فريق أوميغا الأكاديمي يقدم لكم:

دفتر مبادئ الإحصاء

للدكتورة أمل الحلو

(Statistical Principle

Chapter 1

Data Organization تنظيم البيانات

جدول تكراري فئات:

1 - Frequency table without classes.

Organise the following data into frequency table without classes?

Ex A O AB B A O A B AB A O B بيانات نوعية

Solution

* جدول تكراري بدون فئات

* هنا الترتيب غير مهم بالنسبة للفئات

Categorical Data OR Qualitative Data

Blood Type	Freq.
A	4
AB	2
B	3
O	3
	12

Ex 1,0,1,2,3,0,0,5,5,1,2,3. بيانات كمية أو رقمية

* في هذا الجدول نهتم بالترتيب لأن البيانات عددية

* يجوز حذف هذه (لأن تكرارها صفر)

Numerical OR Quantitaive Data.

grade	Freq.
0	3
1	3
2	2
3	2
4	0
5	2

But as quantity (كمية) of datas become Large, grouping into classes become necessery to make data more readable (أسهل قراءة).

2- Frequency table with classes

This table list classes of value along with their Frequency. Based on the following procedure for Constructing the table.

1. Decide on the number(*) of classes [K]

$k = [\sqrt{n}]$ عدد الفئات ونُعطى عادة في السؤال

2. Classes must be both mutually exclusive and all inclusive.

أن تحتوي هذه الفئات على جميع العناصر المعطاة في السؤال بدون حدوث تداخل في الفئات.

3. All intervals have to have same width (Length)

يجب أن يتساوى طول الفئات و عرضها (أي احتواء الفئات على عدد متساوي من العناصر).

Ex

Classes	F	Classes	F
0-2	8	0-1	6
3-5	4	2-5	6
✓ (True)	12	X (False)	12

← والسبب هو اختلاف طول الفترة في الفئة الأولى عنها في الثانية

Ex

Classes	F	Classes	F
0-③	2	0-2	2
③-5	1	3-5	1
X (False)	3	✓ (True)	3

← Their is overlapping لأن هناك تداخل بالفئات

$$\boxed{\text{Width}(W) = \frac{Max - Min}{K}}$$

round up ↑ to have the same accuracy Unite of the observation

Max القيمة العظمى في المشاهدات

Min القيمة الصغرى في المشاهدات

k عدد الفئات

* يجب أن يكون ناتج W نفس عدد منازل Max ، Min أو مساوي لأقلها منزلة
(بحيث a.u لِـ W تماثل a.u للقيم المدخلة)

Accuracy Unite (a.u) وحدة الدقة

* تعتمد على أصغر خانة عشرية

1) 3 , 7 , 9 , 5 ⇒ au = 1
2) 3.1 , 7 , 4 , 8 ⇒ au = 0.1
3) 4.9 , 5 , 9 ⇒ a.u = 0.1
4) 2.01 , 7.95 , 3 ⇒ a.u = 0.01
5) 2.99 , 7.9 , 3.09 ⇒ a.u = 0.01
6) 1050 , 7.7 , 0.91 , 5.73 ⇒ a.u = 0.01

Ex 3 , 7 , 9 , 10 >>> a.u = 1
W = 7.13 ⇝ round up ↑ , W = 8

* a.u لل W تماثل a.u للمشاهدات.

Ex 3.1 , 7.3 , 10 , 11 , 15 , 20 >>> a.u = 0.1
W = 7.01 ⇝ round up ↑ , W = 7.1

ملاحظات عامة

١- يجب أن تدخل جميع المشاهدات في الجدول التكراري دون ادخال أي رقم آخر.
٢- لا يجوز التقريب العادي في مادة الإحصاء فلا تقريب 7.01 ⇜ 7 X

grade	Freq
0	5
1	3
2	1
3	1
4	0
5	1
المجموع	11

grade	Freq
0-2	9
3-5	2
المجموع	11

Ex Construct a freq. table with 7 classes ? (k)

50	40	41	12	11	7	22	44	28	21
19	23	32	51	54	42	Max 88	41	78	56
72	56	12	Min 2	69	30	80	56	29.	33
46	31	39	20	18	29	34	54	73	22
36	39	30	62	54	67	39	31	53	44

① $W = \frac{Max - Min}{k} = \frac{88 - 7}{7} = 11.57$ round up ↑

$W \approx 12$

② (a , b)

↳ Lower class limit → upper class limit

Classes	F
7 − 18	6
19 − 30	10
31 − 42	13
43 − 54	8
55 − 66	5
67 − 78	6
79 − 90	2
	50

③ Sample size = 50 = n ⟶ $n = \sum F$

④ relative Frequency (r.f) ⟶ $r.f = \frac{F}{\sum F} = \frac{F}{n}$

⑤ Mid-point (x) = $\frac{Lower + upper}{2}$

⑥ Cumulative Freq, تكرار تراكمي

تابع ⟶

Class	F	C.F
	6	6
	10	16
	13	29
	8	37
	5	42
	6	48
	2	50
	50	

الجدول كاملاً

Classes	F	mid point (x) منتصف الفئة (المركز)	actual class limit حدود الفئات الفعلية	relative Freq. (r.f)	Cumulative Freq. (c.F)
7 - 18	6	12.5	6.5 - 18.5	6/50	6
19 - 30	10	24.5	18.5 - 30.5	10/50	16
31 - 42	13	36.5	30.5 - 42.5	13/50	29
43 - 54	8	48.5	42.5 - 54.5	8/50	37
55 - 66	5	60.5	54.5 - 66.5	5/50	42
67 - 78	6	72.5	66.5 - 78.5	6/50	48
79 - 90	2	84.5	78.5 - 90.5	2/50	㊿
	50			50/50 = 1	

