Quaulitative Date.

- Descriptive statislzis
  - Tablelenpreseatume
    - Freq. Dist
    - Cum Freq. Dist
  - Graplucal presetinb.
    - Historgams
    - Ogive.

Section 2.4

Crosstabulation and scatter Diagrams.

* Crosstabulation :-

* Example :- We want to study the relationship between age and H.I = helth insurance → تأمين الصحي

Raw DATA :-

	age	H.I
1	27	Yes
2	34	No
⋮	⋮	⋮
1200	45	yes

1200 Elements
1200 observation
Sample size = n = 1200

2 varaible → age, H.I

Age → Quatitave → Ratio
Helth insurance → Qulitative → nominal.

Freq. Dist. for H.I

H.I	freq.
Yes	1020
No	180
Total	1200

→ Raw Data.

Freq. Dist for age

Age	Freq.
18 – 34	552
35+	648
Total	1200

Cross Tabulation.

age \ H.I	yes	No	Total
18 – 34	~~1020~~ 450	102	552
35 +	570	78	648
Total	1020	180	1200

Joint frequencies.

marginal frequenies.

marginal Frequenies.

Sample size

نسب الأسطر

Def :- Row percentage = $\frac{\text{Row}}{\text{Row total}} \times 100$

age \ H.I	yes	No	Total
18 – 34	$\frac{450}{552} \times 100 = 81.52$	$\frac{102}{552} \times 100 = 18.48$	100
35+	$\frac{570}{648} \times 100 = 87.96$	$\frac{78}{648} \times 100 = 12.04$	100

→ Row percentage
[the larger the aga the higher the percentage of people with H.I]

Def :- Column percentage = $\frac{\text{Column}}{\text{Column total}} \times 100$

نسب الأعمدة

age \ H.I	yes	No
18 - 34	$\frac{450}{1020} \times 100 = 44.12$	$\frac{102}{180} \times 100 = 56.67$
35+	$\frac{570}{1020} \times 100 = 55.88$	$\frac{78}{180} \times 100 = 43.33$
total	100	100

→ Column percentage
[the higler the percentage of people with H.I, the lerger the age].
كلما زادت نسبة المؤمنين يتزيد نسبة صغار السن.

Conclusion 80 الاستنتاج :

The lerge the age, the higher the percentage of people with H.I, and vice versa

2 Variabelen (Crosstabulation) تلخيص ل *

TABULAR اكمابل

RAW DATA ↓ Cross tabulation.

- Presentations
- Descriptive statistics.

Cross tabulation:

- Row percentage ↓ Conclusion
- Column percentage ↓ Conclusion

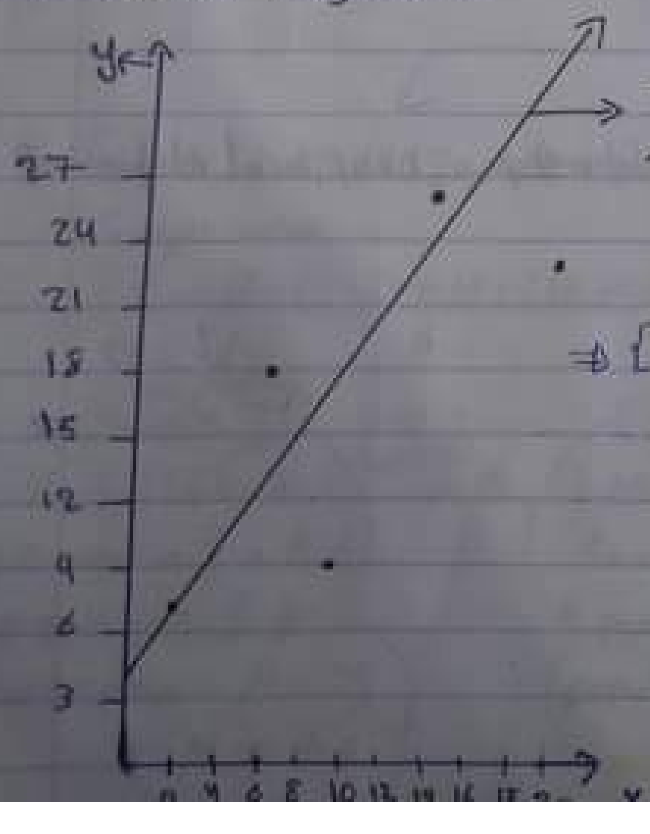
* Scatter Diagrams and Trendlines :-

→ Example :-

x	2	6	9	13	20
y	7	18	9	26	23

* 2 variable, x and y / * 5 observation
* 5 Elements / * sample size , n = 5

→ Scatter Diagrams :-

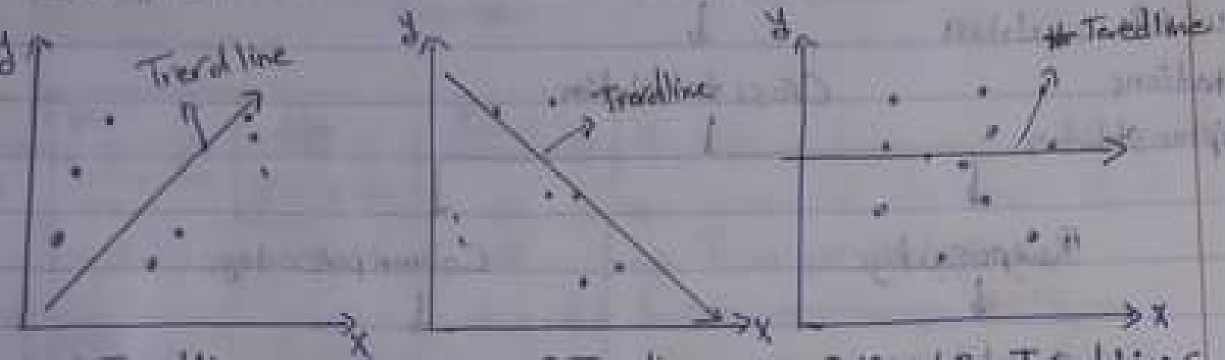


→ Trendline slope > 0
There is a positive liner relationship between x and y
⇒ { x↑ ⇒ y↑ }
كل ما زادت x تزيد y

x	y
2	7
6	18
9	9
13	26
20	23

"17"

* Trend line :-



DATASET 1: Trend line	DATASET 2: Trendline	DATASet 3:- Trend liner
Slope > 0	slope < 0	Slope = 0
→ positive linear relashionship between X and y	→ negative linear relationship between X and y	→ No Apparent relationship betwene x and y

*تلخيص (Scatter Diagroun)

DATA
↓
Scatter Diagroun
↓
Trendline.

+ve linear relationship | -ve linear relationship | no apparut relationship.

Chapter 3 :- Descriptive statistics :- numerical measures :- قياس العددي

مقياس وصفي

Section 3.1

measures of location :- مقياس الموقع

* Example :-

Sample 1 :- 4, 7, 12, 11, 9 ⇒ أعمار أشخاص في المدرسة

Sample 2 :- 18, 17, 20, 19, 21, 22 ⇒ أعمار أشخاص في الجامعة أو في الثانوية

* المقاييس اللي رح نتعامل فيها بالشابتر هي :-

1)- mean 2)- median 3)- mode 4)- percentiles 5)- Quartiles.

* Sample 1 :- 4, 7, 12, 11, 9 Find the sample mean :- أوجد الوسط الحسابي

Def :- Sample mean :- Sample : $x_1, x_2 \dots x_n$

Sample size : n

Sample mean :-

$$\bar{x} = \frac{\sum x_i}{n} \Longrightarrow \Rightarrow \frac{\text{مجموع القيم}}{\text{عددها}}$$

Solution :- Sample mean :-

$$\bar{x} = \frac{4+7+12+11+9}{5} = \frac{43}{5} = 8.6$$

* Sample 2 : 18, 17, 20, 19, 21, 22 Find the sample mean :-

Sample mean =

$$\bar{x} = \frac{18+17+20+19+21+22}{6} = \frac{117}{6} = 19.5$$

Sample 1

4, 7, 12, 11, 9

Sample mean = $\bar{x}$ = 8.6

population 1

Population mean :- μ = ??

$\mu \approx 8.6$

Replesal

Sample 2

18, 17, 20, 19, 21, 22

Sample mean = $\bar{x}$ = 19.5

population 2

population = μ = ??

$\mu \approx 19.5$

Statistical inference

Def ⇒ population mean

population : $x_1 \dots x_N$

population size = N

population mean = $\frac{\sum x_i}{N}$

Note:-

* $\bar{X}$ is called a sample statistic.

* μ is called a population parameter.

قيمة تخزيبية

* $\bar{X}$ is a point estimator for μ.

Note:-
if the sample represents the population (random) then we may say that $\mu \approx \bar{X}$
This is an example of statistical inference.

Def :- median :- The median is the value in the middle after we sort the data.
ترتيب

* القيمة إلى موجودة في الوسط ويجب ترتيب الارقام حتى أتمكن من البحث عن القيمة الصحيحة ويجب أخراج قيمة واحدة فقط في حال وجود قيمتين بأخذ الأمزيج بين القيمتين.

Example :-

Sample 1 :- 4, 7, 12, 11, 9 Find the median

$n = 5$ odd
عدد فردي يوجد قيمة واحدة فقط

Sloution:-
Sample 1 (sorted) :- 4, 7, 9, 11, 12
median = 9

Sample 2 :- 18, 17, 20, 19, 21, 22 Find the median.

$n = 6$ even
عدد زوجي يوجد قيمتين يجب أن أخذ الأمزيج بين القيمتين

Slootion:-
Sample 2 (sorted) :- 17, 18, 19, 20, 21, 22

$$\text{median} = \frac{19+20}{2} = \frac{39}{2} = 19.5$$

Sample 1
4, 7, 12, 11, 9
mean : $\bar{x} = 8.6$
median = 9

Sample 2
18, 17, 20, 19, 21, 22.
mean : $\bar{x} = 19.5$
median = 19.5

→ Sample 3 :-
4, 7, 12, 11, 9, 20 Find the mean and median.

mean :- $\bar{x} = \frac{4+7+12+11+9+20}{6} = \frac{63}{6} = \underline{\underline{10.5}}$

median :- Sample 3 (Sorted) :- 4, 7, 9, 11, 12, 20 (n = 6 even) عدد زوجي

$\text{median} = \frac{9+11}{2} = \frac{20}{2} = \underline{\underline{10}}$

→ Sample 4 :-
18, 17, 20, 19, 21, 22, 50 Find the mean and median.

mean = $\bar{x} = \frac{18+17+20+19+21+22+50}{7} = \frac{167}{7} = 23.86$.

median :- Sample 4 (sort) = 17, 18, 19, 20, 21, 22, 50
median = 20

* Note :-

* The mean is sensitive to extrem values.

* on the other hand, the median is robust to extreme values.

ملاحظة :- إذا تم زيادة قيمة كبيرة لم تتناسب مع الارقام الموجود تصبح هناك مشكلة كبيرة أو فجوة كبير بين الإجابات ويصبح هناك اختلاف كبير بين الارقام كم حدث معنا في (Sample 3 and Sample 4) لكن عندما قارنا الاجابات بـ (Sample 1 and sample 2) هناك فرق بين الاجابات كبير، ويجب أن نزيد أرقام في نسبة معقولة قريبة من الارقام الموجودة.

Def :- mode :- The most frequent value(s).

هي القيمة الأكثر تكراراً بين الأرقام ويمكن أن تكون رقم أو عدة أرقام أو يمكن أن لا يوجد قيمة.

Example :- Sample 1, 2, 3, and 4 Find the mode.

Sample 1 → No mode.
Sample 2 → No mode.
Sample 3 → No mode.
Sample 4 → No mode.

→ Sample 5 :- 4, 7, 12, 11, 9, 7, 8, 9, 9 Find the mode. $3 > 2$

mode = 9 uni-modal data :- قيمة واحدة تكرارها أكثر من الأرقام الباقية

→ Sample 6 :- 4, 7, 12, 11, 9, 7, 8, 9, 9, 7 Find the mode. $3 = 3$

modes = are 7 and 9 ← multi-modal data :- قيمتين تكرارهن نفس المرات

* mean
* median
* mode

→ measures of central location.

* Percentiles :-

Data set (not ordered)

order → min, ... → max

$\underbrace{\hspace{2em}}_{20\%}$ ≃ $\underbrace{\hspace{2em}}_{\%30}$

$\underbrace{\hspace{2em}}_{10\%}$ $\underbrace{\hspace{4em}}_{80\%}$

[illegible]

min, ... → max

$\underbrace{\hspace{4em}}_{= p\%}$ ↑ $\underbrace{\hspace{2em}}_{\simeq (100-p)\%}$

p^{th} percentiles

* Percentiles :- Sorted Data

min .... p^{th} percentiles .... max

$\underbrace{\hspace{4em}}_{\text{at least } p\%}$ $\underbrace{\hspace{4em}}_{\text{at least } = (100-p)\%}$

$$0 \leq p \leq 100$$

P_p : p^{th} percentiles

* Example :- Sorted Data [illegible]

[illegible]

min .... ↓ .... max

$\underbrace{\hspace{4em}}_{\text{at least } 70\%}$ $\underbrace{\hspace{2em}}_{\text{at least } 30\%}$

According to the def

P_{70} = 70^{th} percentiles.

(Roughlily speaking ----)

min .... P_{70} .... max.

$\underbrace{\hspace{6em}}_{\simeq 70\%}$ $\underbrace{\hspace{2em}}_{\simeq 30\%}$

Example :-

→ Simple 1 : 4, 7, 12, 11, 9 , Find 85^{th} percentile.

Solution :-

1. sorted data : 4, 7, 9, 11, 12.
(location) 2. index = $i = \frac{85}{100} \times 5 = 4.25$ Take 5 (roundup) * اذا الجواب أجى كسر بأخذ الرقم الي بعده.
3. $P_{85} = 12$

→ Simple 1 : 4, 7, 12, 11, 9 , Find 40^{th} percentile. ⇒ P_{40}

Solution :-

1. Sorted : 4, 7, 9, 11, 12
(location) 2. index = $i = \frac{40}{100} \times 5 = 2$ (Take 2 and 3) اذا أجى عدد صحيح بأخذ الموقع الثاني والثالث وبعد ذلك بأخذ المجموع ثم حتى يطلع معي P_{40}.
3. $P_{40} = \frac{7+9}{2} = 8$

* p^{th} Percentile (P_p) :-

* Sort the data set from smallest value to largest value.

* index = $i = \frac{p}{100} \times n$, n : sample size.

* 1. i not integer ⇒ P_p = the data value in the position i after rounding up.
(عدد غير صحيح)

2. i integer ⇒ P_p = average of data values in position i and next position.
(عدد صحيح)

* Example:

→ Sample 2 : 18, 17, 20, 19, 21, 22, Find 70th percentile, P_{70} = ??

1. sorted data :- 17, 18, 19, 20, 21, 22
2. index :- $i = \frac{70}{100} \times 6 = 4.2 \rightarrow$ (take 5)
3. $P_{70} = 21$

→ Sample 2: 18, 18, 17, 20, 21, 22, Find 50th percentile, P_{50} = ??

1. Sorted data = 17, 18, 19, 20, 21, 22
2. index : $i = \frac{50}{100} \times 6 = 3$
3. $P_{50} = \frac{19 + 20}{2} = 19.5$.

Sample 1	Sample 2
4, 7, 9, 11, 12.	17, 18, 19, 20, 21, 22.
→ $P_{85} = 12$.	→ $P_{70} = 21$
→ $P_{40} = 8$	→ $P_{50} = 19.5$

* Quartiles :- يقسموا البيانات الى ارباع

1^{st} Quartiles , $Q_1 = P_{25}$

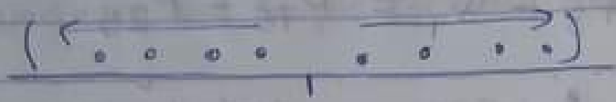
2^{nd} Quartiles : $Q_2 = P_{50} \rightarrow$ medien.

3^{rd} Quartiles : $Q_3 = P_{75}$

$$\text{min}\ \overbrace{\qquad\qquad}^{25\%}\ Q_1\ \overbrace{\qquad\qquad}^{25\%}\ Q_2\ \overbrace{\qquad\qquad}^{25\%}\ Q_3\ \overbrace{\qquad\qquad}^{25\%}\ \text{max}$$

Section 3.2
measures of variability

Data set 1 

Data set 2 

* Data set 2 has higher variability than Data set 1
الاتساع في البيانات المجموعة الثانية أكثر من بيانات المجموعة الأولى.

* Rang ✓
* IQR ✓
* Variance ✓
* Standard deviation ✓
* CV ✓

المقاييس إلي ح نتعاملهم في هذا الدرس هم :- → measures of variability.

→ **Def** Range:-

Range = largest value − smallest value.

Example: Sample 1 :- 4, 7, 12, 11, 9, Find the range.

Solution ⇒ Range = largest value − smaller value = 12 − 4 = 8

Example:- Sample 3: 4, 7, 12, 11, 9, 20, Find the range.

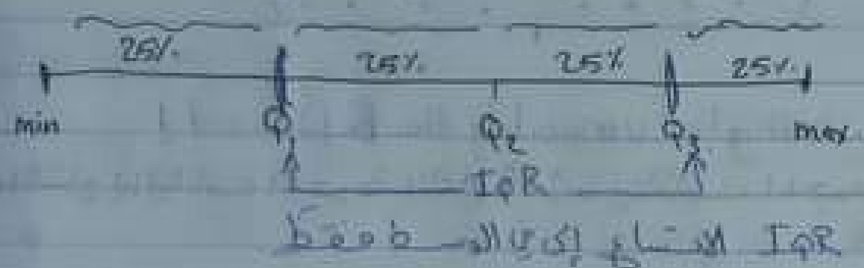
Solution ⇒ Range = largest value − smaller value = 20 − 4 = 16

Notes:
Range is sensitive (حساس) to extreme value.
رانج بيهمش كل البيانات.

* IQR

→ <u>Def</u>. IQR :- Interquartile Range.

$$IQR = Q_3 - Q_1 = P_{75} - P_{25}$$



Example:-

→ Sample 1 : 4, 7, 12, 11, 9, Find IQR

Solution:- Sorted Data = 4, 7, 9, 11, 12.

P_{25} → index = $\frac{25}{100} \times 5 = 1.25 \approx$ Take ② ⇒ $P_{25} = 7$

P_{75} → index = $\frac{75}{100} \times 5 = 3.75 \approx$ Take 4 ⇒ $P_{75} = 11$

$$IQR = Q_3 - Q_1 = P_{75} - P_{25} = 11 - 7 = 4.$$

→ Sample 3: 4, 7, 12, 11, 9, 20, Find the IQR.

Solution: Sorted Data = 4, 7, 9, 11, 12, 20

P_{25} → index = $\frac{25}{100} \times 6 = 1.5 \approx$ Take ② → $P_{25} = 7$

P_{75} → index = $\frac{75}{100} \times 6 = 4.5 \approx$ Take ⑤ → $P_{75} = 12$

$$IQR = 12 - 7 = 5$$

Sample 1	Sample 3
4, 7, 9, 11, 12	4, 7, 9, 11, 12, 20
• Range = 8	• Range = 16
• IQR = 4	• IQR = 5

→ Def :- Sample variance and sample standard deviation.

Data set (Sample) : $x_1, x_2, \ldots\ldots, x_n$
Sample size = n.

* Sample variance : $S^2 = \dfrac{\sum (x_i - \bar{x})^2}{n-1}$

* Sample standard deviation : $S = \sqrt{\dfrac{\sum (x - \bar{x})^2}{n-1}}$

→ Def :- population variance and population standard deviation Data sete (population) : $x_1, x_2, \ldots\ldots, x_N$

Population size : N

* Population variance : $\sigma^2 = \dfrac{\sum (x_i - \mu)^2}{N}$

* Population standard deviation : $\sigma = \sqrt{\dfrac{\sum (x_i - \mu)^2}{N}}$

note so

S : Sample standard deviation → Sample statistic.

σ : population standard deviation → population parameter.

S is a point estimator for σ.

Note :-

S^2 : Sample variance $\longrightarrow$ Sample statistic.

σ^2 : population variance $\longrightarrow$ poplation parameter.

S^2 is a point estimator for σ^2.

x **Def** :- CV

CV : Coeffecient of variation

$$CV = \left(\frac{\text{Standard deviation}}{\text{mean}} \right) \times 100\,\%$$

Example :-

Sample 1	sample 3
4, 7, 9, 11, 12	4, 7, 9, 11, 12, 20
Range = 8	Range = 16
IQR = 4	IQR = 5

* Find : Variance, st.dev, and CV for Sample 1 and Sample 3.

Recall :-

Sample :-	population :-
$S^2 = \frac{\sum (x_i - \bar{x})^2}{n-1}$	$\sigma^2 = \frac{\sum (x_i - \mu)^2}{N}$
$S = \sqrt{S^2}$	$\sigma = \sqrt{\sigma^2}$
$CV = \left(\frac{S}{\bar{x}} \right) \times 100\%$	$CV = \frac{\sigma}{\mu} \times 100\%$

Example:-

Sample 1 :- 4, 7, 9, 11, 12

Sample mean = $\bar{X} = \frac{\Sigma x_i}{n} = \frac{4+7+9+11+12}{5} = \frac{43}{5} = 8.6$

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

$$= \frac{(4-8.6)^2 + (7-8.6)^2 + (9-8.6)^2 + (11-8.6) + (12+8.6)^2}{5-1}$$

$$= \frac{41.2}{4} = 10.3$$

Sample st.dev = $S = \sqrt{S^2} = \sqrt{10.3} = 3.21$

Sample CV = $\frac{S}{\bar{X}} \times 100\% = \frac{3.21}{8.6} \times 100\% = 37.33\%$

Sample 3 :- 4, 7, 9, 11, 12, 20 → طريقة الحل على الآلة حاسبة

Sample mean = $\bar{X} = 10.5$

Sample variance $S^2 = 29.9$

Sample St.dev → $S = 5.47$

$CV = \frac{5.47}{10.5} \times 100\% = 52.10\%$

١- بكبس على زر [mode] وبعدها ع الرقم 2

٢- وبعدها أدخل البيانات وبعد كل عدد أضغط على M+

٣- وبعدها على AC.

٤- وبعدها ع [Shift] وبعدها ع الرقم 2.

٥- لإيجاد $\bar{X}$ بكبس على الرقم واحد وبعدها =.

٦- وبعدها نفس طريقة رقم ٤

٧- لإيجاد S بكبس على الرقم 3 وبعدها =.

٨- لإيجاد S^2 بكبس على Ans وبعدها تربيع =.

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

١٠- يجب أن أقرب لأقرب منزلتين عشريتين.

طريقة حل Sample على الآلة حاسبة.

Note :-

Sample 3 has more variability than Sample 1 Because $CV_3 > CV_1$.

* نفرض أنو المثال السابق (population) طريقة الحل على طريقة (population).

Example :-

population 1

4, 7, 9, 11, 12

median = 9

Rang = 8

IQR = 4

$\mu = 8.6$

$\sigma^2 = 8.24$

$\sigma = 2.87$

$CV = 33.37\%$

population 2

4, 7, 9, 11, 12, 20

median = 10

Range = 16

IQR = 5

$\mu = 10.5$

$\sigma^2 = 24.92$

$\sigma = 4.99$

$CV = 47.52\%$

طريقة حل population على الآلة حاسبة.

١- نكبس على زر (mode) وبعدها على الرقم ٢
٢. وبعدها أدخل البيانات وبعد كل رقم على زر M+
٣- وبعدها على AC.
٤- وبعدها (shift) على الرقم ٢
٥- لإيجاد μ نكبس على الرقم واحد وبعدها =.
٦- وبعدها نفس طريقة رقم ٤.
٧- لإيجاد σ نكبس على الرقم 2 وبعدها =
٨- لإيجاد (σ^2) نكبس على (Ans) وبعدها تربيع x^2.
٩- لإيجاد

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

١٠. دقرب لأقرب منزلتين عشريتين.

Section 3.3.

Measures of Dist. Shape, Relative location, and detecting outliers.

Distribution shape:

* Histogram.

* Compare mean vs. median.

mean > median ⇒ skewed to right.
mean = median ⇒ summetric.
mean < median ⇒ skewed to left.

This doesn't work for all cases!.

* Skweness :-

Def. skweness $= \frac{n}{(n-1)(n-2)} \sum \frac{(x_i-\bar{x})^3}{s^3}$

Skewness > 0 ⇒ postively skwed ⇒ skewed to right.
skewness = 0 ⇒ not skewed ⇒ summetric.
skewness < 0 ⇒ negatively skewed. ⇒ skewed to left.

It works for all cases.

* Z-score :-

Def:- Z-score for a data value x is defined as :-

$$\boxed{Z = \frac{X - \text{mean}}{\text{st. dev}}}$$

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

population: $Z = \frac{X-\mu}{\sigma}$.

Example:-

1- Tawjehi grades $\mu = 78$ / $\sigma = 11$, some one scored 97. Find the Z-score?

$$Z = \frac{x - \mu}{\sigma} = \frac{97 - 78}{11} = 1.73.$$

2. Tawjehi grades $\mu = 78$ / $\sigma = 11$, some one scored 65, Find the Z-score?

$$Z = \frac{x - \mu}{\sigma} = \frac{65 - 78}{11} = -1,18.$$

Example:-

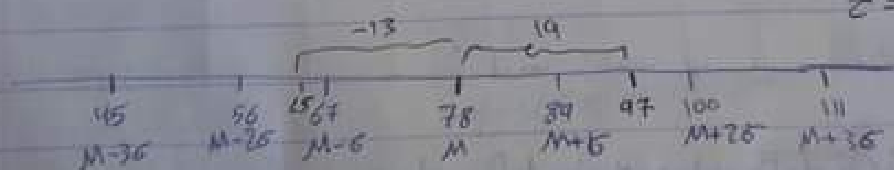
High schooles Exam (out of 20). ; $\mu = 13$ / $\sigma = 1.5$, some one 17 out of 20
Find the Z-score.

$$Z = \frac{x - \mu}{\sigma} = \frac{17 - 13}{1.5} = 2.67$$

EX:- $\mu = 78$, $\sigma = 11$

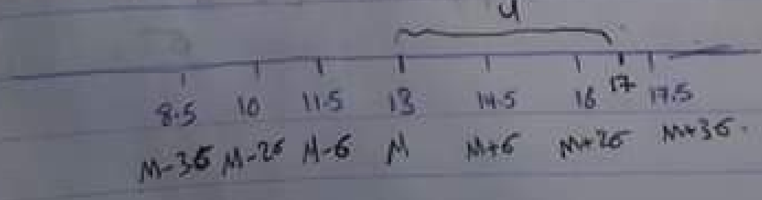
$X = 97$
$Z = 1.73$

$X = 65$
$Z = -1.18$

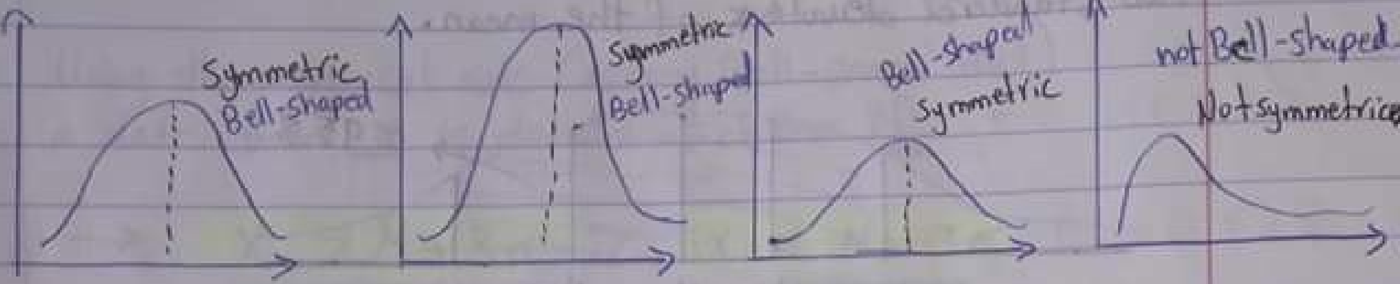


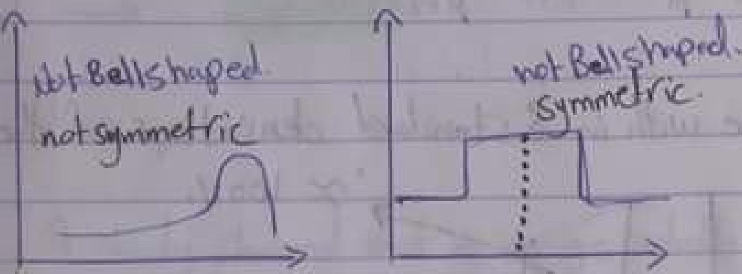
EX:- $\mu = 13$ $\sigma = 1.5$

$X = 17$
$Z = 2.67$



Bell-shaped Distribution :-





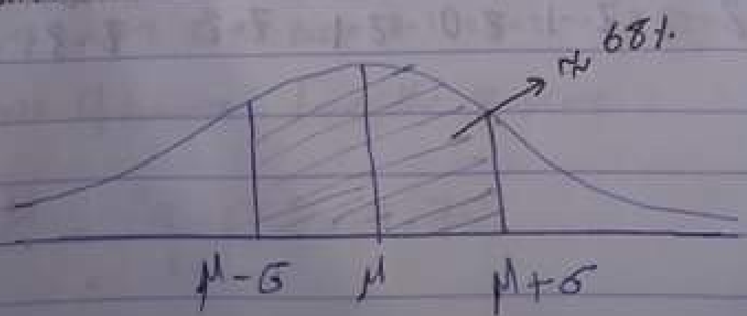
Distribution

- symmetric
  - Bell-shaped
  - Not Bell-shaped.
- not symmetric.

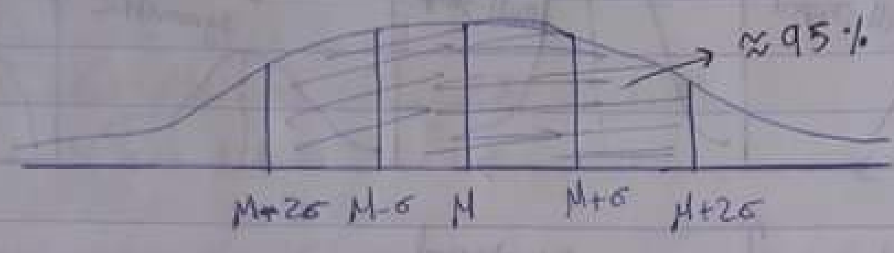
* Emperical Rule:-

for data having a bell-shaped Distribution: -

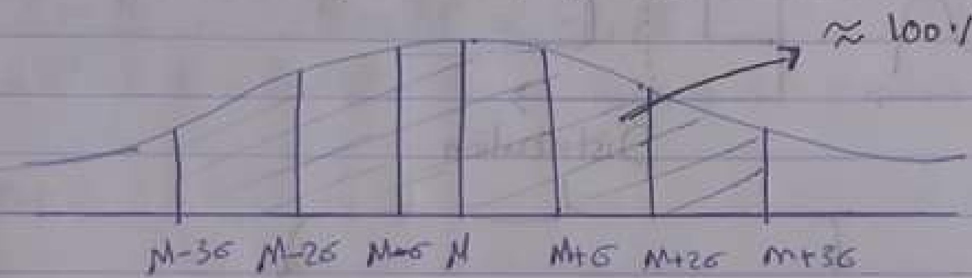
1)- Approximately 68% of the data will be withein one standard deviation of the mean.



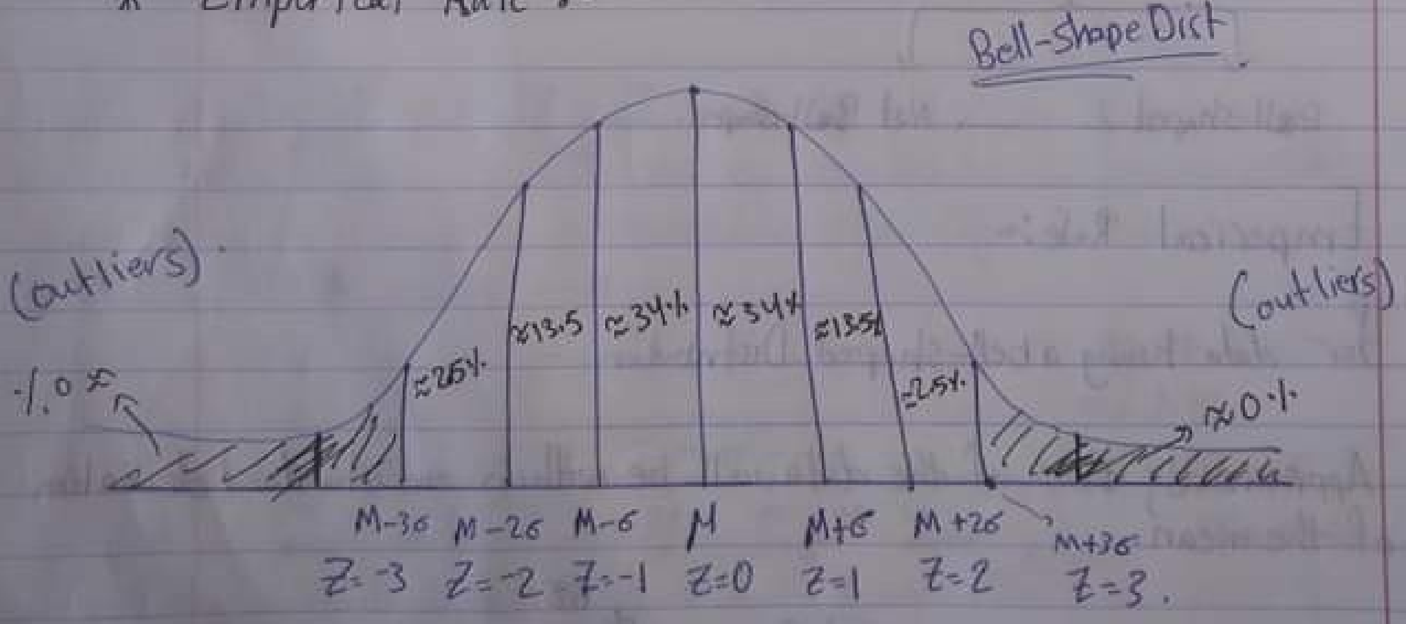
2). Approximately 95% of the data well be with in two standerd deuations of the mean.



3). Almost all data will be with in 3 standard deniations of the mane.



* Emperical Rule :-



* How do we detect outliers ?

Using the empeuical rule (assumy Bell-shaped dist) a date value $\underline{x}$ is deteded as ontlher $\underline{if}$.

$\rightarrow$ $X > \mu + 3\sigma$ or $X < \mu - 3\sigma$.

$\rightarrow$ $Z > 3$ $Z < -3$

* Example :-

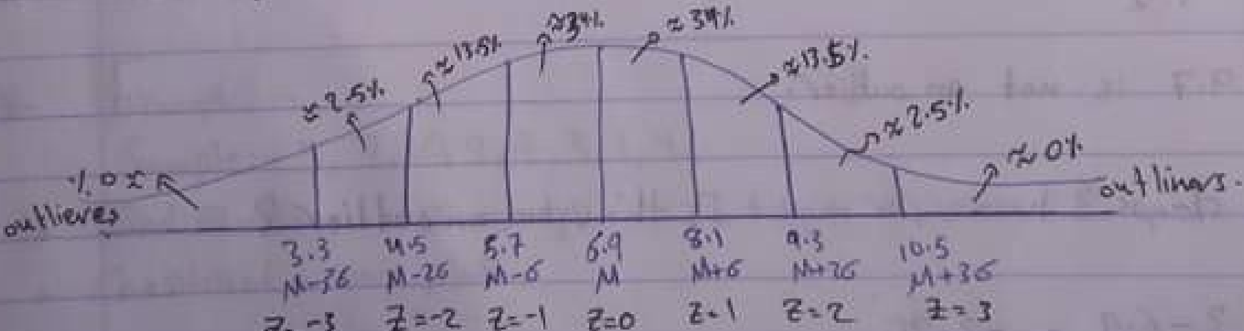
A study showed that on avarge adults sleep 6.9 hours per night. Suppoes that the St. dev is 1.2 hours.

Assume that the mean and St. dev are true for the population. Assume that the data bell shaped ?

$\bar{X} = 6.9 \longrightarrow \mu = 6.9$

$S = 1.2 \longrightarrow \sigma = 1.2$

Bull-Shaped $\Rightarrow$ $\mu = 6.9$ / $\sigma = 1.2$.



1). A peron sleeps 9 houres. Find the Z-score.

$$Z(9) = \frac{X - \mu}{\sigma} = \frac{9 - 6.9}{1.2} = 1.75.$$

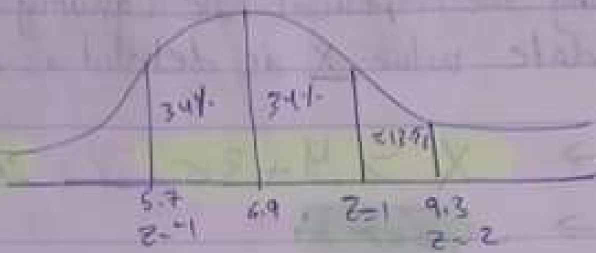
2). A person sleeps 4 houres. Find the Z-scores.

$$Z(4) = \frac{X - \mu}{\sigma} = \frac{4 - 6.9}{1.2} = 2.42.$$

3). what percentage of the people sleep between 5.7 houres to 9.3 houres ?

$$Z(5.7) = \frac{5.7 - 6.9}{1.2} = -1$$

$$Z(9.3) = \frac{9.3 - 6.9}{1.2} = +2$$

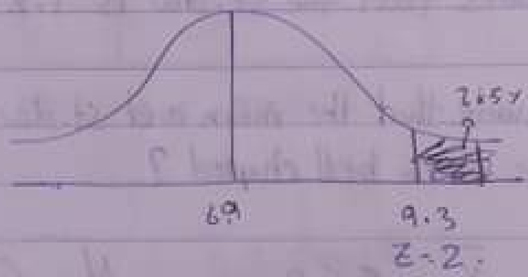


Percentage ≈ 34% + 34% + 13.5% = 81.5%

4). what percentage of people sleep more than 9.3 houres ?

$$Z(9.3) = \frac{9.3 - 6.9}{1.2} = 2$$

percentage = 2.5%



5). Assume some one sleeps 9.7 houres. It this value an outlier ?

$$Z(9.7) = \frac{9.7 - 9.6}{1.2} = 2.33$$

→ The value 9.7 is not an outlier.

6). Some one sleeps 3 houres per night Is this value an outlier ?

$$Z(3) = \frac{3 - 6.9}{1.2} = -3.25$$

→ The value 3 is an outlier.

Section 3.4

Exploratory Data Analysis.

* Five-number Summary : min, Q_1, Q_2, Q_3, max.

* Box plot.

* Box plot :-

1- Five-number summary.

2. $IQR = Q_3 - Q_1$

3. Upper limit of boxplot : $UL = Q_3 + 1.5\,IQR$.

4. Lower limit of box plot : $LL = Q_1 - 1.5\,IQR$.

5. outlier : $X > UL$ or $X < LL$

* Example :-

Sample:- 5, 7, 10, 9, 7, 3, 1, 4.

- Find Five-number Summary.
- Construct a box plot.

Solution :- Sorted Sample :- 1, 3, 4, 5, 7, 7, 9, 10

* Five number Summary :- 1 (min), 3.5 ($\frac{3+4}{2}$), 6 ($\frac{5+7}{2}$), 8 ($\frac{7+9}{2}$), 10 (max)

$\Rightarrow$

• box plot :-

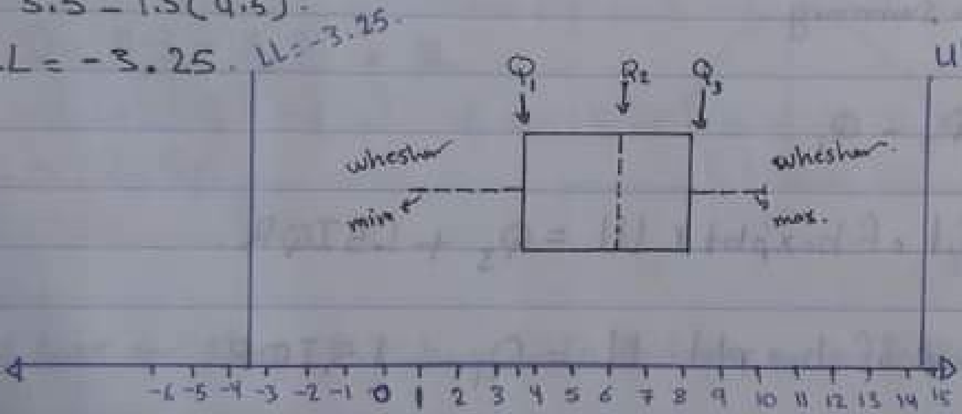
← Example :-

ordered sample :- 1, 3, 4, 5, 7, 7, 9, 10
Five-number Summary :- 1, $\underline{3.5}$, 6, $\underline{8}$, 10

$IQR = Q_3 - Q_1$
$= 8 - 3.5 = 4.5$

$UL = Q_3 + 1.5(IQR)$.
$UL = 8 + 1.5(4.5)$.
$UL = 14.75$.