آخر class هو مجموع التكرار الكلي

Actual class Limit الحدود الفعلية

actual Lower Limit =
Lower Limit − 0.5

actual upper limit =
upper Limit + 0.5

* إضافة أو طرح 0.5 لا تؤثر في تداخل القيم مع بعضها لعدم وجود تكرارات للفئات الجديدة

سؤال جدول تكراري فيه 5 فئات ومجموع التكرارات 30 والتكرار التراكمي للفئة الرابعة 20.

Find the relative Freq. (r.f) for the 5th class?

Classes	F	Cf
(-)		
(-)		
(-)		
(-)		20
(-)	10	30
	30	

Solution is $\frac{10}{30} = \frac{1}{3}$

Representing data graphically

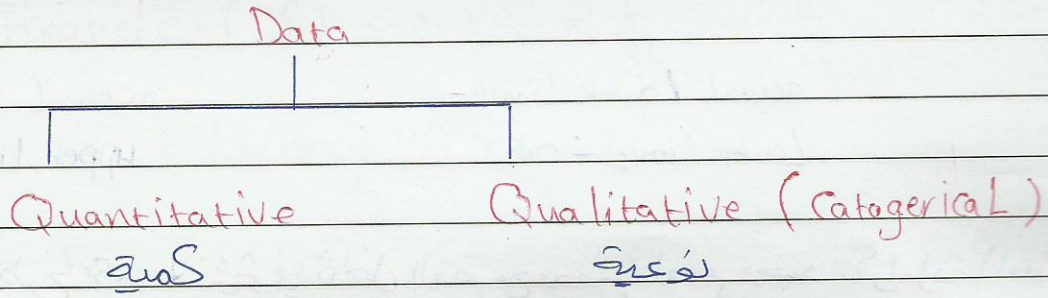
1- Bar charts

Bar charts can used to represent all Kinds of data

Blood Type	F
A	5
B	3
AB	2
O	1

* تستخدم لتمثيل جميع أنواع البيانات سواء أكانت نوعية أو كمية

* الرسم يكون على شكل خطوط عريضة



2- Histogram

Histogram Can used to Represent Quantitative data Just

* Relationship between actual boundries and Freq; / r.f

sketch the histogram for the table below (Construct a histogram?)

Classes	F	actual Class boundries
6-9	3	5.5 - 9.5
10-13	2	9.5 - 13.5
14-17	1	13.5 - 17.5
18-21	4	17.5 - 21.5

- How many observation are Less than 13.5? 5 obs.
- How many observation are between (b/w) (9.5 - 13.5)? 2 ob
- How many observation are between (5.5 - 17.5)? 6 obs.

3- Polygon

Polygon Can used to Represent Quantitative data only.

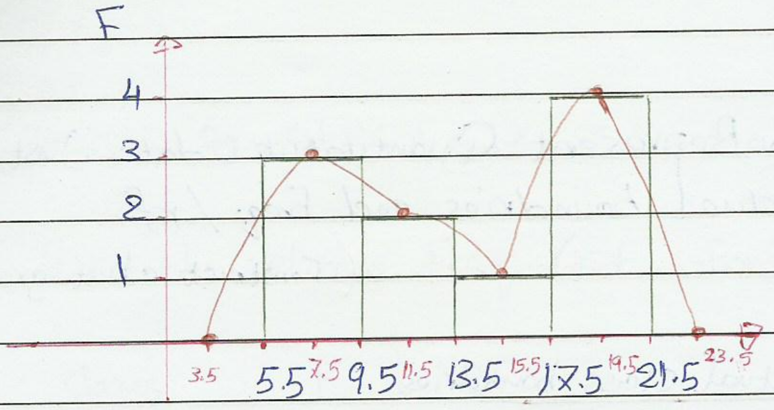
* Relationship Freq. or r.f with mid point.

ملحوظات

- إذا كانت رسمة Histogram موجودة ونريد رسم Polygon نستعمل الرسمة الموجودة لطريقة Histogram كما هو موضح بالشكل (A)

- منحنى Polygon مغلق لذا لابد من افتراض فترة قبل أول فترة موجودة وكذلك فترة بعد آخر فترة موجودة ويكون تكرار هاتان الفترتان صفراً.