$LL = Q_1 - 1.5(IQR)$.
$LL = 3.5 - 1.5(4.5)$.
$LL = -3.25$.



* Example :-

Sample :- −4, 1, 3, 4, 5, 7, 7, 9, 10, 20

1- Find five-number summery 2- Constorct - box plot.

Soution :-

Stared Sample :- −4, 1, 3, 4, 5, 7, 7, 9, 10, 20.

$Q_1 = P_{25} = \frac{25}{100} \times 10 = 2.5$

$Q_2 = P_{50} = \frac{50}{100} \times 10 = 5 \; \frac{(5+7)}{2}$

$Q_3 = P_{75} = \frac{75}{100} \times 10 = \underline{9}$

1- Five-number summery :- −4, 3, 6, 9, 20.

2- $IQR = Q_3 - Q_1$

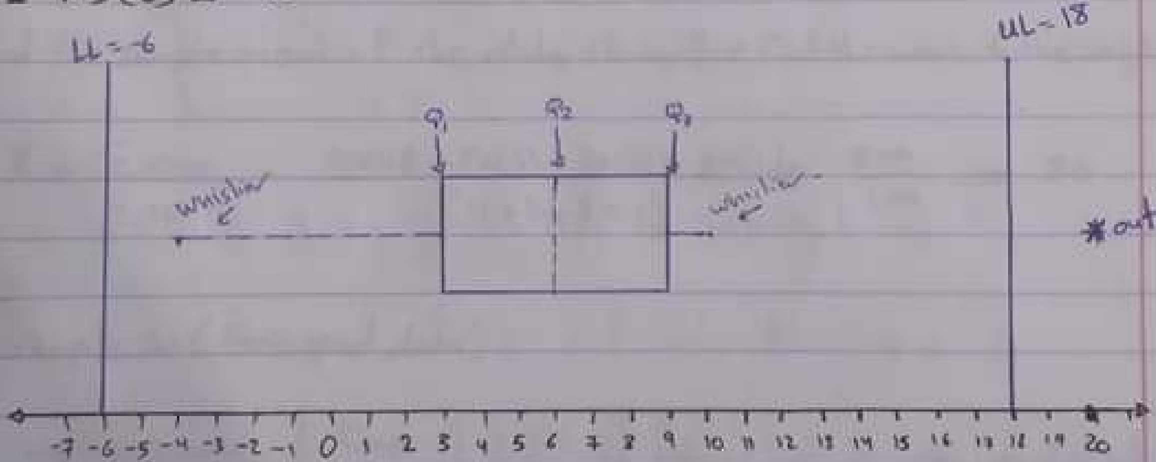
$IQR = 9 - 3 = 6.$

$UL = Q_3 + 1.5(IQR) =$

$UL = 9 + 1.5(6) = 18$

$LL = Q_1 - 1.5(IQR) =$

$LL = 3 - 1.5(6) = -6$



Sactione 3.6

The weighted mean and working with grouped data.

* Example :- (weighted data).

x_i : data value.
w_i : weight.

x_i	w_i
90	4
80	3
70	2
60	1

* Sample mean (without weights).

$$\bar{x} = \frac{\sum x_i}{n}.$$

* Sample mean of weighted data.

$$\bar{x} = \frac{\sum x_i w_i}{\sum w_i}.$$

1)- Find the sample mean assuming the values x_i have no weights?

$$\bar{x} = \frac{\sum x_i}{n} = \frac{90+80+70+60}{4} = 75.$$

2)- Find the sample mean of the data shown? or Find the mean of the weighted data?

$$\bar{x} = \frac{\sum x_i w_i}{\sum w_i} = \frac{90(4)+80(3)+70(2)+60(1)}{(4+3+2+1)} = \frac{800}{100} = 80.$$

* Example :- (Grouped data).

waiting time	freq.
0 – 4	4
5 – 9	8
10 – 14	5
15 – 19	2
20 – 24	1
Total	20

M_i	f_i
2	4
7	8
12	5
17	2
22	1
Total	20.

m_i : midpoint of class

$$M_i = \frac{UL_i + LL_i}{2}.$$

* Sample mean of grouped d

$$\bar{x} = \frac{\sum M_i f_i}{\sum f_i} = \frac{\sum M_i}{n}$$

Calculate:-

$\bar{x} = 9$ $\quad$ $M = 9$

$S^2 = 30$ $\quad$ $\sigma^2 = 28.5$

$S = 5.48$ $\quad$ $\sigma = 5.34$.

* قوانين . "Roul"

* Sample variance of grouped data :-

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

* Sample standand devielion of grouped data :-

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

* population mean of grouped data $= \mu = \frac{\sum M_i f_i}{\sum f_i} = \frac{\sum M_i F_i}{N}$

* population variance of grouped data $= \sigma^2 = \frac{\sum (M_i - \mu)^2 f_i}{\sum f_i} = \frac{\sum (M_i - \mu)^2 f_i}{N}$

* population st. dev of grouped data $= \sigma = \sqrt{\frac{\sum (M_i - \mu)^2 f_i}{\sum f_i}} = \sqrt{\frac{\sum (M_i - \mu)^2 f_i}{N}}$.

$\bar{X}$ ponit eslimator for μ

S^2 ponit eslimator for σ^2

S ponit eslimator for σ

Sample statistue:

population parameters.

Section 3.5 (Measures of Association between two varaibles).
Section 12.2 (least - Squares Method).

3.5 :- Def :- Sample covariance.

→ Sample : $(x_1, y_1), \dots, (x_n, y_n)$.

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

⇒ Def :- population covariance :-

→ population :- $(x_1, y_1), \dots, (x_N, y_N)$

$$\rightarrow \sigma_{xy} = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{N}$$

* S_{xy} → Sample covariance
n → Sample size
S_{xy} Sample statistic

* σ_{xy} → population covariance
N → population size
σ_{xy} population paramelter.

S_{xy} is a point estimator for σ_{xy}

Example :-

x	y	$x-\bar{x}$	$y-\bar{y}$	$(x-\bar{x})(y-\bar{y})$
1	2	-1	-1	1
2	3	0	0	0
3	4	1	1	1
Total				2

$\bar{x} = 2$
$\bar{y} = 3$
$n = 3$

$$S_{xy} = \frac{2}{3-1} = \frac{2}{2} = 1.$$

$S_{xy} > 0$ ⇒ There is a positive liner relationship between x and y.
$S_{xy} = 0$ ⇒ " " no " " " " " "
$S_{xy} < 0$ ⇒ " " a negative " " " " " ".

"44"

Def:-

Sample correlation coeffcient

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

r_{xy} : Sample statistic

→ Def:-

→ population correlation coeffcient.

$$\rho_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$$

ρ_{xy} = population parameta.

r_{xy} is a point estimtor for ρ_{xy}.

Example:-

x	y
1	2
2	3
3	4

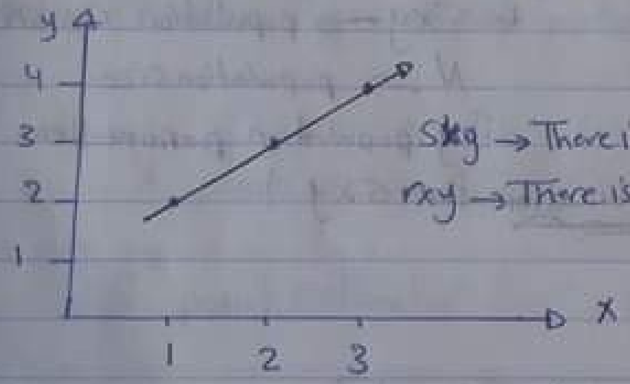
$\bar{x} = 2$ $S_x = 1$

$\bar{y} = 3$ $S_y = 1$

$n = 3$

$S_{xy} = 1$

$$* \; r_{xy} = \frac{S_{xy}}{S_x S_y} = \frac{1}{(1)(1)} = \boxed{1}$$



S_{xy} → There is a positive liner relationship between x and y

r_{xy} → There is a perfect positive liner relationship between x and y.

strong ← weak → ← weak strong →

strong negative	weak negtive	weak positive	Strong positive

-1 0 1 r_{xy}

perfect negative liner relationship.

no liner relationship.

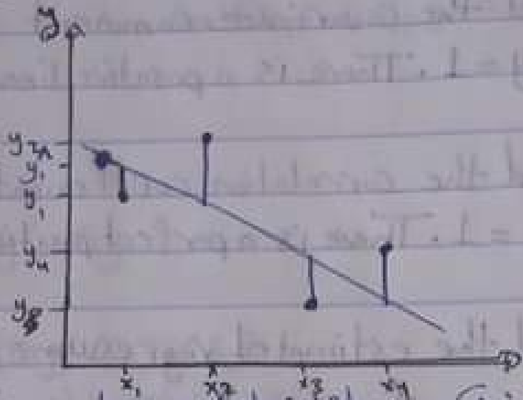
perfect positive liner relationship.

12.2 :- $(x_1, y_1), \ldots, (x_n, y_n)$.

$y_i \rightarrow$ observed value at x_i

$\hat{y}_i \rightarrow$ estimated value at x_i

$\hat{y}_i = b_0 + b_1 x_i$



least squares problem $\rightarrow \min \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$

b_0 : y intercept $\rightarrow$ (A)

b_1 : slope $\rightarrow$ (B)

* After solving the least-squares problemes :-

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

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

$\boxed{\hat{y} = b_0 + b_1 X} \longrightarrow$ Estimated Regression Equation.

* Example:-

x	y	$\hat{y}$
1	2	2
2	3	3
3	4	4
4	-	5

$\bar{x} = 2 \quad S_x = 1 \quad S_{xy} = 1$

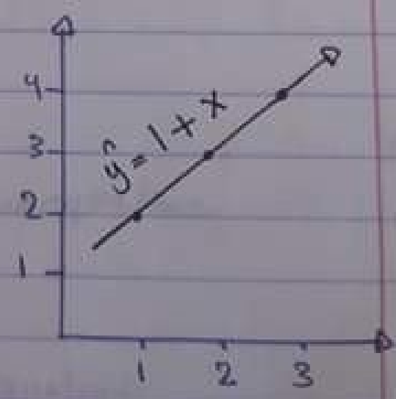
$\bar{y} = 3 \quad S_y = 1 \quad r_{xy} = 1$

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

$= 1 + 1(x)$

$= 1 + X$

$\hat{y} = 1 + X$



$b_1 = \frac{1}{(1)^2} = 1$

$b_0 = 3 - (1)(2) = 1$

1)- Find the covariance. comment.

$S_{xy} = 1$. There is a posative liner relactionshap between x and y.

2)- Find the correlation coeffeciant. comment.

$r_{xy} = 1$. There is a perfect positived liner relastonshap between x and y.

3)- Find the estimated regresum equation.

$\hat{y} = 1 + x$

4)- Find the y-intercept of the estimated regeuisel equation.

$b_0 = 1$

5)- Find the slope of the estimated regresued equations.

$b_1 = 1 \longrightarrow$ slope > 0

6)- Estimated y when $x = 4$

$\hat{y} = 1 + 4 = 5$

مهم *

Example :-

طريقة الحل باستخدام الآلة حاسبة :-

X	2	6	9	13	30
y	7	18	9	26	23

Find the :- $\bar{x}$, S_x, $\bar{y}$, S_y, S_{xy}, r_{xy}.

$$r_{xy} = \frac{S_{xy}}{S_x \cdot S_y} \Rightarrow S_{xy} = (r_{xy}) \cdot (S_x) \cdot (S_y)$$
$$S_{xy} = (0.65) \cdot (10.84) \cdot (8.38)$$

1. Find $\bar{x} = 12$.
2. Find $S_x = 10.84$.
3. Find $\bar{y} = 16.6$
4. Find $S_y = 8.38$
5. Find $S_{xy} = 59.05$ هذا الرقم مع تقريب → (58.75) هذا الرقم بدون تقريب
6. Find $r_{xy} = 0.65$
7. Comment on point 5.
   $S_{xy} = 59.05$ → There is a positive liner relastionshope betwen X and y.
8. Comment on point 6.
   $r_{xy} = 0.65$ → There is a moderate positaiv liener relation ship between x and y.
9. what is the estimated regrewsion eqatun. A = 10.6 / B = 0.5.

$$\hat{y} = b_0 + b_1 x = 10.6 + 0.5x.$$

10. Find the y-intercept of the estimatet regressiun equatione.
   $b_0 = 10.6$
11. Find the slope of the estimated regressium equatione.
   $b_1 = 0.5$.
12. Estimate y when $x = 25$.
   $\hat{y} = 10.6 + 0.5(25) = 23.1$.
13. Estimate y when $x = 9$.
   $\hat{y} = 10.6 + 0.5(9) = 15.1$.

A way to classify weak, moderate, strong relationship

* Chapter 4 :- Introduction to Probabilly :- علم الاحتمالات

Section 4.1

تجارب Experiments, counting rules, and assiging probabilitides.

الاحتمالات

تطبيق نواتج معرفة تعطي واضح تجربة عملية

Def: Experiment : process that generates well-defined out comes.

Probabilitec experiment → Possible out comes.

Experiment	Experiment out come	Sample space
Toss a Coin Roll a die	Head, Tail 1, 2, 3, 4, 5, 6.	$S = \{H, T\}$ $S = \{1, 2, 3, 4, 5, 6\}$.
play a game	Win, loss, Tie	$S = \{W, L, T\}$.
select a given part for inspection	Good, ... Defective.	$S = \{G, D\}$.

Def :-

Sample point :- is one of possible out comes.

Sample space :- is the set of all possible outcomes.

note: $\underbrace{S}_{\text{Sample space}} = \{\underbrace{H}_{\text{sample point}}, \underbrace{T}_{\text{Sample point}}\}$.

* Counting Rules:

عد قواعد تجربة متعددة الخطوات

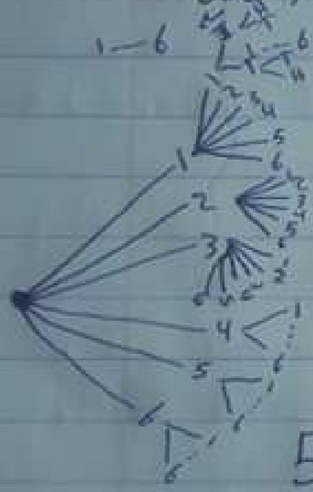
* Example 1 :- (Counting Rule of multistep Expermenet).

Experiment :- roll, die, twice and a coin 3 tiemes ?
write the Sample space.
How many elemantes are there in the Sample Space?

Roll	Roll	Toss	Toss	Toss
die	die	Coin	Coin	Coin
1-6	(1-6)	(H,T)	(H,T)	(H,T).

Tree diagram :

die , die , coin, Coin , coin



$S = \{ (1,1,H,H,H), (1,1,H,H,T), (1,1,H,T,H)$
$(1,1,H,T,T) \ldots\ldots, (6,6,T,T,T).$

Multi - step Experiment.

5 - step experimend:-

$\# S = (6)(6)(2)(2)(2) = 288.$

* K - step experiment Counting Rull

$\# S = (n_1)(n_2) \ldots\ldots (n_k).$

no. of possible out comes in step 1 | no. of possible out comes in step 2 | no. of possible outcome in step k.

* Example 2:-

we want to construct a password. The lenghn of the password is 7. we are allowed to use English letters (case sensitive) and numberes. How many passwordes can we construct?

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

A, B ... Z
a, b ... z
0, 1 ... 9

		حرف
A, B ..... Z	$\rightarrow$	26
a, b ..... z	$\rightarrow$	26
0, 1 9	$\rightarrow$	10
total		62 possible outcome.

$$\# \text{passwords} = (62)(62)(62)(62)(62)(62)(62) =$$
$$= 62^{\wedge}7$$
$$= 3.52 \times 10^{12}.$$

* طريقة الآلة حاسبة اذا طلع معي الرقم 3.52×10^{12} ← وبدي ياه رقم عادي بكبس.
[Shift $\longrightarrow$ ENG].

* Example 3.- (Counting Rule for combinations) ← قواعد العد للجمع .

you are going on a trip. you are allowed to take with you a limited number of the times. you have a total of 20 items. you are allowed to take only 4 item form the 20 items. How many possibilites do you have ?

?	?	?	?
20	19	18	17
	20-1	20-2	20-3

← بدي أختار ٤ أشياء من الأغراض إلي عددها ٢٠ وكل ما أختار غرض بنقص واحد لانو تم إختياره .

$$\# S = \frac{20 \times 19 \times 18 \times 17}{4 \times 3 \times 2 \times 1} = 4845$$

فكتوريال

المضروب

$4! = 4 \times 3 \times 2 \times 1 = 24$

بقسمة عليه عشان يكون ترتيب غير مهم

* Counting Rule for Combinations :-

You select n items for N items.

- The item is selected only once.
- The order is not important.

ترتيب

$$\text{Number of Combinations} = C^{N}_{n} = \binom{N}{n} = \frac{N!}{(N-n)!\, n!}$$

* Example 3 :- Soulation for (Counting Rule for Combinations).

$N = 20$, $n = 4$

$$\# S = C^{20}_{4} = \binom{20}{4} = \frac{20!}{16!\,4!} = \frac{20 \times 19 \times 18 \times 17 \times \cancel{(16 \times 15 \cdots 1)}}{\cancel{(16 \times 15 \cdots 1)}\,(4 \times 3 \times 2 \times 1)}$$

$$= \frac{20 \times 19 \times 18 \times 17}{4 \times 3 \times 2 \times 1} = 4845.$$

على طريقة الآلة حاسبة ← بوضع الرقم (N) وبعدها (nCr) وبعدها رقم (n) وبعدها يساوي

$\# S = C^{20}_{4} = 4845$ $\boxed{20}\ \boxed{nCr}\ \boxed{4} = 4845$ (✓)

* Example 4:-

We are interested in making a team of 3 members out of 25 members.

a - How many teams can we constuct ? The teammember can assign taeir roles later. ← order is not important. ترتيب هنا غير مهم

→ Combinations : $n = 3$, $N = 25$.

Number of teams $C^{25}_{3} = \binom{25}{3} = 25 \text{ ncr } 3 = 2300$.

b - How many team can we constrect ? The team should have a team leader, a finincial office, and a PR officer. ← order is important. ترتيب هنا مهم

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

25 24 23

$\# S = 25 \times 24 \times 23 = 138,00$

* Counting Rule for Permutations :-

Select n items from N item.

- The items is selected only once.
- The order is important.

Number of Permutations $= P^{N}_{n} = \dfrac{N!}{(N-n)!}$

Example 4:- Counting Rule for permutations ?

(b) $n = 3$, $N = 25$, Permutations.

$\# S = P^{25}_{3} = \dfrac{25!}{(25-3)!} = \dfrac{25!}{22!} = \dfrac{25 \times 24 \times 23 \times \cancel{(22 \times 21 \cdots 1)}}{\cancel{(22 \times 21 \times 20 \cdots 1)}} = 25 \times 24 \times 23 = 13800.$

* طريقة استعمال آلة حاسبة :- بنضغ رقم (N) وبعدها شفت + ncr وبعدها رقم (n)

$\# S = P^{25}_{3} = \boxed{25}\ \boxed{npr}\ \boxed{3} = 13800.$

Shaft $\boxed{ncr}$

.احتمال

* Rules for assigning Probabilites :-

Example :-

Experiment : Toss a coin

$S = \{H, T\}$.

* Assign probabiliteies for the Sample point ?

$P(H) = \frac{1}{2} = 0.5$

$P(T) = \frac{1}{2} = 0.5$

→ The classical metual for assiging probabites. Equally likely outumeres.

Example :- Experimant : Toss a Coin.

$S = \{H, T\}$. Assign prob. for the Smple points.

outcome	Freq	Relative Freq
H	7	0.35
T	13	0.65
Total	20	1.00.

$P(H) = \frac{7}{20} = 0.35$

$P(T) = \frac{13}{20} = 0.65$.

→ The relative freq. method for assigning pont.

Example :- Experiment : Toss a coin.

$S = \{H, T\}$. Assgin perb. for the sample ponite.

$P(H) = 0.49$

$P(T) = 0.51$

→ The subjetive method for assiging probabilites.

* Rules for assigning Porbabilites :-

$S = \{ E_1, E_2 \ldots\ldots E_n \}.$

S = Sample space
E_i = Sample points.