Closed curve ; منحنى مغلق

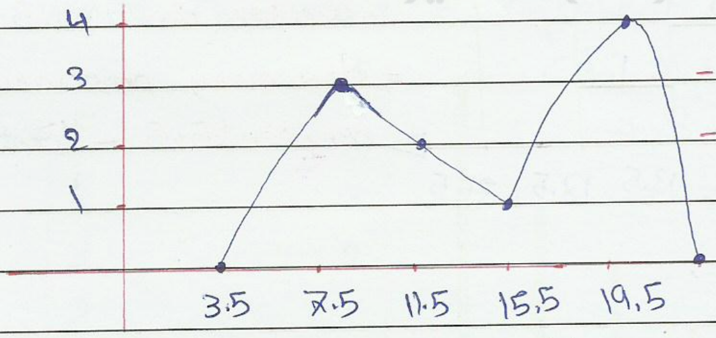


شكل (A)

في حالة عدم توفر رسم لطريقة Histogram نستخدم الجدول

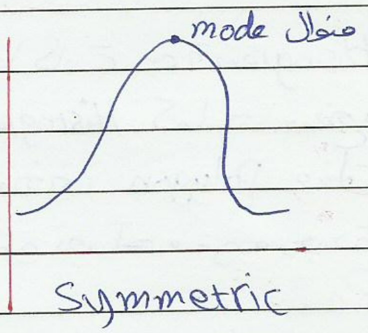
Sketch the Polygon of the following ?

	Classes	F	X (mid point)
		0	3.5
	6-9	3	7.5
+4	10-13	2	11.5
+4	14-17	1	15.5
+4	18-21	4	19.5
		0	23.5

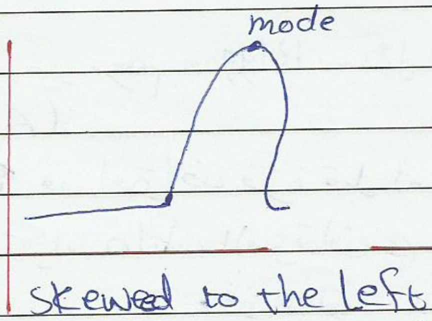


- it has 2 mode منوال
- slightly symmetrical
متماثل لكن ليس تماماً

graph
→ symmetrical متماثل
→ a symmetrical غير متماثل

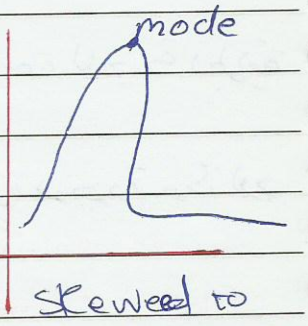


Symmetric



skewed to the left

*حسب الذيل إلى أي جهة



Skewed to the right

عدديًا البيانات تلخيص

Summarising the data numerically

1- Measures of central tendency مقاييس النزعة المركزية

- a. average or mean الوسط الحسابي
- b. mode المنوال
- c. median الوسيط
- d. percentiles المئينات (تستخدم عندما تتجمع البيانات في جهة واحدة)

2- Measures of Variation مقاييس التشتت

- a. Range = Max − Min المدى
- b. Inter Quartile Range (IQR) المدى الربيعي
- c. Variance (S^2) التباين
- d. Standard deviation (S) الانحراف المعياري

Mean / Average

$X : X_1 , X_2 , X_3$ عبارة عن المشاهدات

$\sum X = X_1 + X_2 + X_3$ مجموع المشاهدات

$n = 3$ # of observation عددهم

$\sum X^2 = X_1^2 + X_2^2 + X_3^2$

$$\boxed{\bar{X} = \frac{\sum X}{n}}$$ يستخدم فقط اذا كانت data مفرودة

$\hookrightarrow$ X-bar تُقرأ

Ex ① 5, 6, 7, 3 Find the mean?

$$\bar{X} = \frac{\sum X}{n} = \frac{5+6+7+3}{4} = \frac{21}{4} = 5.25$$

② 7, 3, 11 Find the mean?

$$\bar{X} = \frac{\sum X}{n} = \frac{7+3+11}{3} = \frac{21}{3} = 7$$

Ex Find the mean?

X	F
3	50
2	20
1	100
	170

$$\boxed{\bar{X} = \frac{\sum(XF)}{n}} \quad , \quad n = \sum F$$

تستخدم اذا كانت البيانات مرتبة في جدول

$$\bar{X} = \frac{3(50) + 2(20) + 1(100)}{170}$$

$$\bar{X} = \frac{290}{170} = \frac{29}{17} \approx 1.706$$

Ex Compute the mean?

Classes	F	X	XF
0-4	2	2	4
5-9	1	7	7
10-14	3	12	36
15-19	4	17	68
	10		115

$$\bar{X} = \frac{\sum(XF)}{n}$$

$$\bar{X} = \frac{115}{10} = 11.5$$

Median

Def It is the middle number when observation are arranged in an increasing order.

هو القيمة التي تقع في منتصف بيانات مرتبة ترتيب تصاعدي

Case 1 if (n) is odd فردي

median is the middle number

Case 2 if (n) is even زوجي

median is the average of the middle two number.

Ex ① 5 , 1 , 3 , 10 , 11 Find the Median?

1, 3, 5, 10, 11 $\rightsquigarrow$ Median = 5

② 5, 1 , 3 , 10 Find the median?

1, 3, 5, 10 $\rightsquigarrow$ Median $= \frac{3+5}{2} = \frac{8}{2} = 4$

③ 1 , 3 , 3 , 10 Find the median?

$\rightsquigarrow$ Median $= \frac{3+3}{2} = 3$

Mode

Def. It is the value that occurs with greatest Freq among the other observation. القيمة الأكثر تكراراً ما بين المشاهدات

Ex ① 2, 5, 7, 2, 3, 7, 9, 7 Find the mode?

The mode = 7

② 2, 2, 2, 2 Find the mode?

No mode, because no other observation.

③ 1, 3, 5, 3, 5, 4 Find the mode?

there are two mode → mode 1 = 3
→ mode 2 = 5

Ex ①

X	F
2	5
3	5
4	5

Find the Mode?

No mode.

②

X	F
2	5
3	1
4	7

mode = 4

③

X	F
2	7
3	1
4	7

mode1 = 2

mode 2 = 4

Ex

classes	F	mid Point
(0-4)	3	2
(5-9)	2	7
(10-14)	3	10

Find the mode?

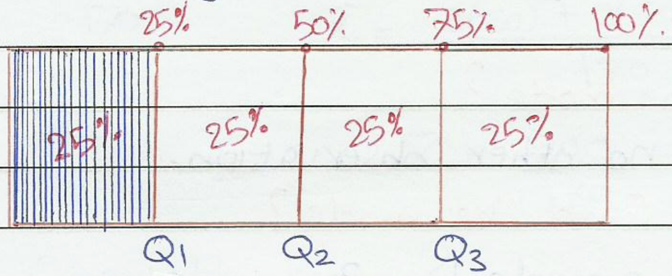
نجد المركز لكل فئة أولاً

mode 1 ≈ 2

mode 2 ≈ 12

Percentiles *مهم جداً* ملتوية

<u>Def</u> Percentiles are useful if data is badly skeweed.
Let $X_1, X_2, \ldots, X_n$ be a set of measurment arranged in increasing order



Q : Quartile

Q1 : 25th percentile ربع أول
Q2 : 50th percentile = ربع ثاني = (median)
Q3 : 75th percentile ربع ثالث

<u>Def</u> Let $0 < P < 100$ the pth percentile is anumber X S.t p% of all measurments fall <u>below</u> تحت the P-th percentile and (100-P)% Fall <u>above</u> فوق it.

هي القيمة التي يوجد أسفل منها P% وفوقها (100-P)%

Procedure

① order data from smallest to Largest.
② determine the Location موقع $L = np$, n : حجم العينة
③ - If (L) is integer عدد صحيح , $X_{pth} = \dfrac{\text{value}(L) + \text{value}(L+1)}{2}$.
- If (L) is Fraction كسر , then <u>round it up</u> ↑ to the next integer Value.

<u>Ex</u> Given $\overset{1}{2}, \overset{2}{5}, \overset{3}{8}, \overset{4}{10}, \overset{5}{11}, \overset{6}{14}, \overset{7}{17}, \overset{8}{20}$

Find the ① 30th percentile
② 50th percentile
③ 75th percentile

ملحوظة: البيانات هنا مرتبة ترتيباً تصاعدياً

$n=8$

① $P = \frac{30}{100} = \frac{3}{10}$

$L = np$

$L = 8\left(\frac{3}{10}\right) = 2.4$ round it up ↑

$L = 3$

$X_{30} = \boxed{8} \rightsquigarrow$ أي الرقم الموجود في الرتبة ٣٠ هو 8

② $P = \frac{50}{100}$ OR $\frac{1}{2}$

$L = np$

$L = 8\left(\frac{1}{2}\right) = 4$

$X_{50} = \frac{\text{Value}(4) + \text{Value}(5)}{2} = \frac{10+11}{2} = \boxed{10.5}$

No round up for X_{p}
Just for Location

③ $P = \frac{75}{100}$

$L = np$

$L = 8\left(\frac{75}{100}\right) = 6$

$X_{75} = \frac{(L_6) + (L_7)}{2} = \frac{14+17}{2} = \boxed{15.5}$

Ex. لديك البيانات التالية [وين ما بتكون البيانات مرتبة (عموديًا أو أفقيًا) نعتمد ذلك] Note

مرتبة ↓	18	31	43	47	54	58	66	73
	21	32	44	48	55	61	67	77
	21	33	45	49	55	62	69	80
	27	35	45	52	56	63	70	81
L_5 →	29	41	46	54	57	65	71	82 → $L=40$

∴ البيانات مرتبة تصاعديًا وحجم العينة $(n) = 40$

Determiane ① 25th percentile (P_{25}) ③ 75th percentile (P_{75})
② 50th percentile (median) ④ 73 th percentile (P_{73})

① $L = np$, $P = \frac{25}{100} = \frac{1}{4}$

$L = 40\left(\frac{1}{4}\right) = 10$

$$X_{25} = \frac{(L_{10}) + (L_{11})}{2} = \frac{41+43}{2} = \frac{84}{2} = \boxed{42}$$

② $L = np$, $P = \frac{50}{100} = \frac{1}{2}$ * تعريف آخر للـ median أنه هو 50th Percentile