We want to assigning the probinbiltes, $P(E_i)$ for the Sample point E_i.

① $0 \leq P(E_i) \leq 1$

② $\sum_{i=1}^{n} P(E_i) = 1$

Section 4.2

Events and their probabilities.

Def:- Subset from the sample space is called an Event. حدث

Def:- P(Event) = Sum of probabilities of the sample points that are in the event.

Ex:- $S = \{E_1, E_2, E_3\}$.

$P(E_1) = 0.20$, $P(E_2) = 0.65$, $P(E_3) = 0.15$

$A = \{E_1, E_3\} \leftarrow$ Event

$P(A) = P(E_1) + P(E_3) = 0.20 + 0.15 = 0.35$.

$B = \{E_1, E_2\} \rightarrow$ Event.

$P(B) = P(E_1) + P(E_2) = 0.20 + 0.65 = 0.85$.

$C = \{E_1\} \rightarrow$ Event.

$P(C) = P(E_1) = 0.20$.

$D = \{E_1, E_2, E_3\} \longrightarrow$ Event.

$P(D) = P(S) = 1$

$M = \{\ \} = \phi \leftarrow$ Event.

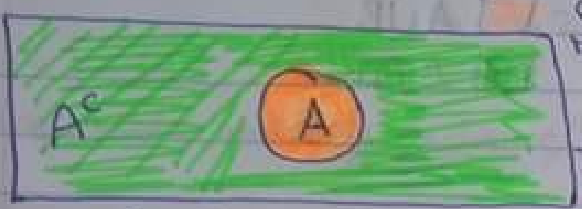
$P(M) = P(\phi) = 0$

Section 4.3

Some basic relationship of probability

* Complement law :-

S: Sample space
A: event
A^c: complement of event A.

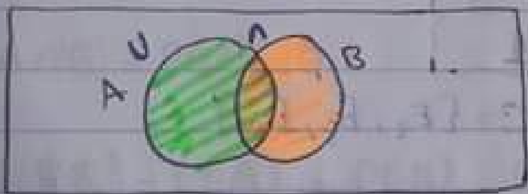


→ Venn diagram.

→ $A^c = \{ x \in S : x \notin A \}$.

$$\left.\begin{array}{l} \to P(A) + P(A^c) = 1 \\ \to P(A) = 1 - P(A^c) \\ \to P(A^c) = 1 - P(A) \end{array}\right\} \to \text{Complement law.}$$

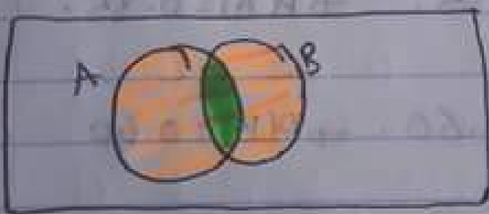
* The addition law



A
B

$A \cup B$: A union B.
$A \cap B$: A intersection B.

→ $A \cup B = \{ x \in S : x \in A$ or $x \in B \}$.
→ $A \cap B = \{ x \in S : x \in A$ and $x \in B \}$.



$A \cup B$:- الاتحاد يعني كل العناصر اللي في A و B
$A \cap B$: تقاطع العناصر المشتركة في A و B يعني بدون تكرار.

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

* Def:- A and B are called mutually exclusive event if:-

$$A \cap B = \{\} = \phi$$

Addition law for mutually exclusive events:-

$$\boxed{P(A \cup B) = P(A) + P(B)} \quad P(A \cap B) = P(\phi) = 0$$

* Example:- $S = \{E_1, E_2, E_3, E_4, E_5\}$.

E_i	E_1	E_2	E_3	E_4	E_5	Total
$P(E_i)$	0.05	0.25	0.10	0.25	0.35	1

Define:- $A = \{E_2, E_3\}$ / $B = \{E_4, E_5\}$ / $C = \{E_2, E_3, E_4\}$

1) $P(E_2) = 1 - (0.05 + 0.10 + 0.25 + 0.35)$

$1 - 0.75 = 0.25$.

2) $P(A) = P(E_2) + P(E_3) = 0.25 + 0.10 = 0.35$. $\Rightarrow P(A) = 0.35$.

3) $P(B) = P(E_4) + P(E_5) = 0.25 + 0.35 = 0.60$. $\Rightarrow P(B) = 0.60$

4). $P(C) = P(E_2) + P(E_3) + P(E_4) = 0.25 + 0.10 + 0.25 = 0.60$. $\Rightarrow P(C) = 0.60$.

5) $P(A^c) = 1 - P(A) = 1 - 0.35 = 0.65$.

) $P(B^c) = 1 - P(B) = 1 - 0.60 = 0.40$.

) $P(C^c) = 1 - P(c) = 1 - 0.60 = 0.40$.

$\Rightarrow$ تابع.

Example :- ← يتبع

8). $P(A \cap C) = P(\{E_2, E_3\}) = P(E_2) + P(E_3)$
$= 0.25 + 0.10 = 0.35.$

9). Are A and C mutually exclusive?

$A \cap C = \{E_2, E_3\} \neq \emptyset \Rightarrow$ A and C are not mutually exclusive.

10). $P(A \cup C) = P(A) + P(C) - P(A \cap C) = 0.35 + 0.60 - 0.35 = 0.60.$

11) Are A and B mutually exclusive events?

$A \cap B = \{\ \} = \emptyset \Rightarrow$ A and B are mutually exclusive.

12) $P(A \cap B) = P(\emptyset) = 0$

13) $P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.35 + 0.60 - 0 = \underline{0.95}.$

Section 4.4
Conclitional Probability.

Def :- let A and B be events form a Sample space S.

* The Conditional probalitaity of the event A given the event B is defned as :- $P(A/\underset{\text{Condition}}{B}) = \dfrac{P(A \cap B)}{P(B)} = P(B) \neq 0.$

* The conditional probalitaty of the event B given the event A is defnedas :- $P(B/\underset{\text{condition}}{A}) = \dfrac{P(A \wedge B)}{P(A)}, \quad P(A) \neq 0.$

مستقلين

Def. A and B are independiont events if :-

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

Def. of condition at prob...

multiplication law.

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \Rightarrow P(A \cap B) = P(A|B) \cdot P(B).$$

$$P(B|A) = \frac{P(A \cap B)}{P(A)} \Rightarrow P(A \cap B) = P(B/A) \cdot P(A).$$

* Multiplication law for independent events.

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

* Example :-

$S = \{E_1, E_2, E_3, E_4, E_5\}$.

$A = \{E_2, E_3\}$, $P(A) = 0.35$

$B = \{E_4, E_5\}$, $P(B) = 0.60$

$C = \{E_2, E_3, E_4\}$, $P(C) = 0.60$

14). $P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(\emptyset)}{P(B)} = \frac{0}{0.60} = 0$

15). Are A and B independent ?

$P(A|B) \neq P(A)$

$0 \neq 0.35$

$\Rightarrow$ A and B are not independent.

16). $P(B/C) = \frac{P(B \cap C)}{P(C)}$ $\quad P(B \cap C) = \{E_4\} = 0.25$

$= \frac{0.25}{0.60} = 0.4167 \rightarrow$ قربت لـ 4 منازل عشرية

17). Are B and C independent ?

$P(B|C) = P(B)$

$0.4167 \neq 0.60$.

$\Rightarrow$ B and C are not independant.

* Example :- Data form Q32 p168.

Healthe insurance :-

age	Yes	No	Total
18-34	750	170	920
& 35+	950	130	1080
Total	1700	300	2000

Cross tabulation.
"frequencies".

1)- Develop the joint probability table.

age	Yes	No	Tota
18-34	0.375 ↳ 750/2000 =	0.085	0.46
35+	0.475	0.065	0.54
Total	0.85	0.15	1.00

relative freq methodes.

Joint prob. Table
"probalilitzei".

2)- Find the joint probaleilites ?

$P(18\text{-}34 \cap Yes) = 0.375$
$P(18\text{-}34 \cap No) = 0.085$
$P(35^+ \cap yes) = 0.475$
$P(35^+ \cap No) = 0.065$

ملاحظة :-
في الاحتمالات بقرب لأقرب أربع منازل عشرية .

3)- Find the marginal probabilites :-

$P(18-34) = 0.46$
$P(35^+) = 0.54$
$P(yes) = 0.85$
$P(No) = 0.15$

4)- What percentege of the population go without health insurance?

$P(No) = 0.15 \Rightarrow$ percentage $\approx 15\%$

5). what is the percentage of people who are 18-34 years old in the population

$P(18-34) = 0.46 \Rightarrow$ percentage $\approx 46\%$

6)- What is the percentage of people who are 18-34 years old and are without halthe insuruns?

no ← without

$P(18-34 \overset{and}{\cap} No) = 0.085 \rightarrow$ percentage $\approx 8.5\%$

7)- A person was selected at ramdom what is the probability that the person is 35^+ years old and has health is insurance?

$P(35^+ \cap yes) = 0.475$

8)- what is the probabilite that a randomly selected person is 35^+ or doesn't have halthinsurance?

$$P(35^+ \overset{or}{\cup} No) = P(35^+) + P(No) - P(35^+ \cap No)$$
$$= 0.54 + 0.15 - 0.065$$
$$= 0.625$$

9)- A person (with health insurance) was selected at random, what is the probab- that he/she is 18-34 year old ?

→ yes

$$P(18\text{-}34 / \text{yes}) = \frac{P(18\text{-}34 \cap \text{yes})}{P(\text{yes})} = \frac{0.375}{0.85} = 0.4412$$

10) If a person who is 35^+ years old is selected what is the prop. that he/she has health insurance ?

$$P(\text{yes} / 35^+) = \frac{P(\text{yes} \cap 35^+)}{P(35^+)} = \frac{0.475}{0.54} = 0.8796$$

11) are the event yes and 35^+ indepentent ? why?

$P(\text{Yes} / 35^+) = 0.8796.$

$P(\text{yes}) = 0.85$

$P(\text{yes} / 35^+) \neq P(\text{yes})$

⇒ yes are 35^+ are not independent.

Another Solution :-

1# $P(35^+ / \text{yes}) = \frac{0.475}{0.85} = 0.5588$

$P(35^+) = 0.54$ ⇒ $P(35^+ / \text{yes}) \neq P(35^+)$

→ 35^+ and yes are not indepent.

2# $P(\text{yes}) = 0.85$ $P(35^+) = 0.54$

$P(\text{yes}) \cdot P(35^+) = 0.459.$

$\neq P(\text{yes} \cap 35^+) = 0.475$

yes and 35^+ are not indepentent.

12). Are the variables age and health insurentes indedent ?

35+ and yes are not indepentant.

→ age and H.I are not indepantant.

Chapter 5 :- Discrete Probability Distributions

section 5.1

Random Variable

Example :-

Experiment : Toss a coin
Sample Space : $S = \{H, T\}$.

$P(H) = \frac{1}{2} = 0.5 \quad P(T) = \frac{1}{2} = 0.5$

Define :- X : number of heads.

$H \rightarrow x = 1$ قياس رقمي.
$T \rightarrow x = 0$
$P(H) = 0.5 \longrightarrow P(X = 1) = 0.5$
$P(T) = 0.5 \longrightarrow P(X = 0) = 0.5$

قياس رقمي ←   عدد ←

Def :- A random variable is a numerical measure (description) of the out come (outcomes) of an experiment.

* Example :-
Experiment : Toss a coin twice.

$S = \{(H,H), (H,T), (T,H), (T,T)\}$.
$\quad x=2 \quad x=1 \quad x=1 \quad x=0$

Define :-
X : no. of heads.
possible values of X :-
$(0, 1, 2)$

Def :- we say that the random variable x is discrete if the possible values of X are countable. * العناصر قابلين للعد.

Def :- we say that the random variable x is Continuose if the possible values of X are uncountable. * العناصر غير قابلين للعد.

* Example :- شحنة تلفونات ← تفحص
Experiment :- Inspect a shipment of 100 phones.
random variable X : number of defelive phones

* possible value of X :- $0, 1, 2, 3, 4, 5 \ldots\ldots, 100$ ← بقدر أسرد الأرقام

⇒ X is a discrete random variable.

* Example :-
Experiment : fill a soft drink can.
random varaible X : amoumt of soft drink filled in the can.

* possible value of X :- $0 \leq X \leq 330$
أي رقم ممكن يكون بين 0 و 330 ∴ الرقم غير قابل للحصر.

→ X is a Continuous random varable.

Section 5.2

Discrete Probability Distrubations.

* Example:-

Experiment :- Toss a coin twice. (independently).

$S = \{(H,H), (H,T), (T,H), (T,T)\}$.

$x=2 \quad x=1 \quad x=1 \quad x=0$

$P(H,H) = 0.25$ $(\frac{1}{2} \times \frac{1}{2})$, $P(H,T) = 0.25$, $P(T,H) = 0.25$, $P(T,T) = 0.25$

X : number of heads (random variable), possible value of X : 0, 1, 2

$P(X=0) = 0.25$, $P(X=1) = 0.25 + 0.25 = 0.50$, $P(X=2) = 0.25$.

X	P(X)
0	0.25
1	0.50
2	0.25
Total	1.00

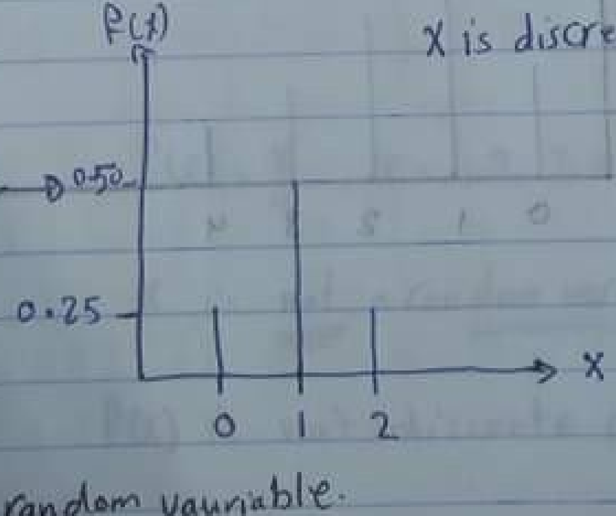
← Probability Distribution.

probability function

$$f(x) = \begin{cases} 0.25 & , x = 0, 2 \\ 0.50 & , x = 1 \\ 0 & , \text{else} \end{cases}$$

X is discrete but not uniform.

Graphical Presentation of f(x)



X : discrete random variable.

f(x) : discrete probability functions

we have a discrete prob. distribution.

* propention of Discrete Prob. function :-

1) $f(x) \geq 0$ for all x.

2) $\sum_{x} f(x) = 1$

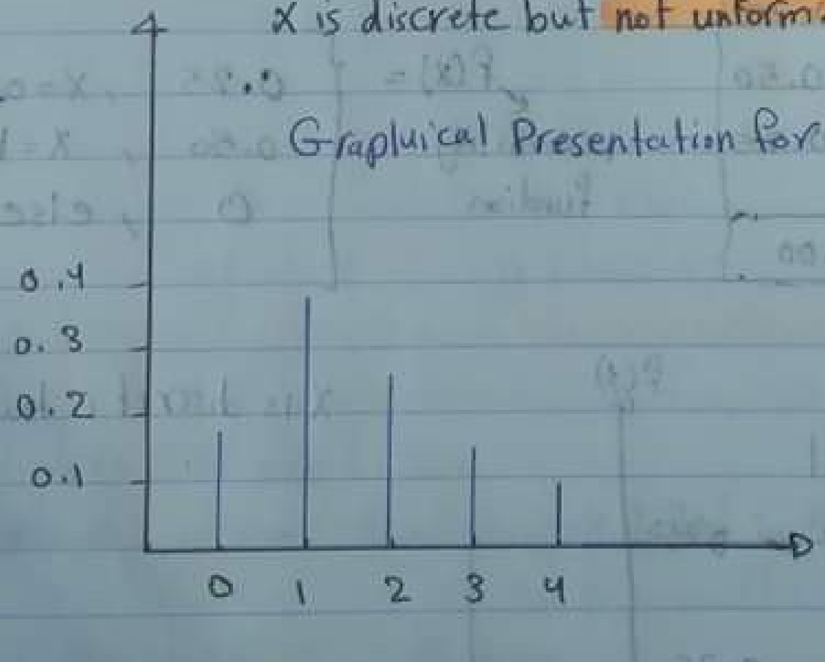
* Example :- Gevien

x	$f(x)$
0	0.18
1	0.38
2	0.25
3	0.13
4	0.06
Total	1.00

* possible values of X : 0, 1, 2, 3, 4 : X is a discrete randome varable.
$f(x) \geq 0$ for all x, $\sum_{x} f(x) = 1$

$\therefore$ $f(x)$ is a discrete prob. distrbutions.
X is discrete but not unform.

Graplical Presentation for $f(X)$:



* Example:- Given.

X	$f(x)$
1	1/6
2	1/6
3	1/6
4	1/6
5	1/6
6	1/6
Total	1

* possible values of X :- $1, 2, 3, 4, 5, 6$.

→ X is a discrete random varable.

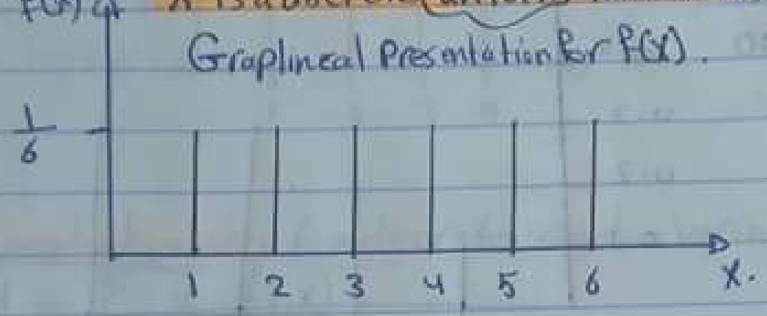
$f(x) \geq 0$ for all x, $\sum_x f(x) = 1$

$f(x)$ is a discrete prop. function.

لو كل الاحتمالات متساوية

X is a discrete unform random variables.

Graphical Presentation for $f(x)$.



Def:- we say that X is a discrete unform random varable and $f(x)$ is a discrete unform prob. function if:

$f(x) = \frac{1}{n}$, n: number of possible values of the random varable X

Example

x	$f(x)$
1	0.25
2	0.50
3	0.75
4	1.00
Total	2.50

$f(x) = \frac{x}{4}$, $x = 1, 2, 3, 4$.

X is not a random varable

$f(x)$ is not discrete prob. function

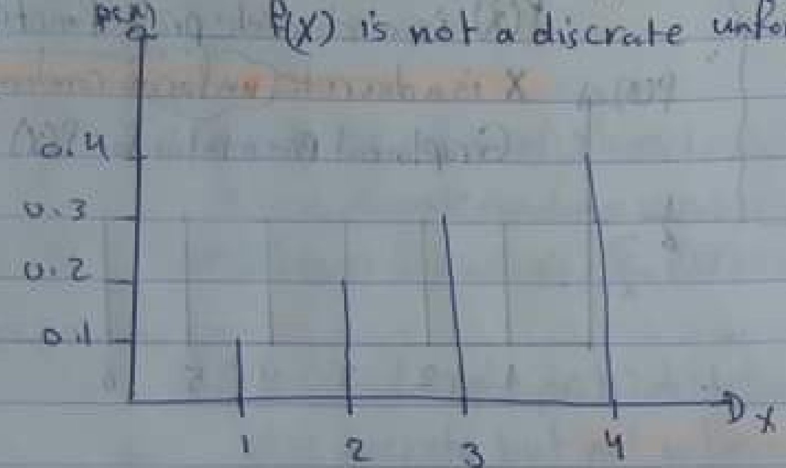
Example:- $f(x) = \frac{x}{10}$, $x = 1, 2, 3, 4$.

x	$f(x)$
1	0.1
2	0.2
3	0.3
4	0.4
Total	1.00

$X \rightarrow$ is a discrate randome vatuber

$f(x) \rightarrow$ is a discrate prob. function.

$f(x)$ is not a discrate unform prob. fuction

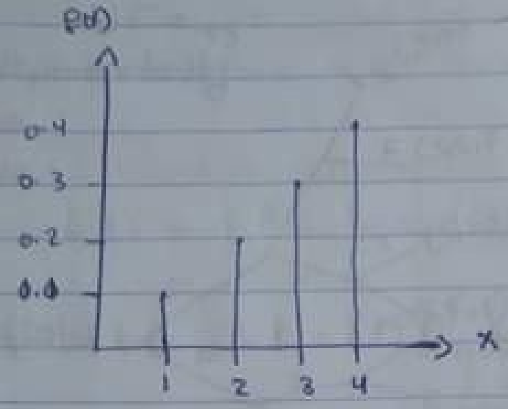


Section 5-3

Expected value and variance

Example:-

x	f(x)
1	0.1
2	0.2
3	0.3
4	0.4
Total	1.0



Def:- The expected value of a discrete random variable X

$E(X) = \mu = \sum_x x \cdot f(x)$

Def:- The variance of a discrete random variable X

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

Standard deviation $= \sigma = \sqrt{Var(x)}$.

X	f(x)	x. f(x)	$(x-3)$ $(x-\mu)$	$(x-\mu)^2$	$(x-\mu)^2 \cdot f(x)$
1	0.1	0.1	-2	4	0.4
2	0.2	0.4	-1	1	0.2
3	0.3	0.9	0	0	0
4	0.4	1.6	1	1	0.4
Total	1.0	3	—	—	1.00

1). Expeted value of x:-

$E(x) = \mu = 3$

2). Variance of X:-

$Var(x) = \sigma^2 = 1$

3). St. dev. of x:-

$\sigma = \sqrt{1} = 1$

Example:- gavine.

X	f(x)
0	0.18
1	0.39
2	0.24
3	0.14
4	0.04
5	0.01
Tota	1.00

X → discriete randon varaiable. <u>but not</u> discrete uniform.

f(X) → discriete prob. function but not discarte unform.

* طريقة الحل على الآلة حاسبة أسهل.

Find:-

1)- Expected value of X:-

$E(x) = \mu = 1.5$

2) varaince of X:-

$Var(x) = \sigma^2 = 1.25$

) st. dev of X:-

$\sigma = 1.12$

Section 5.4

توزيع الاحتمالي ذي الحدين . Binomial Probability Distribution.

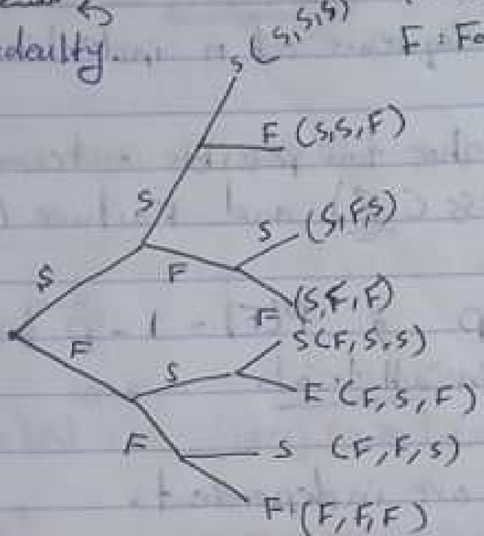
Example:-

Investment 3 times independendaulty. ← مستقل.

احتمال النجاح P(success) = 0.60.

احتمال الفشل P(Failure) = 0.40.

S : Successes.
F : Failure.



outcome	x	prob.		prob.	
(S,S,S)	3	(0.60)(0.60)(0.60)	$= (0.60)^3(0.40)^0$	= 0.216	ما تكررت
(S,S,F)	2	(0.60)(0.60)(0.40)	$= (0.60)^2(0.40)^1$	= 0.144	مكررات
(S,F,S)	2	(0.60).(0.40).(0.60)	$= (0.60)^2(0.40)^1$	= 0.144	مكررات
(S,F,F)	1	(0.60).(0.40).(0.40)	$= (0.60)^1(0.40)^2$	= 0.096	مكرر
(F,S,S)	2	(0.40).(0.60).(0.60)	$= (0.60)^2(0.40)^1$	= 0.144	مكررات
(F,S,F)	1	(0.40).(0.60)(0.40)	$= (0.60)^1(0.40)^2$	= 0.096	مكرر
(F,F,S)	1	(0.40).(0.40).(0.60)	$= (0.60)^1(0.40)^2$	= 0.096	مكرر
(F,F,F)	0	(0.40).(0.40).(0.40)	$= (0.60)^0(0.40)^3$	= 0.064	ما تكررا

Sum = 1.000.

Def:- random variable of x:-

x : number of time the investment has "success".

* possible value of x:- [0, 1, 2, 3]

*

x	f(x)
0	0.064
1	(0.096 + 0.096 + 0.096) = 0.288
2	(0.144 + 0.144 + 0.144) = 0.432
3	0.216
Total	1.000.

Def:- Binomial Experiment is an experiment such that :-
عبارة عن تجربة

* 1). It has a sequence of n indentical trials.

* 2). Each trial has tow possible outcomes :- Success (S) and Failure (F).

* 3). $P(S) = p$, $P(F) = 1 - p$, both probablites don't change in all trial.

* 4). The trieles are independent.

Def:- The binomial random variables X
X : number of successe in a binomial experimenets.

Not: possible value of $X : 0, 1, 2, 3, \ldots n$

→ Binomial prob. function :-

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

→ Expected value of binomial random variable :-

$$E(X) = \mu = n \cdot p$$

→ Variance of binomial random variable :-

$$Var(X) = \sigma^2 = n \cdot p \cdot (1-p).$$

→ Stander deviation of binomial random varaiable :-

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

* Example :-

It's known that 35% of basket ball players are left handed. In a Sample of 10 players, what is the probability that exactly 1 players is the left handed ?

احتمال

1- $n = 10$ 2- S :- left-handed 3- $P(S) = p = 0.35$ 4- players independent.

X : number of left hand players.

Find :- $F(x) = C^n_x \, p^x (1-p)^{n-x}$

$f(1) = C^{10}_1 (0.35)^1 (1-0.35)^{10-1}$

*طريقة الحساب على الآلة حاسبة :-

$(10C1) \cdot (0.35)^1 \cdot (0.65)^9 =$

$f(1) = C^{10}_1 (0.35)^1 (0.65)^9 = 0.072491694 = 0.0725$

* Example :-

We are making an investment 3 independent times. The prob. that the investment will make profit is 0.60. Find the prob. of :-

1- no profit 2- profit in one investment 3- profit in 2 investments 4- profit in all inv

* $n = 3$ * S : profit * F : no profit * $P(S) = 0.60$ / $P(F) = 0.40$ * investment inde

X : number of investment that make profit.

* حل السؤال على الآلة حاسبة أسهل :-

$f(x) = C^n_x \, p^x (1-p)^{n-x}$

1- $f(0) = C^3_0 (0.60)^0 (0.4)^3 = 0.064$

2- $f(1) = C^3_1 (0.6)^1 (0.4)^2 = 0.288$

3- $f(2) = C^3_2 (0.6)^2 (0.4)^1 = 0.432$

4- $f(3) = C^3_3 (0.6)^3 (0.4)^0 = 0.216$

* Example:-
It's know that 35% of basket ball players are left-handed. In a Sample of 10 players, what is the prob. that more than one players is the left handed?

1- $n = 10$ 2- $p = 0.35$ 3- X: number of left hand players. (→ bonomial Random variable)

$x: 0, 1, 2, 3 \ldots\ldots 10.$

$P(x > 1) = P(2) + P(3) + P(4) + P(5) + \ldots\ldots + P(10)$ أختصار كل هدول

$= 1 - (P(0) + P(1))$. بقانون المتممة.

$= 1 - C_0^{10}(0.35)^0(0.65)^{10} - C_1^{10}(0.35)^1(0.65)^9$

بآخذ ٤ منازل عشرية الفاصلة $= 0.914045561 \approx 0.9140$

* Example:-
In a Sample of 20 players, what is the expected number of left-handed player? It's know that 35% of basket ball players are left handed?

$n = 20$, $p = 0.35$, x: number of left-handed players.

Find the Expected value of Bonomial:

$E(x) = \mu = n \cdot p = (20) \cdot (0.35) = 7$ players.

* Find the variances and st. dev. for the no. of the left

$Var(x) = \sigma^2 = np(1-p) = (20) \cdot (0.35) \cdot (0.65) = 4.55$ players.

$\sigma = \sqrt{np(1-p)} = \sqrt{(20)(0.35)(0.65)} = 2.13$ players

* طريقة استخدام الآلة حاسبة لتمارين (3.5 and 12.2)

x	2	6	9	13	30
y	7	18	9	26	23

Find the :-

رموز الكلمات

Sxy → Sample covariance.
rxy → Sample correlation coeffcation.
Sx → Sample santines (x)
Sy → Sample santines (y).

1 → mode ↔ $\boxed{3}$ → REG
2 → nembuer $\boxed{1}$ → Lin
3 → ندخل البيانات → الأولى جنب M+ ويضع بينهم الفاصلة اي جنب وهكذا باقي البيانات وبعدها نضغط
4 → Shift → $\boxed{2}$
5 → Find $\bar{x}$ → Shift → $\boxed{2}$ → number $\boxed{1}$ → = $\bar{x} = 12$
6 → Find Sx → Sheft → $\boxed{2}$ → number $\boxed{3}$ → = Sx = 10.84 (Sx $\xrightarrow{\text{رمزها}}$ $x6n-1$).
7 → Find $\bar{y}$ → shift → $\boxed{2}$ → وبعدها على سهم → number $\boxed{1}$ = $\bar{y} = 16.6$.
8 → Find Sy → shift → $\boxed{2}$ → وبعدها سهم → number $\boxed{3}$ = Sy = 8.38 (Sy $\xrightarrow{\text{رمزها}}$ $y6n-1$).
9 → ~~Find Sxy~~ shift → $\boxed{2}$ وبعدها بكبس مرتين على → A, B, R. نحصل على
10 → Find rxy → shift → $\boxed{2}$ وبعدها على سهم بكبس مرتين → number $\boxed{3}$ → Rxy = 0.65
11 → Sxy → الآلة حاسبة ما تعطيني ياه جاهزة انا احسبها

$$Rxy = \frac{Sxy}{Sx \cdot Sy} \Rightarrow Sxy = (rxy) \cdot (Sx) \cdot (Sy).$$

$$Sxy = (0.65)(10.84) \cdot (8.38)$$

$$Sxy = 59.05$$ إذا بدي الرقم بشكل أدق بضرب الأرقام مباشرة على الآلة

12 → Sxy → بشكل أدق → shift → $\boxed{2}$ → number $\boxed{3}$ → (x) → shift → $\boxed{2}$ → number $\boxed{3}$ (x) shift → $\boxed{2}$ → $\boxed{3}$
Sxy = $x6n-1 \times y6n-1 \times r$ = 58.75 → بشكل دقيق (بدون تقريب)
13 → Find the estimated regression equatoun ⇒ $\hat{y} = b_0 + b_1 x$ = $\hat{y} = 10.6 + 0.5x$
لهالمعادلة كاملة على الآلة حاسبة موجودة برموز → A + B
→ shift → $\boxed{2}$ → على السهم مرتين بختار → A وبعدها B → A → numbr $\boxed{1}$ = 10.6 / B → number $\boxed{2}$ = 0.5
14 → Find the y-intercept of the estimated regssion equatonn = b_0 ↔ A = 10.6.
المقطع العادي
15 → Find the slope of the estamated regession equatonn = b_1 ↔ B = 0.5.
الميل
16 → Eestimat y when x = 25 ↔ $\hat{y}(25) = 10.6 + 0.5(25) = 23.1$.

علي

طريقة استخدام الآلة حاسبة لشابتر 4

4.1

* إذا طلع معي الرقم هكذا (3.52×10^{12}) بتحوله لرقم عادي بكبس على ← ENG ← بعدها shift

* Counting Rule of Combinations = $C_n^N = \binom{N}{n} = \dfrac{N!}{(N-n)!\,n!}$ رقم ← n / رقم ← N

على الآلة حاسبة بوضع N → $\boxed{nCr}$ → n =

20 → $\boxed{nCr}$ → 2 = 4845 مثل

→ [order is not important.] في حال الترتيب غير مهم

* Counting Rule for Permutations = $P_n^N = \dfrac{N!}{(N-n)!}$ → رقم ← N / رقم ← n

على الآلة حاسبة بوضع ← N → shift → nCr → n =

25 → shift → nCr → 3 = 13800 مثل

(shift + nCr) = P

→ [order is important.] في حال ترتيب مهم

طريقة استخدام الآلة حاسبة لشابتر 5

5.3

X	0	1	2	3	4	5	Total
P(x)	0.18	0.39	0.24	0.14	0.04	0.01	1

1_ mode → $\boxed{2}$ → M+ بيطلع خانة قيم X وبعدين كل قيمة

2 → بنصعد على السهم إلى فوق الوسط وبنكبس على السهم (اليمين) وبعدها بنكبس مرتين وبنغير القيم كل شي
تبعت كل قيمة كما هو موضح في الجدول وبعدها بنحط المعطيات.

3 → Find → Expected value → $E(x) = (\mu)$ → shift → $\boxed{2}$ → number(1) → $\bar{x}$

4 → Find → standard deviation → σ → shift → $\boxed{2}$ → number(2) → $x\sigma n$.

5 → Find → varaince → σ^2 → shift → $\boxed{2}$ → number(2) وبعدها تربيع → x^2 → $(x\sigma n)^2$.

Capter 6 :- Continuous probaplity Distrobution :. توزيع الاحتمال المستمر

Section 6.1

Continuous → Uniform Probaplity Distrobution - توزيع الاحتمال المنتظم

* Example: A car is traveling from one city to another. The travel time is 120 min to 140 min. It's knowe that the car can arrived at any moment.

1). what is the random variabl of interest? Is it countinues or discrete?

X: Car's travel time in minut → $(120 \leq X \leq 140)$

→ X Continuous random variables.

Discrete :-	Continuous
$f(x)$ = prob. function	$f(x)$ = prob. density function.
• $f(x) \geq 0$	• $f(x) \geq 0$
• $\sum f(x) = 1$	• $\int f(x)\,dx = 1$ = area under the curve of $f(x)$
→ $P(X=a) = f(a)$	→ $P(X=a) = \int_a^a f(x)\,dx = 0$
→ $P(X=b) = f(b)$	→ $P(X=b) = \int_b^b f(x)\,dx = 0$
→ $P(a \leq X \leq b) = \sum_{x=a}^{b} f(x)$.	→ $P(a \leq X \leq b) = \int_a^b f(x)\,dx$ — area under $f(x)$ from a to b.

2). Find the prob. that the car takes 130 min to travel from the first city to the second city?

X = travel time → Continuous.

$P(X=130) = 0$

3). What is the prob. that the car take less than 130 min to complete the trip?

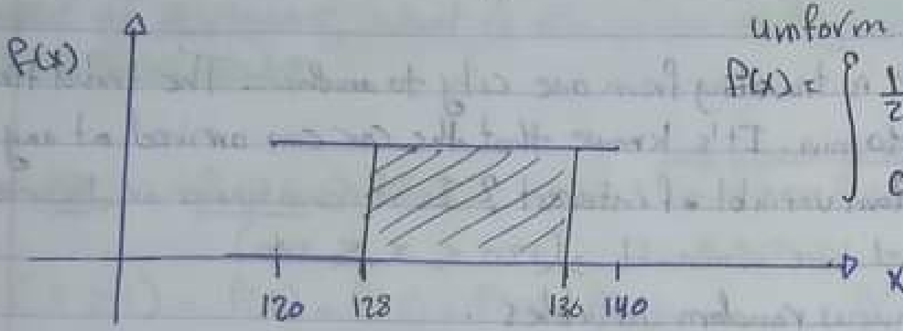
X: travel time, Continuous. uniform, continuous

$\int f(x)\,dx = 1$ — area under the curve of $f(x)$ is 1

$P(X < 130) = \underbrace{(10)}_{\text{الطول}} \times \left(\frac{1}{20}\right) = 0.5$

$f(x) = \begin{cases} \frac{1}{20}, & 120 \leq x \leq 140 \end{cases}$

4). Find the prob. that trip time is between 128 and 136 minten ?



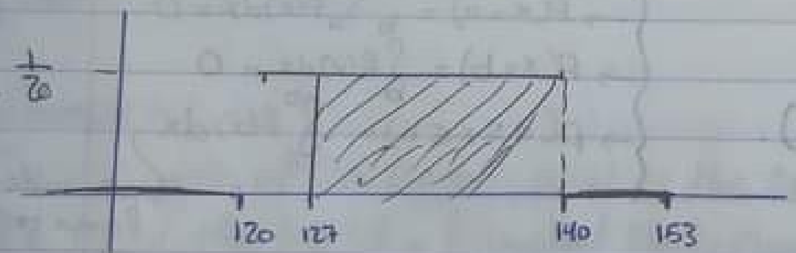
Continuoes uniform $120 \le x \le 140$

$$f(x) = \begin{cases} \frac{1}{20}, & 120 \le x \le 140 \\ 0, & \text{else.} \end{cases}$$

$P(128 < x < 136)$ = are under the curve frome $x = 128$ to $x = 136$

$$= \frac{1}{20}(136 - 128) = \frac{8}{20} = 0.40$$

5). Find the prob. that trip time is between 127 and 153 minutes ?

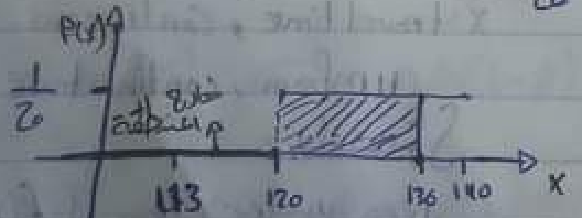


$$f(x) = \begin{cases} \frac{1}{20}, & 120 \le x \le 140 \\ 0, & \text{else.} \end{cases}$$

$$P(127 < x < 153) = \frac{1}{20}(140 - 127) + 0(153 - 140) = \frac{13}{20} = 0.65.$$

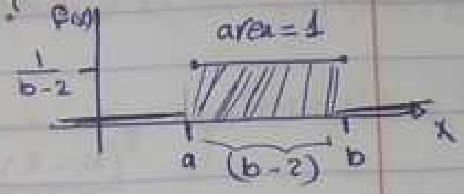
6). $P(x = 127) = 0$

7). $P(113 \le x \le 136) = \frac{1}{20}(136 - 120) = \frac{1}{20}(16) = 0.8$



8). $P(113 < x < 136) = 0.8$

* Continuoseus fumfrom prob. distraibution?



random variable :- $a \leqslant x \leqslant b$

prob. dusiting fundum : $f(x) \begin{cases} \frac{1}{b-a}, & a \leqslant x \leqslant b \\ 0, & \text{else.} \end{cases}$

Expected value of cont. unform randomverable = $\mu = E(x) = \frac{a+b}{2}$

Variaunes of cont. unform randome varable: $\sigma^2 = Var(x) = \frac{(b-a)^2}{12}$

Bache example?

9). what is the expected travel time?

$\mu = E(x) = \frac{a+b}{2} = \frac{120+140}{2} = 130$ minuten.

10). what is the varianece of the travel time?

$\sigma^2 = Var(x) = \frac{(b-a)^2}{12} = \frac{(140-120)^2}{12} = \frac{400}{12} = 33.33 \text{ min}^2$

11). what is the st. dev of the traveal time?

$\sigma = \sqrt{Var(x)} = \sqrt{\frac{400}{12}} = 5.77$ mimuntes.

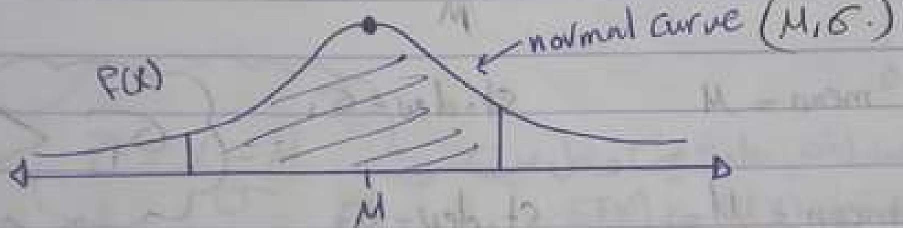
Section 6.2

Normal Probability Distribution :- توزيع الطبيعي

* Normal curve :- مش مهم حفظ المعادلة

Prob. density function :- $f(x) = \frac{1}{\sqrt{2\pi}\cdot\sigma} \cdot e^{\frac{-(x-\mu)^2}{2\sigma^2}}$, $-\infty < x < \infty$

x :- normal random variable with mean μ and st. dev. σ.



* propetien of normal dist. :- خصائص

1- Every normal distribiution is characterized by two parmeters μ and σ.