$L = 40\left(\frac{1}{2}\right) = 20$

$$X_{50} = \frac{(L_{20}) + (L_{21})}{2} = \frac{54+54}{2} = \boxed{54}$$

③ $L = np$, $P = \frac{75}{100} = \frac{3}{4}$

$L = \overset{10}{\cancel{40}}\left(\frac{3}{\cancel{4}}\right) = 30$

$$X_{75} = \frac{(L_{30}) + (L_{31})}{2} = \frac{65+66}{2} = \boxed{65.5}$$

④ $L = np$

$L = 40\left(\frac{73}{100}\right) = 29.2$ round it up ↑

$L = 30$

$X_{73} = \boxed{65}$

Ex

X	F	(L) C.f
5	2	2
7	3	5
9	3	8
11	4	12
	12	

Find ① 25th percentile
② 50th percentile
③ 75th percentile
④ 70th percentile

* حتى لو كان جدول نتأكد من أن البيانات مرتبة

* C.f = 2 ← أي أنها تحتوي كلاً من L_1 و L_2

* C.f = 5 ← أي أنها تحتوي كلاً من L_3 و L_4 و L_5 وهكذا

* في مثل هذا الجدول نستطيع فرد البيانات بالشكل التالي ثم نكمل الحل المعتاد

5, 5, 7, 7, 7, 9, 9, 9, 11, 11, 11, 11
1 2 3 4 5 6 7 8 9 10 11 12

① $L = np$, $P = \frac{25}{100} = \frac{1}{4}$

$L = 12\left(\frac{1}{4}\right) = 3$

$$X_{25} = \frac{(L_3) + (L_4)}{2} = \frac{7+7}{2} = \textcircled{7}$$

② $L = np$, $P = \frac{50}{100} = \frac{1}{2}$

$L = 12\left(\frac{1}{2}\right) = 6$

$$X_{50} = \frac{(L_6) + (L_7)}{2} = \frac{9+9}{2} = \textcircled{9}$$

③ $L = np$, $P = \frac{75}{100} = \frac{3}{4}$

$L = 12\left(\frac{3}{4}\right) = 9$

$$X_{25} = \frac{(L_9) + (L_{10})}{2} = \frac{11+11}{2} = \textcircled{11}$$

④ $L = np$, $P = \frac{70}{100} = \frac{7}{10}$

$L = 12\left(\frac{7}{10}\right) = 8.4$ round it up ↑ to become $L = 9$

$$X_{70} = \text{Value}\,(L_9) = \textcircled{11}$$

Ex

class	F	C.F
4 – 8	3	3
9 – 13	2	5
14 – 18	4	9
19 – 23	4	13

Find — هام جداً ومتكرر

① P_{60}

② P_{50} (Median)

③ Percentage of data that are greater than 16.

نسبة البيانات اللي أكبر من القيمة 16

Note : في حالة الجداول التكرارية لفئات لا نستخدم Location / round up

بحيث نركز على C.F & upper class limit

① 60th percentile

L = np

$L = 13\left(\frac{60}{100}\right) = 7.8$ don't round it up

actual upper	C.F
8.5	3
13.5	5
$\boxed{x}$	$\boxed{7.8}$
18.5	9
23.5	13

$$\frac{7.8-5}{9-5} = \frac{x-13.5}{18.5-13.5}$$

$$\frac{2.8}{4} = \frac{x-13.5}{5}$$

$\frac{2.8}{4}(5) + 13.5 = x$ OR $\frac{4}{4}(x-13.5) = \frac{5(2.8)}{4}$

$x = 17$

$x - 13.5 = \frac{5}{4}(2.8)$

$x = \frac{5}{4}(2.8) + 13.5$

$x = 17$

وهذا يعني: أن 7.8 من المشاهدات أقل من 17

أو يعني أن 60% من المشاهدات تقع أسفل 17 و 40% من المشاهدات تقع فوق 17

② Median

L = np

$L = 13\left(\frac{1}{2}\right) = 6.5$

upper limit	C.F
13.5	5
x	6.5
18.5	9

$$\frac{6.5-5}{9-5} = \frac{x-13.5}{18.5-13.5}$$

$x = 15.375$

③ أولاً : نحسب نسبة البيانات التي هي أقل أو تساوي 16

ثم نطرح من 100 .

actual upper	C.F
8.5	3
13.5	5
16	x
18.5	9
23.5	13

* $\frac{x-5}{9-5} = \frac{16-13.5}{18.5-13.5}$

* $\boxed{x = 7}$ * of obs. ≤ 16

* $\frac{7}{13}(100\%) = 53.8\%$

* $100\% - 53.8\% = 46.2\%$

Ex

x	F	C.F
2	5	5
3	2	7
4	7	14

Find median?