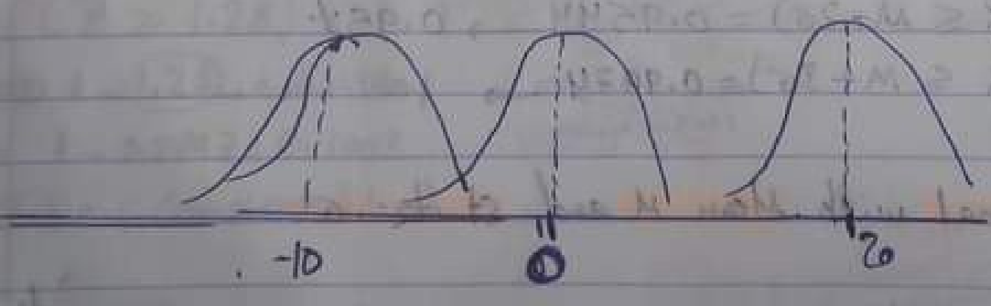
2- The highest point on the normal curve is obtained at the mean.

μ = mean = median = mode.

In this case we have :- بس في حالة كلها متساوية

3. The mean μ could be negative, zero, positive.

(μ measures the location).



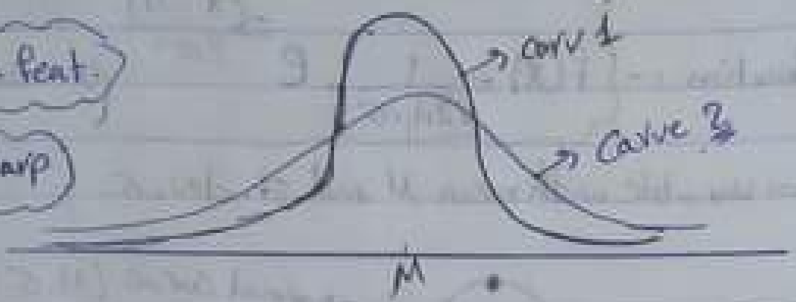
4. The normal curve is symmetric $\boxed{\text{Skewness} = 0}$

⇒ * propeitien of normal dist. تكملة الخصائص

5). The st. dev. σ measure the variability, where $\sigma > 0$

Carv 2 is more flat.

Carv 1 is more sharp

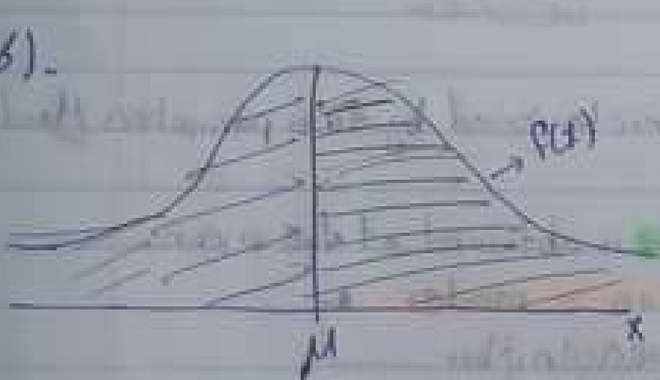


Carve 1: mean = μ st. dev = σ_1

Carve 2 :- mean = μ st. dev = σ_2

$\sigma_1 < \sigma_2$

6).



Total area under the curve = 1

$P(X < \mu) = 0.5$] symmetry.
$P(X > \mu) = 0.5$

$P(X = \mu) = 0$] cont. dist

$P(X \leq \mu) = 0.5$

$P(X \geq \mu) =$

7). $P(\mu - \sigma \leq X \leq \mu + \sigma) = 0.6826 \rightarrow 0.68\%$

$P(\mu - 2\sigma \leq X \leq \mu + 2\sigma) = 0.9544 \rightarrow 0.95\%$

$P(\mu - 3\sigma \leq X \leq \mu + 3\sigma) = 0.9974 \rightarrow 100\%$

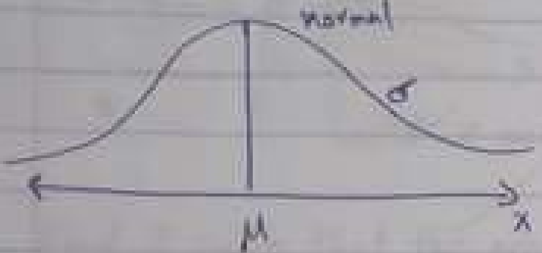
Not :- X is normal with Mean μ and st. dev. σ

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

Z : is standerd normal → Table.

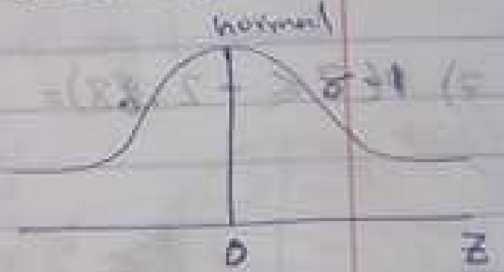
X : normal random variable with mean μ and st.dev. σ.

$E(X) = \mu$, $Var = \sigma^2$



Z : standard normal random variable.

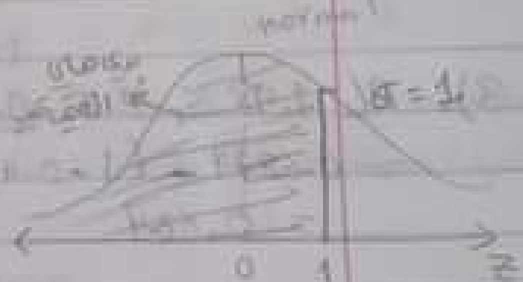
$E(Z) = 0$, $Var(Z) = 1$



$$Z = \frac{X - \mu}{\sigma}$$ → هاي المعادلة هي حلقة الوصل بين الجهتين $E(X) \Leftrightarrow E(Z)$.

Example:-

let Z be a standard normal random variable.



1- $P(Z \leq 1) \Rightarrow 0.8413$ ← حسب الجدول طلعت من الجدول

2- $P(Z \leq 1.5) = 0.9332$ (Table)

3- $P(Z < 0.67) = 0.7486$ (Table)

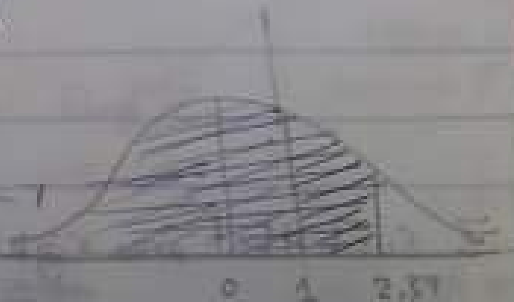
4- $P(Z \geq 1.28) =$

$\Rightarrow 1 - (1.28)$ المنطقة المطلوبة

لكن الجدول بعطيني هاي المنطقة مباشرة

$= 1 - 0.8997 = 0.1003$.

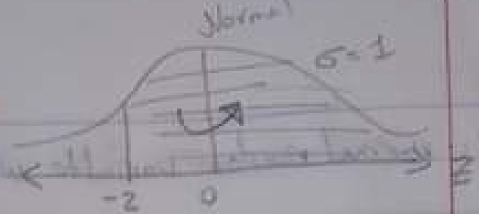
(Table + Complement law).

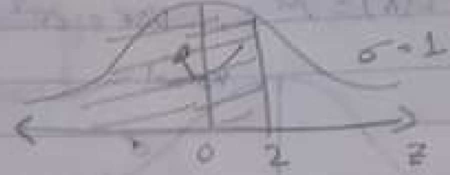


5) $P(1 < Z \leq 2.37) =$

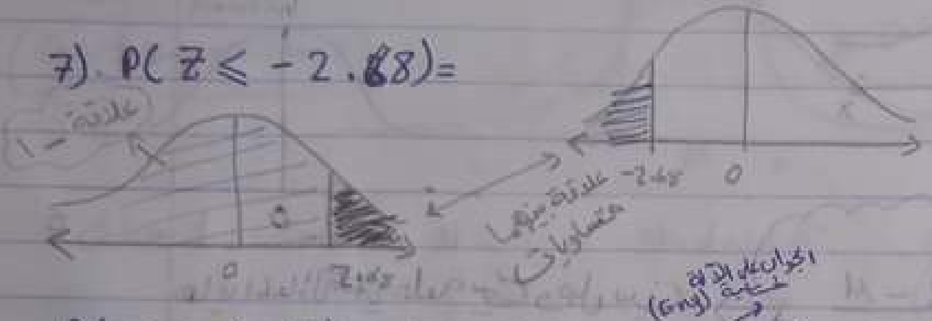
$0.9904 - 0.8413 = 0.1491$.

6)- $P(Z \geq -2) = 0.9772$.





7). $P(Z \leq -2.68) =$

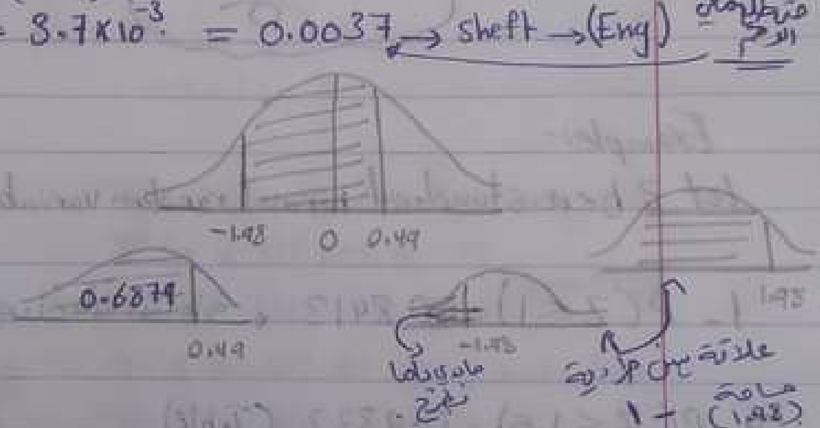


$P(Z \leq -2.68) = 1 - 0.9963 = 3.7 \times 10^{-3} = 0.0037 \rightarrow$ shift $\rightarrow$ (Eng)

8)- $P(-1.98 \leq Z \leq \underline{0.49})$ (Table) $=$

$= 0.6879 - (1 - 0.9761) =$

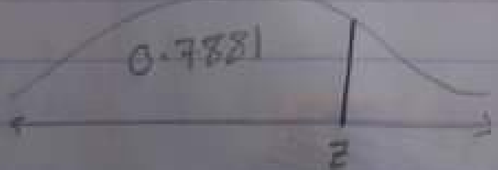
$= 0.664$.



* Example :-

let Z be a standard normal random variable :-

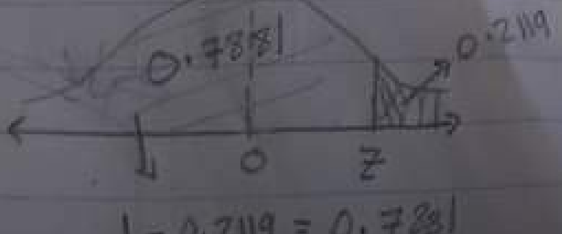
1- Find Z such that the area left to Z is 0.7881.



$Z = 0.80$

2)- Find Z such that area right to Z is 0.2119.

$1 - \square$



$Z = 0.80$

$1 - 0.2119 = 0.7881$

"84"

3). Find Z such that the area left to Z is 0.9

Closet area = 0.8997

Z = 1.28.

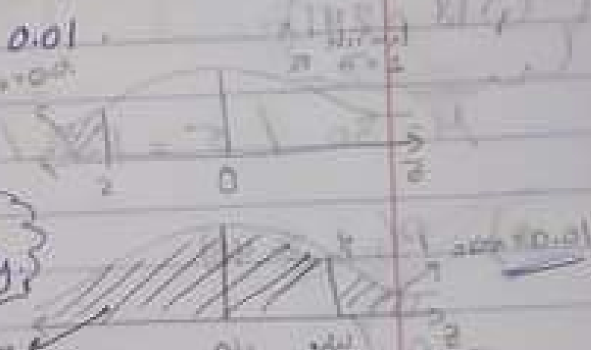
Z	0.08	0.09
1.2	0.8997 0.0003	0.9015 0.0015

ملحوظة: إذ وقعت رقمين لهما نفس البعد نأخذ الأربح بين الرقمين.

4)- Find Z such that the area left to Z is 0.01.

Z	0.02	0.03
2.3	0.9898 0.0002	0.9901 0.0001 الأقرب

Z = -W
Symmetry.



area = 0.99 complement

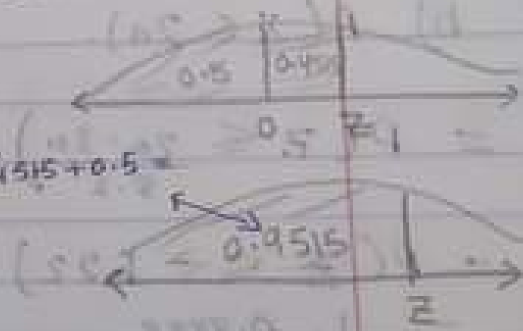
Closet area = 0.9901

W = 2.33

Z = -2.33.

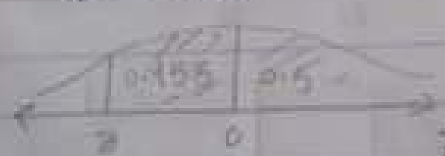
* العلاقة بين Z و W متماثلات في القيمة ومختلفات في الإشارة.

5) Find Z such that area between 0 and Z is 0.4515.



0.4515 + 0.5 = 0.9515

Z = 1.66

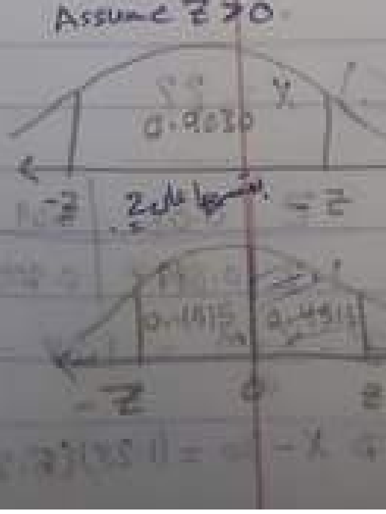


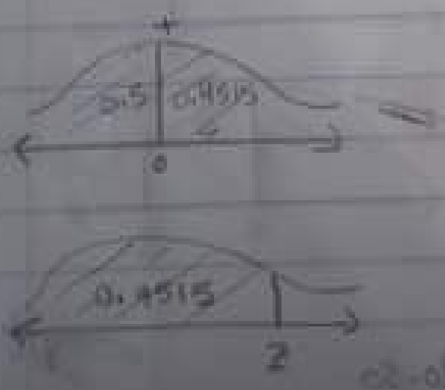
Z = -1.66

∴ Z = ± 1.66.

6). The area from -Z to Z is 0.9030. Find Z.

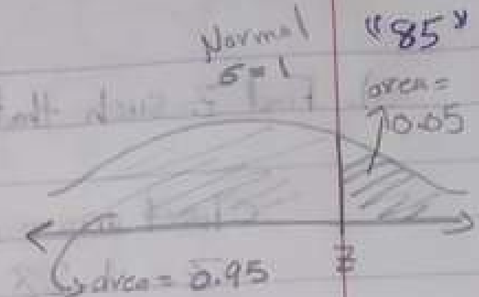
Assume Z > 0.





Z = 1.66.

"85"



⑦. Find z such that the area right to z is 0.05.

z	0.04	0.05
1.6	0.9495	0.9505
	0.0005	0.0005

متساويات Closest area: both.

$$z = \frac{1.64 + 1.65}{2} =$$

Q18 p241:

$\mu = 30$ / $\sigma = 8.2$ / x : stok price / x: is a normal random variable.

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

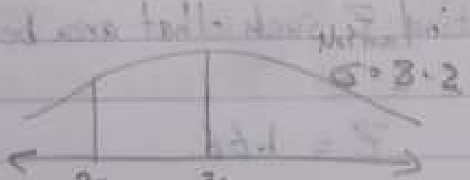
a)- $P(x \geq 40)$

$$= P\left(\frac{x-\mu}{\sigma} \geq \frac{40-30}{8.2}\right)$$

$= P(z \geq 1.22)$.

$= 1 - 0.8888 = 0.1112$.

b)- $P(x \leq 20)$.

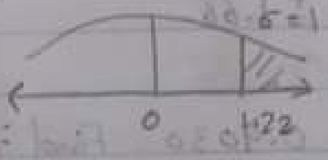


$= P\left(z \leq \frac{20-30}{8.2}\right)$.

$= P(z \leq -1.22)$

$= 1 - 0.8888$

$= 0.1112$.

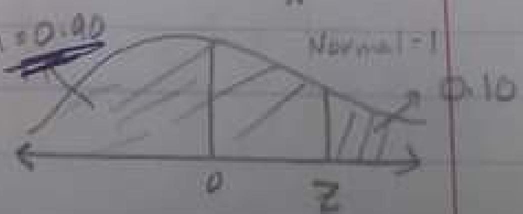


c) $x = ??$ Top 10%.

z	0.08	0.09
1.2	0.8997	0.9015

⇒ closest area = 0.8997

$z = 1.28$

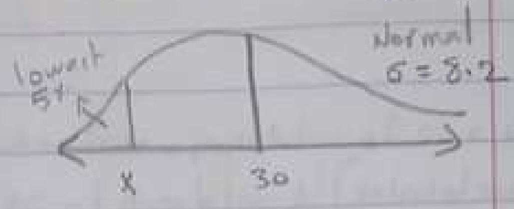


$$z = \frac{x - \mu}{\sigma} = 1.28 = \frac{x - 30}{8.2}$$

$\Rightarrow x - 30 = (1.28)(8.2) \Rightarrow x = 30 + (1.28)(8.2) = 40.50$.

* Question:-

How low does a stoch price have to be to puta a company in the lowest 5% ?



Z	0.04	0.05
1.6	0.9495	0.9505

$Z = -1.645$.

$$Z = \frac{X - \mu}{\sigma} \Rightarrow -1.645 = \frac{X - 30}{8.2}$$

$$\Rightarrow X - 30 = (-1.645)(8.2)$$

$$X = 30 + (-1.645)(8.2)$$

$$X = 16.51$$

Chapter 7 :- Sampling and Sampling Distrubution :- كيفية اختيار العينات

Saction 7.2

Selecting a Sample.

* Sampling from a finite population
→ A Simple random Sample of size n taken from a population of size N is a sample
Size of sample n Size of population N [selected such that each possible Sample of size has the Same prob. of being selected].

Frame
1 ① ⑤ ㉗ ㊼ ② stonen
2
3 $\boxed{7}$ $\boxed{9}$ $\boxed{20}$ $\boxed{101}$ $\boxed{37}$ papeltion.
4
5 Random number Table
⋮ Calcalators / Computores → random numbers generations.

N=200 طريقتها ↓

$\boxed{\text{العدد}}$. $\boxed{\text{Shift}}$ $\boxed{\cdot}$ =
Rank

Note :-

* Sampling without replacement: 1, 7, 9, ~~1~~, 3, ~~7~~, ~~9~~, ~~3~~, 8 → 5 elemanes بدون تكرار الرقم مثلاً

* Sampling with replacement: 1, 7, 9, 1, 3. مع تكرار الرقم

* Sampling from a process :-

Example -: Factory → production → quality Contoral.

→ Random Sample :-
* Sampled elements should be independent.

* Sampled elements have the same prob. dist.

Section 7.3
Point Estimation

Sample statistic :- is a point eslimotor for the population parameten.

* Sample population: the population that we take Sample from.

* Target population:- the population that we want to make inference about.

Saction 7.7

other Sampling Methouds P284-286.

* Stratified Random Sampling.
* Cluster Sampling.
* Systemtic Sampling.

probabalitic Sample metuodes.
طرق أخذ عينات احتمالية.

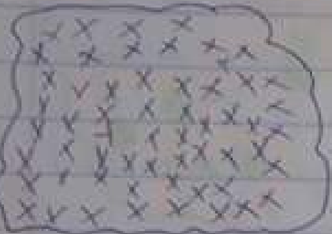
* Convenience Sampling.
* Judgement Sampling.

Non-probalitic Sample methoudes
طرق أخذ عينات غير إحتمالية.

Captor 8 : Interval Estimation

Section 8.1

Population mean : σ know.



population
$\mu = ??$

Take a Sample : $x_1, x_2 \ldots\ldots x_n$ ← random Sample

Sample size : n

Find apoint estimator for μ : $\bar{X}$

$\Rightarrow \mu \approx \bar{X}$ → statzshial infarance.

Sample should be random.

Confidence interval / interval estimation.

$$(1-\alpha)\,CI = \bar{X} \mp Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} = \left[\bar{X} - Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}},\ \bar{X} + Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right]$$

CI : Confindence interval

$1-\alpha$: Confindence coeffeaient.