نتأكد أن البيانات مرتبة

$n=14$, $P = \frac{1}{2} = \frac{50}{100}$

$L = np$

$L = 14\left(\frac{1}{2}\right) = 7$

$$\text{median} = \frac{\text{Value}(L_7) + \text{Value}(L_8)}{2}$$

$$= \frac{3+4}{2}$$

$$\boxed{3.5}$$

ملاحظة : القيم الشاذة (Extreme value) (إما الكبيرة جداً أو الصغيرة جداً) لا تؤثر في كل من median و mode إنما تؤثر فقط في Mean

Ex 1, 3, 3, 3, 5

$\bar{x} = \frac{1+3+3+3+5}{5} = 3$

Median = 3

Mode = 3

Ex 100, 3, 3, 5

3, 3, 5, 100

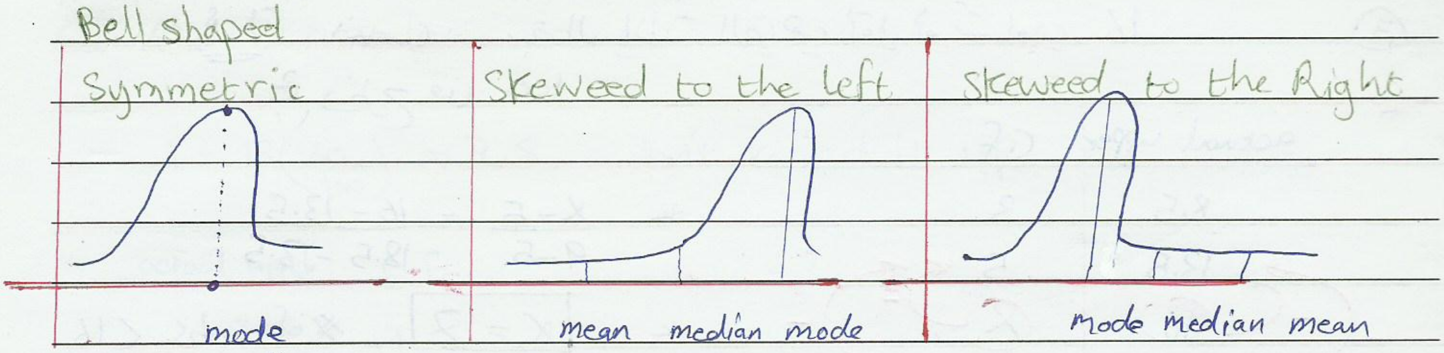
$\bar{x} = \frac{3+3+5+100}{4} = 22.8$

Median = $\frac{3+3}{2} = 3$

Mode = 3

* نلاحظ أن Mean لا يستخدم عندما تكون هناك قيم شاذة

Bell shaped


Symmetric | Skeweed to the left | Skeweed to the Right

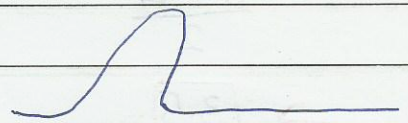
mean < median < mode   mode < median < mean

← إذا mean على اليسار يكون الرسم ذيله ملتوي إلى اليسار

Mode : أكثر القيم تكراراً (القمم)
Median : في المنتصف
Mean : يوجد في المنطقة التي توجد فيها معظم المشاهدات

Ex Mean = 3 , Median = 5 , mode = 5.5   كيف تكون الرسمة ؟
mean < median < mode
Skeweed to Left

Ex


Skeweed to Right

mean = 5 , mode = 0 , median ?
ⓐ 7   [✓ ⓑ 1]   ⓒ 0   because median always between mode & mean.

Ex mode = 10 , median = 2 , mean = 3
mean < median < mode
Skeweed to the Left.

Ex 15 و 13 و 12 و 17 يوجد تجانس
15، 30 و 70 و 100 لا يوجد تجانس
150، 140 و 144 و 200 لا يوجد تجانس
100 و 100 و 100 و 100 Variation = 0 لا يوجد تشتت (تجانس تام)

Range

Ex 2 و 3 و 5 و 4 و 32 , find Range?

Range = Max − Min

$32 - 2 = \textcircled{30}$

Inter Quartile Range (IQR)

$IQR = Q_3 - Q_1$

Q_3 : 75th percentile
Q_1 : 25th percentile

Ex 2, 3, 4, 5, 32 , find IQR?

Q_1 : $L = np$

$L = 5\left(\frac{25}{100}\right) = 1.4$ round it up ↑ $= 2$

$Q_1 = 3$

Q_3 : $L = np$

$L = 5\left(\frac{25}{100}\right) = 3.9$ round it up ↑ $= 4$

$Q_3 = 5$

$IQR = Q_3 - Q_1 = 5 - 3 = \textcircled{2}$

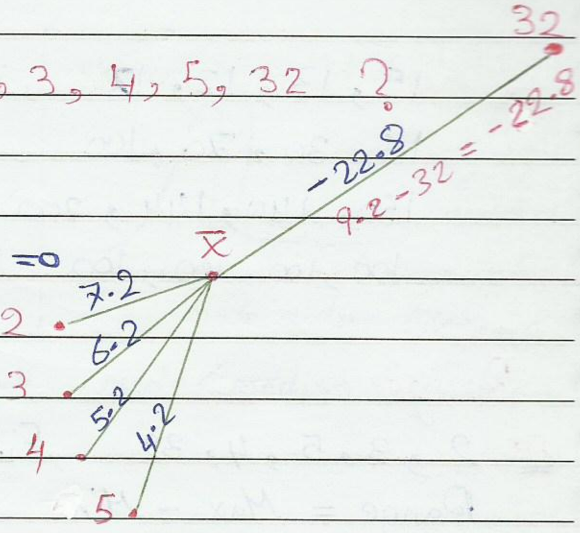
ملحوظات :

- IQR يستخدم للتخلص من القيم الكبيرة جداً أو الصغيرة جداً لأن وجدة
- IQR نحسب Range بعيداً عن الأطراف بعد التأكد من أن الـ data مرتبة

Variance and Standard deviation

Ex Find the variance for 2, 3, 4, 5, 32 ?