$(1-\alpha)\,100\%$: Confidence level

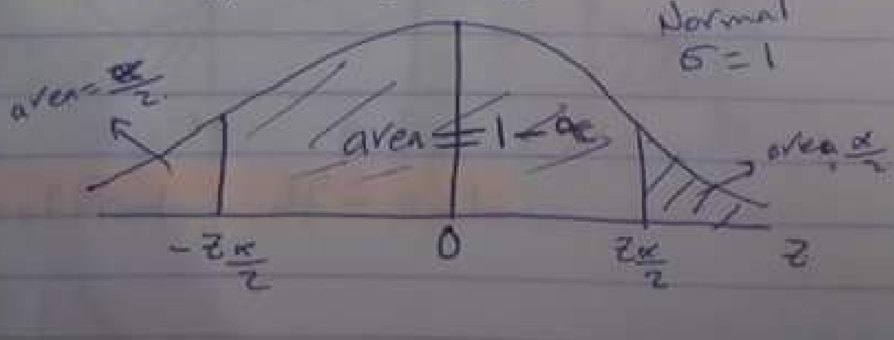
$\bar{X}$: Sample mean.

$\frac{\sigma}{\sqrt{n}}$: standard erroe

$Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} = E$: margin of erone.

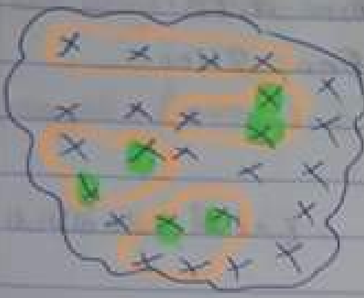
σ : population st. dev

* $Z_{\frac{\alpha}{2}}$: Standand normal value sech that
(the area right to $Z_{\frac{\alpha}{2}}$ is $\frac{\alpha}{2}$



Section 8.2

Population mean : σ unknown.



population
$\mu = ??$

take a sample : $x_1, \ldots, x_n$.
Sample size : n.
Find a point estimator for μ : $\bar{x}$
$\Rightarrow \mu \approx \bar{x}$ ← statistical interval.
↑ Sample should be random.

Confidence interval / Interval Estimation :-

$$(1-\alpha)CI = \bar{x} \mp t_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}} = \left[\bar{x} - t_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}} , \bar{x} + t_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}} \right]$$

CI : confidence interval

$(1-\alpha)$: confidence coeffecient.

$(1-\alpha)100\%$:- Confidence level

$\bar{x}$: Sample mean.

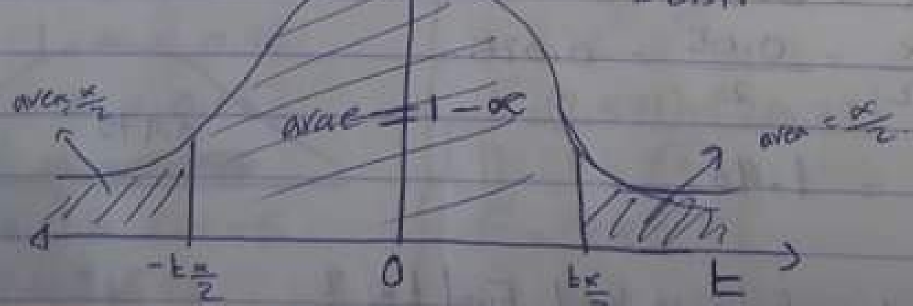
$\frac{S}{\sqrt{n}}$: Standard error.

$E : t_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}}$: margin of error.

S : Sample st. dev.

$t_{\frac{\alpha}{2}}$: t value such that the area right to $t_{\frac{\alpha}{2}}$ is $\frac{\alpha}{2}$.

t Dist.



degrees of freedom : $df = n-1$.

Example :- Some one is intrested in the mean salary for a group of employees. The person collected a sample and the salaries were :-

Salaries :- 2300, 2200, 2700, 2900, 2500, 2400.

① - Find a point estimator for $M = \bar{X}$

$\bar{X} = 2500.$ على الآلة حاسبة بأخذ الأرقام.

②. Assuming that the population standard deifition is equal to 300, Find the standard error :- $\sigma = 300$.

$$\text{Standard error} = \frac{\sigma}{\sqrt{n}} = \frac{300}{\sqrt{6}} = 122{,}47.$$

③ - Assuming that we dont know the population st. dev, find the standard error. unknow. $\leftarrow \sigma = ??$, standard error = ??

$$\text{Stander error} = \frac{S}{\sqrt{n}} = \frac{260.77}{\sqrt{6}} = 106{,}46$$

على الآلة الحاسبة ثبت رقم 2 وبعدها رقم 3

يستخدم لما يكون معلوم

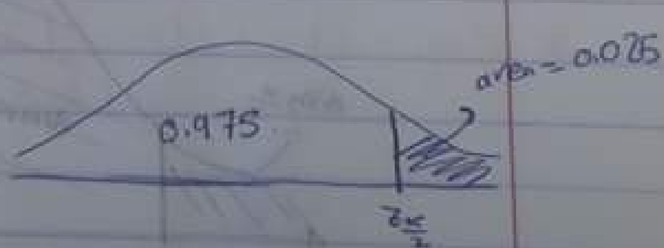
④ - Using 95% confidence level, find $Z_{\frac{\alpha}{2}}$? or $t_{\frac{\alpha}{2}}$ ومنه نستنتج يكون مجهول

$1 - \alpha = 0.95 \longrightarrow (1-\alpha)100\% = 95\%$

$\alpha = 0.05$

$\frac{\alpha}{2} = \frac{0.05}{2} = 0.025.$

$Z_{\frac{\alpha}{2}} = 1.96.$



⑤. Using 95% confidence level, Find $t_{\frac{\alpha}{2}}$?

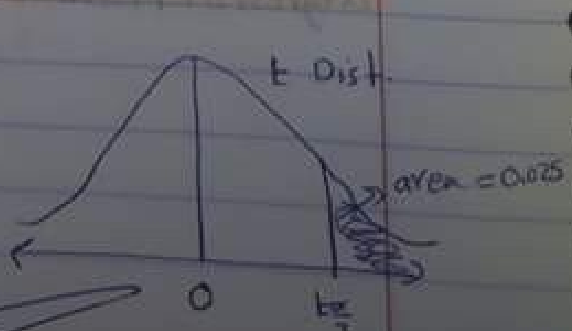
$1 - \alpha = 0.95$ $(1-\alpha)100\% = 95\%$.

$\alpha = 0.05$

$\frac{\alpha}{2} = \frac{0.05}{2} = 0.025$

$df = n - 1 = 6 - 1 = 5$

$t_{\frac{\alpha}{2}} = 2.571.$



6. Find the margh of error, assuming the population st. dev. is 300 and using 95% confichese level.

margh of error :- $E = \begin{cases} Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} & , \sigma \text{ Know} \leftarrow \sigma = 300 \rightarrow \text{معروفة} \\ t_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}} & , \sigma \text{ unknow.} \leftarrow S = 260.77. \end{cases}$

σ know, $\sigma = 300 \Rightarrow E = Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$

$Z_{\frac{\alpha}{2}} = 1 - \alpha = 0.95$

$\alpha = 0.05$

$\frac{\alpha}{2} = 0.025$

$Z_{\frac{\alpha}{2}} = 1.96.$

$E = Z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$

$E = \frac{(1.96) \cdot 300}{\sqrt{6}} = 240.05.$

7. Find the margh of error using 95% level of confidence:

Assume the population st. dev is unknow. → $\boxed{S}$

$(n=6, S=260.77, t_{\frac{\alpha}{2}} = ??)$

margin of error :- $E = t_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}}$

$t_{\frac{\alpha}{2}} = (1-\alpha)100\% = 95\%$

$1-\alpha = 0.95$

$\alpha = 0.05$

$df = n-1 = 6-1 = 5$

$t_{\frac{\alpha}{2}} = 2.571$

$E = t_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}}$

$E = \frac{(2.571) \cdot (260.77)}{\sqrt{6}}$

$E = 273.71.$

(8). Assuming the population st. dev is 300.
Find the 95% confidence interval for the population mean.
Comment on the result. σ know.

$(1-\alpha)\,CI = \bar{X} \mp Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$

$95\%\ CI = 2500 \mp 1.96 \cdot \frac{300}{\sqrt{6}}$

$95\%\ CI = 2500 \mp 240.05$

$= [2259.95, 2740.05]$

$(2500 - 240.05)$ $(2500 + 240.05)$

* We are 95% confident that μ is between 2259.95 and 2740.05

(9). Find the 95% confident interval for the population mean.
Comment on the result. σ unknow.

$(1-\alpha)\,CI = \bar{X} \mp t_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}}$

$95\%\ CI = 2500 \mp 2.571\ \frac{260.77}{\sqrt{6}}$

$= 2500 \mp 273.71.$

$= [2226.29, 2773.71].$

* We are 95% confident that μ is between 2226.29 and 2773.71

(10). Find the 90% confidence interval for μ.

$n = 6$, $\bar{X} = 2500$, $S = 260.77$. σ unknow.

$(1-\alpha)\,CI = \bar{X} \mp t_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}}$

$= 2500 \mp 2.015 \frac{260.77}{\sqrt{6}}$

$= 2500 \mp 214.51$

$= [2285.49, 2714.51]$.

$1-\alpha = 0.90$
$\alpha = 0.10$
$\frac{\alpha}{2} = 0.05$
$df = n-1 = 6-1 = 5$
$t_{\frac{\alpha}{2}} = 2.015$.

(11). Assuming the population st. dev is 300. Find the 90% confidence interval for μ. σ know.

$n = 6$, $\bar{X} = 2500$, $S = 260.77$, $\sigma = 300$.

$(1-\alpha)\,CI = \bar{X} \mp Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$

$= 2500 \mp 1.645 \frac{300}{\sqrt{6}}$

$= 2500 \mp 201.47$.

$= [2298.53, 2701.47]$

t_table : $df = \infty \rightarrow Z$ values.
$1-\alpha = 0.90$
$\alpha = 0.10$ $(df = \infty)$
$\frac{\alpha}{2} = 0.05$
$Z_{\frac{\alpha}{2}} = 1.645$.

σ known $CI \sim Z_{\frac{\alpha}{2}}$

90% $CI = [2298.53, 2701.47]$

95% $CI = [2259.95, 2740.05]$

σ unknown $CI \rightarrow t_{\frac{\alpha}{2}}$

90% $CI = [2285.49, 2714.51]$.

95% $CI = [2226.29, 2773.71]$

Notes:—

1- The above methods are used assuming.

- The sample taken are random.
- The dist. of the data must be normal or take a large sample $n \geq 300$

$\left\{\begin{array}{l} t\text{-table} \\ df=\infty \\ Z \text{ value} \end{array}\right.$

$$(1-\alpha)\,CI = \begin{cases} \bar{x} \mp Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} & \sigma \text{ known} \\ \bar{x} \mp t_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}} & \sigma \text{ unknow.} \end{cases}$$

2- If the confidensec level is lieght, the CI will be widen.

3- If the sample size is larger, the CI will be shorter.

Section 8.3

Determing the Sample size.

$$(1-\alpha)\,CI \text{ for } \mu = \begin{cases} \bar{x} \mp Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}, & \sigma \text{ known.} \\ \bar{x} \mp t_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}}, & \sigma \text{ unknown.} \end{cases}$$

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

random up.

σ known, $n = \left(\frac{Z_{\frac{\alpha}{2}}\sigma}{E}\right)^2$ random up. $n \rightsquigarrow Z_{\frac{\alpha}{2}}$

σ unknown:- $n = \left(\frac{Z_{\frac{\alpha}{2}}\sigma^*}{E}\right)^2$ random up. $\rightarrow n \rightsquigarrow Z_{\frac{\alpha}{2}}$

σ^*:- planning value of σ we use it to find n when σ is unknown.

* How do we get σ^*?

1. previous studien.
2. Pilot study.
3. Judgement / Guess.
4. $\sigma^* = \frac{\text{Range}}{4}$

Example :- Sample of salaries :-

Preious study "98" → Pilot study

2300, 2200, 2700, 2900, 2500, 2400.

1)- we want to construct a 99% confidence interval of the population mean. we want the marge of error to be 150. Find the sample size requared to get the about CI.

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

E = 150

σ unknow

$\sigma^* = S$

= 260.77

$1-\alpha = 0.99$

$\alpha = 0.01$

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

$Z_{\frac{\alpha}{2}} = 2.576$

$df = \infty$

$$= \left(\frac{(2.576)(260.77)}{150} \right)^2$$

= 20.06 round up (n = 21.)

2)- we want to construct a 99% confiduce interval for the population mean. we want the marge of error to be 150. Find the sample size requned to get the about CI.

Assuming the population st.dv is 300.

$$n = \left(\frac{Z_{\frac{\alpha}{2}} \sigma}{E} \right)^2 = \left(\frac{(2.576)(300)}{150} \right)^2$$

$\sigma = 300$

$E = 150$

$1-\alpha = 0.99$

$\alpha = 0.01$

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

$Z_{\frac{\alpha}{2}} = 2.576$

$df = \infty$

= 26.54 roundo up = تقريب

$\boxed{n = 27}$

Chapter 9 → Hypothesis Tests

Section 9.1

Developing Null and Altenative hyptheses.

H_0 : Null Hoypothesis.

H_a : Alternative Hypothesis.

$H_0: \mu \geq \mu_0$

$H_a: \mu < \mu_0$

Lower Tail Test.

$H_0: \mu \leq \mu_0$

$H_a: \mu > \mu_0$

Upper Tail Test.

One - Tailed Test.

$H_0: \mu = \mu_0$

$H_a: \mu \neq \mu_0$

Tow - Tailed Test.

Not :-

μ_0 : hypothesized value.

"100"

Section 9.2

Type I and Type II errors.

population condition.

Conclusion	H_0 True	H_a True
Accept H_0	Correct Conclusion	Type II error
Reject H_0	Type I error	Correct Conclusion

$\mu = 17$

$H_0 \to$ Fales
$H_a \to$ True

$H_a \to$ True
$H_0 \to$ Fales

Def:- The level of significance is the prob. of making Type I error when the null hypotheais is true and an equality

level of significance : α

* المعطيات الواجب الإطلاع عليها عند الحل:

* Sample :- random & (normal or large enough)
* α : level of significance $\to$ Ch. 9.
* $1-\alpha$: level of confidence $\to$ Ch. 8.
* σ known / σ unknown.
* Tow-Tailed / one-Tailed

Upper Tail   lower Tail.

* Critical Value aproach
* P-Value aproach.

طرق الحل ⟵

إذا ما حدد السؤال الطريقة للحل بأي طريقة نحلها

Section :

Section 9.3 → population mean , σ known.

Section 9.4 → population mean , σ unknown.

Tow Tailed	lower Tailed	upper Tailed.
$H_0: \mu = \mu_0$	$H_0: \mu \geq \mu_0$	$H_0: \mu \leq \mu_0$
$H_a: \mu \neq \mu_0$	$H_a: \mu < \mu_0$	$H_a: \mu > \mu_0$
σ known: CVA, PVA	σ known: CVA, PVA	σ known: CVA, PVA
σ Unknown: CVA, PVA	σ unknown: CVA, PVA	σ unknown: CVA, PVA.

Example :

Some one is intrested in the mean salary for a group of employees.
The person collected a sample and the salaries whe an followes :-

random → [23.00 / 2200 / 2700 / 2400 / 2500 / 2400 .

The person claimen that the population mean is equal to 2600.
With α : 0.05 level of signficance, what is your canclusuin? يدعي

* write the hypotheses :- اكتب الفرضيات

$H_0 : \mu = 2600$ $\qquad \mu_0 = 2600$

$H_a : \mu \neq 2600$ $\qquad$ Tow - Tailed.

* Calculat the test statistic:-

Test about μ when σ unknown

$$t = \frac{\bar{X} - \mu_0}{\frac{S}{\sqrt{n}}} \quad df = n - 1 \longrightarrow \text{Becounse } \sigma \text{ unknow.}$$

$\bar{X} = 2500$

$S = 260.77$

$n = 6$

$\mu_0 = 2600$

$$t = \frac{2500 - 2600}{\frac{260.77}{\sqrt{6}}} = -0.94 \rightarrow \text{تقريبا الأقرب منزلتين}$$

* Rejection Rule for two-Tailed Test about μ (σ unknown).

* Reject H_0 if $|t| \geq t_{\frac{\alpha}{2}}$

* Reject H_0 if $t \geq t_{\frac{\alpha}{2}}$ or $t \leq -t_{\frac{\alpha}{2}}$

* Reject H_0 | Accept H_0 | Reject H_0

$-t_{\frac{\alpha}{2}}$ ↓ critical value

$+t_{\frac{\alpha}{2}}$ ↓ critical value.

* Find critical value :-

$-t_{\frac{\alpha}{2}}$, $+t_{\frac{\alpha}{2}}$

$df = n-1$.

$\alpha = 0.05 \longrightarrow \frac{\alpha}{2} = \frac{0.05}{2} = 0.025$

$df = n-1 \longrightarrow df = 6-1 = 5$

$\longrightarrow t_{\frac{\alpha}{2}} = 2.571$.

Critical value : -2.571 , $+2.571$

Reject H_0 | Accept H_0 | Reject H_0

-2.571 ↓ $-t_{\frac{\alpha}{2}}$: critical value.

-0.94 t : test statistic

$+2.571$ ↓ $+t_{\frac{\alpha}{2}}$: critical value.

Conclusion.

- Accept H_0 ($\alpha : 0.05$)
- $\mu = 2600$ ($\alpha : 0.05$)
- with significance $\alpha = 0.05$, we accept that population mean is 2600.

Example :

Some is intrested in the mean salary for a group of employees. The person collected a sample and the salaries were as follows :-

2300, 2200, 2700, 2900, 2500, 2400

The person claims that the population meanor is equal to 2600. with $\alpha = 0.05$ leveal of the significance, what is your conclusion? Assume that the population st. dev is equal to 300 ?

* write the hypo theses

$H_0 : \mu = 2600$

$H_1 : \mu \neq 2600$

$\bar{x} = 2500$

~~$S = 260.77$~~

$n = 6$

$\mu_0 = 2600$

$\sigma = 300$

* Calculated the test statislic

$$\boxed{Z = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}}}$$ → Test statistic use Test about μ σ known

$$Z = \frac{(2500 - 2600)}{\left(\frac{300}{\sqrt{6}}\right)} \Rightarrow Z = -0.82.$$

* State the Rejection Rule for the Two tented Test about μ (σ know) or

* Reject H_0 if $|Z| \geq Z_{\frac{\alpha}{2}}$

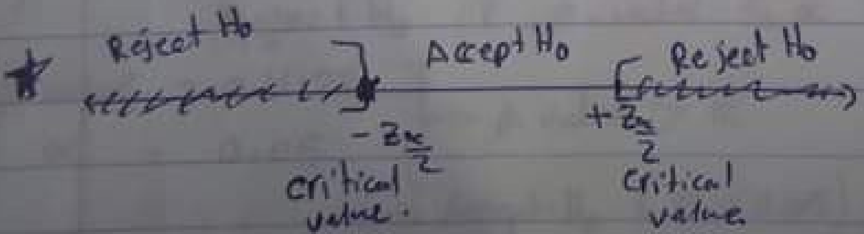
* Reject H_0 if $Z \geq Z_{\frac{\alpha}{2}}$ or $Z \leq -Z_{\frac{\alpha}{2}}$

[Find the cretical values :-

$-Z_{\frac{\alpha}{2}}$ and $+Z_{\frac{\alpha}{2}}$.

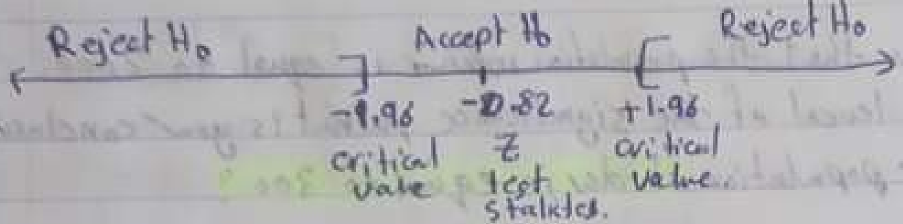
[t - table, df = ∞

*



$$\left.\begin{array}{l}\alpha = 0.05 \Rightarrow \frac{\alpha}{2} = \frac{0.05}{2} = 0.025 \\ df = \infty \text{, t-table}\end{array}\right] \Rightarrow Z_{\frac{\alpha}{2}} = 1.96$$

Critical value :- -1.96 and $+1.96$.



$H_0 : \mu = 2600$

$H_1 : \mu \neq 2600$

Conclusion:-

* Accept H_0 $(\alpha = 0.05)$.
* $\mu = 2600$ $(\alpha = 0.05)$.

* With significance $\alpha = 0.05$ we accept that the population mean is 2600.