$\bar{X} = \frac{2+3+4+5+32}{5} = 9.2$



$\sum(\bar{X}-X) = 7.2+6.2+5.2+4.2-22.8 = 0$

أي أننا لانستفيد منها

$\sum(\bar{X}-x)^2 = 654.8$

$$\boxed{① \; S^2 = \frac{\sum(\bar{X}-X)^2}{n-1} \quad \text{OR} \quad ② \; S^2 = \frac{\sum(X-\bar{X})^2}{n-1}}$$

$(X-\bar{X})^2$
$(2-9.2)^2$
$(3-9.2)^2$
$(4-9.2)^2$
$(5-9.2)^2$
$(32-9.2)^2$

$S^2 = 163.7$ ناتج من أرقام مربعة

$S^2 > 0$

$\boxed{S = \sqrt{S^2}} = S > 0$

$\boxed{S = 12.79}$

لأنه +

Ex. 1, 3, 5, Find S^2 & S ?

$$\boxed{③ \; S^2 = \frac{\sum X^2 - n\bar{X}^2}{n-1}}$$

هذا القانون أسهل استخداماً من القانونين ١, 2

	X^2
1	1
3	9
5	25
	35

$\bar{X} = \frac{1+3+5}{3} = \frac{9}{3} = 3$

$S^2 = \frac{35-3(3)^2}{3-1} = \frac{35-27}{2} = \frac{8}{2} = 4$

$S^2 = 4$, $S = 2$

Ex Compute the standard deviation ?

X	F	$X^2 \cdot f$	$X \cdot f$
1	3	1 (3)	3
3	2	$3^2(2)$	6
5	1	$5^2(1)$	5
	6	46	14

① $\bar{X} = \frac{\sum Xf}{n} = \frac{14}{6} = \frac{7}{3}$

② $\boxed{S^2 = \frac{\sum(X^2 f) - n\bar{X}^2}{n-1}}$ قانون رقم ٤

← هذا القانون هو نفس قانون ③ فقط

تم إضافة F حيث أن الجدول أصبح فيه تكرارات وهو قانون هام

$$S^2 = \frac{46 - 6(\frac{7}{3})^2}{6-1}$$

$S^2 = 2.67$

$S = 1.633$

Ex

	f	mid point X	X.f	$X^2 f$
4 – 8	3	6	6 (3)	$(6)^2 3$
9 – 13	2	11	11 (2)	$(11)^2 2$
14 – 18	4	16	16 (4)	$(16)^2 4$
19 – 23	4	21	21 (4)	$(21)^2 4$
	13		188	3138

Compute the variance & Standard deviation ?

عند إيجاد (Standard deviation, Variance , mean) لجداول تكرارية

نفتكر نلجأ الى الـ mid point

$$\bar{X} = \frac{\sum Xf}{n} = \frac{188}{13} = 14.4615 \approx 14.462$$

$$S^2 = \frac{\sum(X^2 F) - n\bar{X}^2}{n-1} = \frac{3138 - 13(14.462)^2}{13-1} = 34.921$$

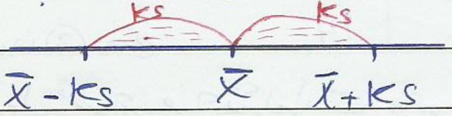
$S = 5.909$

Checheves Rule (Not Symmetric)

For any Sample or population of measurements the proportion measurement that fall within KS of the mean is at least $1 - \frac{1}{k^2}$, $k > 1$ KS : K X (standrad deviation)

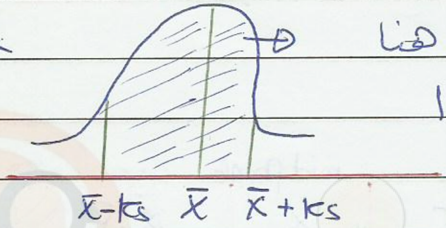
على الأقل

* لأي عينة أو مجتمع نسبة القياسات التي تقع على بعد ks من الوسط الحسابي على الأقل $1 - \frac{1}{k^2}$



$(\bar{x} - ks \; , \; \bar{x} + ks)$

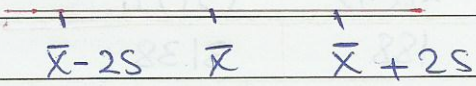
* نسبة البيانات المحصورة هنا هي على الأقل $1 - \frac{1}{k^2}$



* وتكون أكبر مقارنة بالأطراف

① $\boxed{k=2}$ $\rightsquigarrow$ $1 - \frac{1}{k^2} \Rightarrow 1 - \frac{1}{4} = 0.75 \times 100\%$