Example:-

Some is interested in the mean salary for a group of employees. The person collected a sample and the salaries when as follows :-

2300, 2200, 2700, 2900, 2500, 2400

The person claming the population mean is equal to 2600. with $\alpha = 0.05$ level of significance, what is your conclusion? Assume that the population st.dev equal to 300? Use the p-value approach.

$H_0 \Rightarrow M = 2600$

$H_1 \Rightarrow M \neq 2600$

$$Z = \frac{\bar{X} - M}{\frac{\sigma}{\sqrt{n}}} = \frac{(2500 - 2600)}{\left(\frac{300}{\sqrt{6}}\right)} \Rightarrow Z = -0.82$$

* Find the p-value :-

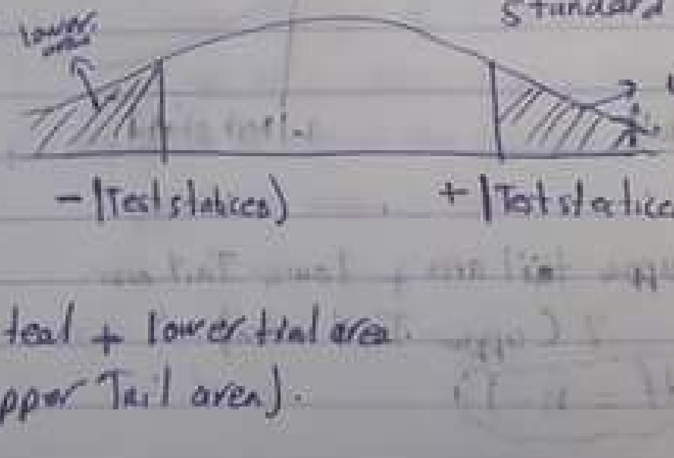
p-value =

$= P(Z < -0.82) + P(Z \geq 0.82)$.

$= 2(1 - 0.7939)$.

$= 0.4122$

standard Normal Dist.



(p-value (Tow Taled test) ←
(about M (σ known)

$p\text{-value} = \text{upper teal} + \text{lower tial area}$.

$= 2(\text{upper Tail area})$.

* State the Rejection Rule using p-value approach ?

Reject H_0 if $p\text{-value} \leq \alpha$

$$\left.\begin{array}{l} p\text{-value} = 0.4122 \\ \alpha = 0.05 \end{array}\right\} \Rightarrow p\text{-value} > \alpha$$

$\Rightarrow$ Accept H_0 $(\alpha = 0.05)$ / 2. $M = 2600$ $(\alpha = 0.05)$.

3. with $\alpha = 0.05$, we accept that the population mean ...

Example :-

Some is interested in the mean salary for a group of employees. The person collected a Sample and the Salaries when as follows.

{ 2300, 2200, 2700, 2900, 2500, 2400 }

The prsone claimen that the population mean is equal to 2600. with $\alpha = 0.05$ level of significouse, what is your conclusion? Use the p-value approach.

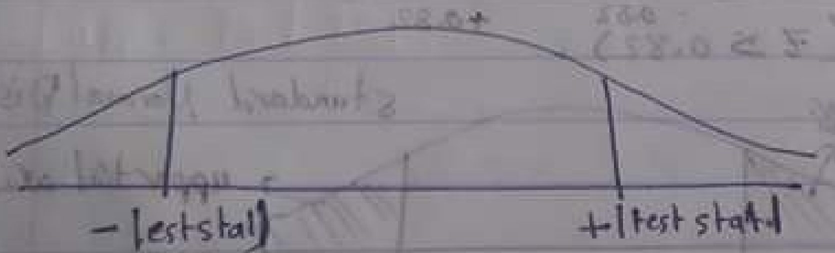
$H_0 : \mu = 2600$

$H_a : \mu \neq 2600$

* $t = \frac{\bar{X} - \mu_0}{\frac{s}{\sqrt{n}}} = \frac{(2500 - 2600)}{\left(\frac{260.74}{\sqrt{6}}\right)} \Rightarrow t = -0.94$

* Find the p-value :-

P-value (Two Tailed test about μ) (σ known)



P-value = upper tail area + lower Tail area.

= 2 (upper Tail area)

(df = n-1)

df	0.20	0.10
5	0.92	1.476

$H_0: \mu = 2600$

$H_a: \mu \neq 2600$

$t = -0.94$

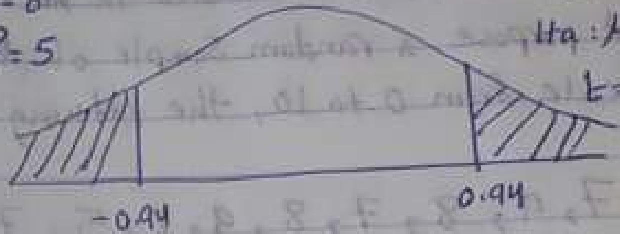
$n = 6$

$df = 5$

← upper tail area

df	0.20	0.10
5	0.92	1.476

↑ 0.94



upper tail area $\in (0.10, 0.20)$

p-value $\in (0.20, 0.40)$

p-value $= P(t > +0.94) + P(t < -0.94)$

$= 2 \cdot$ (upper tail area)

* State the Rejection Rule using the p-value approach :-

Reject H_0 if p-value $\leq \alpha$

$H_0: \mu = 2600$

$H_a: \mu \neq 2600$

$\alpha = 0.05$

p-value $\in (0.20, 0.40)$

→ p-value $\overset{\text{أكبر}}{>}$ α

$\Rightarrow$ Accept H_0 $(\alpha = 0.05)$

$\Rightarrow \mu = 2600$ $(\alpha = 0.05)$

$\Rightarrow$ with $\alpha = 0.05$ the accept that the population mean is 2600.

Example:- Some one is interested in the rating of a certain airport. For that purpose a random sample of the customers was taken on a scale from 0 to 10, the ratings were :-

[7, 9, 8, 7, 8, 9, 6, 5, 7, 7, 9, 9.]

The airport is considered very good if the rating is more than or equal to 7 with $\alpha = 0.05$, is the above airport very good? (one-Tailed Test)

* $H_0: \mu \leq 7$
  $H_a: \mu > 7$

المعطيات
ملاحظة - ندخل المعطيات الآلة حاسبة
كما تعلمنا في شباتر السابقة.

$n = 12$ $\mu_0 = 7$
$\bar{X} = 7.58$ $\alpha = 0.05$
$S = 1.31$ σ unknow

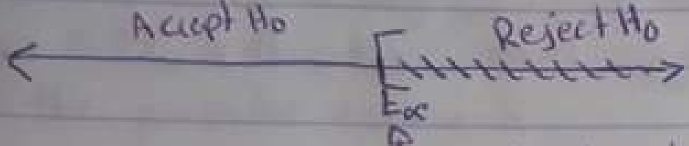
* $t = \dfrac{\bar{X} - \mu_0}{\frac{S}{\sqrt{n}}} = \dfrac{(7.58 - 7)}{\left(\frac{1.31}{\sqrt{12}}\right)} = 1.53$

note :- one tailed Test
* test statistic positive.
→ upper Tail Test.

* State the rejection Rule using the critical value approach.

⇒ upper Tail test (σ unknown)

* Reject H_0 if $t \geq t_\alpha$

* 

$df = n-1$

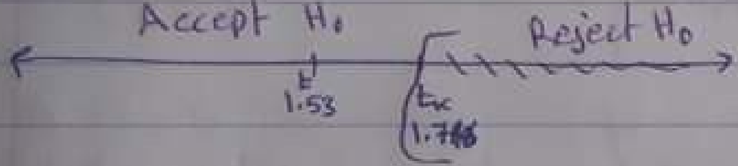
Critical value : $+t_\alpha$

Critical value approach.

- $t = 1.53$
- $\alpha = 0.05$

من الجدول.

$df = n-1 = 12-1 = 11 \Big\} \rightarrow$ Critical value $= t_\alpha = 1.796$

- 

$H_0 : \mu \leq 7$

$H_1 : \mu > 7$

- Conclusion : Accept H_0 ($\alpha = 0.05$) $\rightarrow$ $\mu \leq 7$ ($\alpha = 0.05$)

Example 80

Someone is intrested in the rating of a certain airport. For that purpose a random Sample of Customers was taken, on a scale from 0 to 10, the ratings were?

7, 9, 8, 7, 8, 9, 6, 5, 7, 7, 9, 9.

the airport is consdied very good, of, the rating is more than or equal to 7, with $\alpha = 0.05$, is the above airport very good? ↳ one Tailed Test.

* Use the p-value approach.

- $H_0 : \mu \leq 7$
- $H_1 : \mu > 7$

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

$$= \frac{(7.58 - 7)}{\left(\frac{1.31}{\sqrt{12}}\right)} = 1.53$$

7, 9, 8, 7, 8, 9, 6, 5, 7, 7, 9, 9

$n = 12$ $\mu_0 = 7$

$\bar{x} = 7.58$ $\alpha = 0.05$

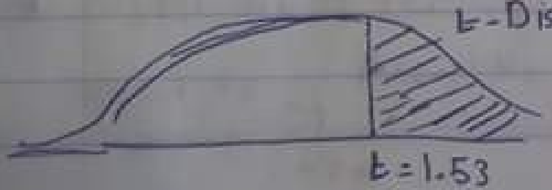
$s = 1.31$ σ = unknown.

* one tailed test.

* test statistic (postive)

→ upper Tailed test

* Find the p-value.



df	0.10	0.05
11	1.363	1.796

↑ t = 1.53

p-value ∈ (0.05, 0.10)

p-value (upper Tail test about μ) (σ unknown).

p-value = area in upper tail

$df = n - 1$
$12 - 1 = 11$

* State the rejection Rule using p-value approch :-

Reject H_0 if p-value $\leq \alpha$

- $\alpha = 0.05$

p-value $\in (0.05, 0.10)$ $\Big\}$ p-value $> \alpha$

Accept H_0 $(\alpha = 0.05)$

$\Rightarrow \mu \leq 7$ $(\alpha = 0.05)$ ..

Question:-

Repeat the previous example assuming the population standard desiation is equal to 2. Use critical value approach, Use p-value approach.

$t \to Z$

$t_\alpha \to Z_\alpha$

p-value rang $\to$ exact p-value

t-table $\qquad$ Z-t

Example:-

A restaurant, says that the service is very good. very good manes a rating equal to or higher than 8. To test this claim we took a sample of customers and the rating were as follows :-

8, 6, 10, 4, 10, 9, 5, 10

At $\alpha = 0.10$, What is your conclusion?

Assume the population st. dev. is equal to 2. /use p-value aprach.

$n = 8$ $\mu_0 = 8$

$\bar{x} = 7.75$ $\alpha = 0.10$

~~$s = 2.43$~~ $\sigma = 2$

σ known

- $H_0 : \mu \geq 8$

  $H_a : \mu < 8$

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

- one tail test
- test statistic → negative
- → lower tail test.

$$Z = \frac{(7.75 - 8)}{\left(\frac{2}{\sqrt{8}}\right)} = -0.35$$

* state the rejection Rule using critical value approach. Lower tail test (σ known)

* Reject H_0 if $Z \leq -Z_\alpha$

Reject H_0 Accept H_0

$-Z_\alpha$

critical value : $-Z_\alpha$ ($df = \infty$, t- table)

critical value approach

- $Z = -0.35$ critical value.
- $\alpha = 0.10$
- $df = \infty$

$-Z_\alpha = -1.282$

Reject Accept H_0

-1.282 -0.35

$-Z_\alpha$ Z

Conclusion:-

Accept H_0 ($\alpha = 0.10$)

$\mu \geq 8$ ($\alpha = 0.10$)

- Find p-value

p-value = lower tail area.
$= 1 - 0.6368$
$= 0.3632$

p-value (lower tail test about μ).
(σ known)

p-value = lower tail area.

(z-table exact-p-value) z-table, $df = \infty$ p-value range.

- state the rejection rule using p-value approach.
Reject H_0 if p-value $\leq \alpha$.

- $\alpha = 0.10$
p-value $= 0.3636$

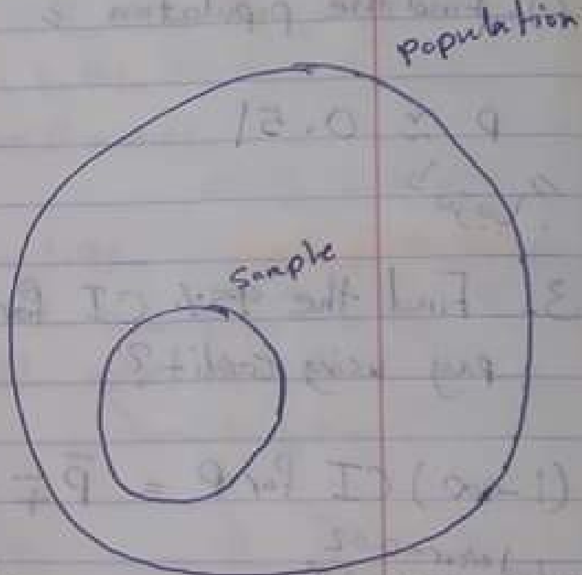
p-value $> \alpha$
Accept H_0 ($\alpha = 0.10$)
$\mu \geq 8$ ($\alpha = 0.10$).

Questions:-
Repeat the pervouse example withoust assuming that σ is known.
Use critical value approuch, use p-value approach.

Chapter 8 → Interval Estimation.
Section 8.4 → population proportion.

Example:- In a certain population we are interested in the ~~population~~ proportion of people who use credit care.
For the sake of estimating ~~are~~ the population porportain we took a random sample.
In the sample 462 out of 900 saind they are crieelt care the main payment

$\bar{P}$: Sample proportion.



P: population proportion.

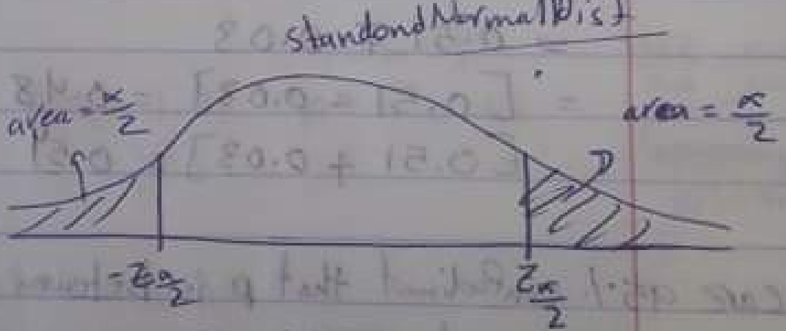
$\bar{P}$ point estimator
stat.

$$(1-\alpha)\ CI\ for\ P = \bar{P} \mp Z_{\frac{\alpha}{2}} \sqrt{\frac{\bar{P}(1-\bar{P})}{n}}$$

$(1-\alpha)$: → Confidence coeffeeeit

$(1-\alpha)100\%$: Confidence level

$\sqrt{\frac{\bar{P}(1-\bar{P})}{n}}$: Standand Error.



$Z_{\frac{\alpha}{2}} \sqrt{\frac{\bar{P}(1-\bar{P})}{n}} = E$: marginof erore

Random Sample and large enough.

Back to Example:-

1- Find the sample proportion of people who prefer to pay using credit card?

$\bar{P} = \frac{462}{900} = 0.51$

2)- Find the population " " " " " " " " " ?

$P \approx 0.51$

"تقريبا"

3- Find the 95% CI for the population porportion of people who prefer to pay using credit?

$(1-\alpha)$ CI for $P = \bar{P} \mp Z_{\frac{\alpha}{2}} \sqrt{\frac{\bar{P}(1-\bar{P})}{n}}$

standard error = 0.02
marginal error = 0.03

$= 0.51 \mp 1.96 \sqrt{\frac{0.51(1-0.51)}{900}}$

$= 0.51 \mp 0.03$

$= [0.51 - 0.03] = 0.48$

$[0.51 + 0.03] = 0.54$

We are 95% confident that p is between 0.48 and 0.54.

$\bar{P} = \frac{462}{900} = 0.51$

$n = 900$

$1-\alpha = 0.95$

$\alpha = 0.05$

$\frac{\alpha}{2} = \frac{0.05}{2} = 0.025$

$Z_{\frac{\alpha}{2}} = 1.96$

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

$n = \frac{(Z_{\frac{\alpha}{2}})^2 \bar{P}(1-\bar{P})}{(E)^2}$

$$n = \frac{(Z_{\frac{\alpha}{2}})^2 P^*(1-P^*)}{E^2}$$ n

round up

تقريب لعدد صحيح

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

$p \approx 0.51$

$\bar{p} = 0.51$

95% CI for P

$= [0.48, 0.54]$.

$E = 0.03$.

$n = 900$

p^* : planing value for P

1- pervious study
2- pilot study.
3. Judgement / guess.
4. $p^* = 0.5$.

Back to Example:-
Find the sample size required to have margin of error $E = 0.01$.

$E = 0.01 \quad n = ?? \quad 1-\alpha = 0.95 \quad Z_{\frac{\alpha}{2}} = 1.96. \quad p^* = 0.5$

$$n = \frac{(Z_{\frac{\alpha}{2}})^2 \, p^*(1-p^*)}{E^2} = \frac{(1.96)^2 (0.5)(0.5)}{(0.01)^2} \Rightarrow n = 9604.$$

$p^* = 0.5$
$0.51 \mp 0.01 = [0.50 \text{ and } 0.51]$.

Back to Example :-

من السؤال p

$E = 0.01 \quad n = ?? \quad 1-\alpha = 0.95 \quad Z_{\frac{\alpha}{2}} = 1.96 \quad p^* = 0.51$

$\frac{462}{900} = 0.51$

previouse pilot study

$$n = \frac{(Z_{\frac{\alpha}{2}})^2 \, p^*(1-p^*)}{E^2}$$

$$n = \frac{(1.96)^2 (0.51)(1-0.51)}{(0.01)^2} = 9600.16$$

$n = 9601$

The End

Chapter 9 → Hypothesis Test.

Section 9.5 → population propotion.

Example :- Conicla the following hypotuesis Test

Two-Tailed test

$H_0 : p = 20$

$H_a : p \neq 20$

n: Sample size.

P_0: hypo thesize valleu

A sample of size 400 provided a Sample population 0.175.
Use $\alpha = 0.05$, What is your Conclusiem ?

$H_0 : p = 0.20$

$H_a : p \neq 0.20$

$n = 400$

$P_0 = 0.20$

$\bar{P} = 0.175$

$\alpha = 0.05$

random Sample

large enough

• Test statistic :-

$$Z = \frac{\bar{P} - P_0}{\sqrt{\frac{P_0(1-P_0)}{n}}} = \frac{0.175 - 0.20}{\sqrt{\frac{0.20(1-0.20)}{400}}}$$

$Z = -1.25$.

$\alpha = 0.05$

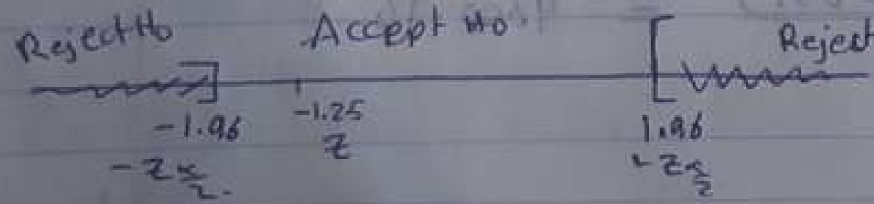
$\frac{\alpha}{2} = 0.025$

t table, df = ∞

$Z_{\frac{\alpha}{2}} = 1.96$

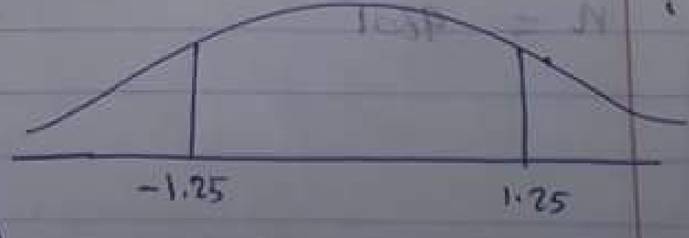
* critical values :- $-Z_{\frac{\alpha}{2}}$, $+Z_{\frac{\alpha}{2}}$

→ critical value :- -1.96, $+1.96$

• 

Accept Ho ($\alpha = 0.05$)

$p = 0.20$ ($\alpha = 0.05$).



• p-Value = area in uppr tail
= + area in lovertial
= $2(1 - 0.8944)$
= 0.2112.

$\alpha = 0.05$

• p-value > α

Accept H_0 ($\alpha = 0.05$)

With signefical $\alpha = 0.05$ we accept Ho that is we accpect that the population porportion the $p = 0.20$.