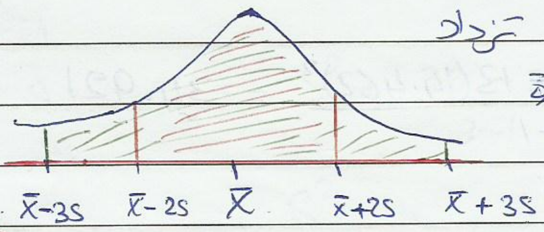
$= \boxed{75\%}$



② $\boxed{k=3}$ $\rightsquigarrow$ $1 - \frac{1}{3^2} = 1 - \frac{1}{9} = 0.889 \times 100\%$

$= \boxed{88.9\%}$

* لاحظي أنه كلما أزدادت قيمة k تزداد النسبة وتزداد المساحة المحصورة



③ $\boxed{k=4}$ $\rightsquigarrow$ $1 - \frac{1}{k^2} = 1 - \frac{1}{16} = 0.932 \times 100\%$

$= \boxed{93.2\%}$

Ex the Score on acertain test have mean = 74 and sd = 4
Give an interval that Containt at least 75% of the scores.

← الكلمة المفتاحية للحل

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

$$\boxed{k=2}$$

Lower limit = $\bar{x}-ks = 74-2(4) = 66$

upper limit = $\bar{x}+ks = 74-2(4) = 82$

$\therefore (66, 82)$

Ex Suppose $\bar{x}=60$, $S=8$, $n=100$

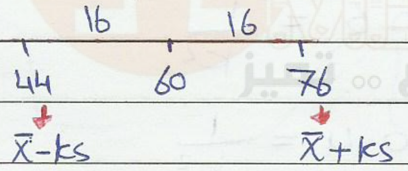
1- At least how many observation are b/w 44 & 76 ?

2- at most how many observation are less than 36 and greater than 84 ?

3- Find an interval that Contains at least 60% of the obs.?

Solution

①



at least = Chechevs

$ks = 16$

$k(8) = 16 \rightsquigarrow \boxed{k=2}$

$$1-\frac{1}{k^2} \Rightarrow 1-\frac{1}{4} = 75\% \text{ of the obs.}$$

لكنه طلب عدد المشاهدات

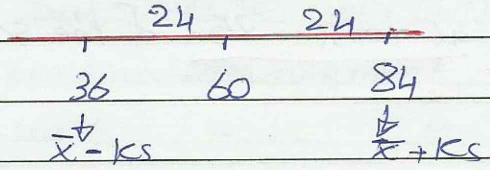
$$\frac{75}{100}\times 100 = \boxed{75 \text{ obs}}$$

ملاحظات :

- نعمل Round up ↑ للنسبة اذا كانت 57.5 مثلاً لتصبح 58 عند ذكر كلمة inside OR at least .

- ٧٥ مشاهدة تقع بين ٤٤ - ٧٦ أما ٢٥ مشاهدة المتبقية توزع على الأطراف بطريقة غير معروفة .

②



$ks = 24$

$k(8) = 24 \rightsquigarrow \boxed{k=3}$

the percentage of data b/w $(36-84) \Rightarrow 1 - \frac{1}{9} = 88.9\%$

number of obs. b/w $= \frac{88.9}{100} \times 100 = 88.9$ round it up ↑

↳ at least / inside فيه $= 89$

* of obs. Less than 36 and greater than 84 $= 100 - 89 = \boxed{11 \text{ obs.}}$

③ 60%

$1 - \frac{1}{k^2} = \frac{60}{100} \Rightarrow 1 - \frac{1}{k^2} = 0.6$

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

$0.4 = \frac{1}{k^2}$

$k^2 = \frac{10}{4}$

$k = \sqrt{\frac{10}{4}}, \quad k > 0$

$\boxed{k = 1.5811}$

Lower Limit $= 60 - 1.58(8)$

upper limit $= 60 + 1.58(8)$

interval $(60 - 1.58(8), 60 + 1.58(8))$

Empirical Rule for (Bell shaped) Data (For Symmetrical)

① $\boxed{K=1}$ → Contains at least $\boxed{68\%}$ of the obs. Fall inside the I (Interval).

② $\boxed{K=2}$ → Contains at least $\boxed{95\%}$ of = = = = = .

③ $\boxed{K=3}$ → = = = $\boxed{99\%}$ of = = = = = .

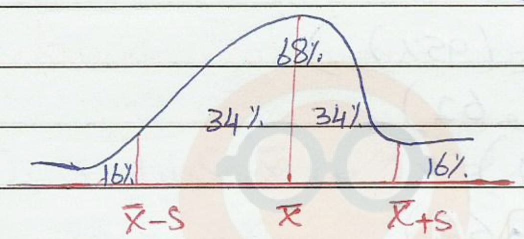
Chercheves → عند عطاء الفترة يكون $\bar{X}$ بالمنتصف

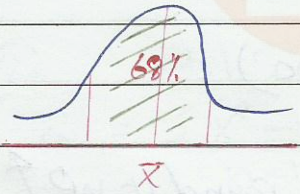
Empirical → ليس من الضروري أن تكون $\bar{X}$ بالمنتصف في الفترة المعطاة لأن الشكل symmetrical

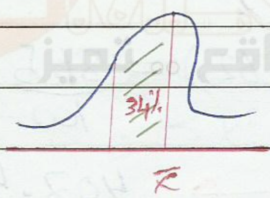
1- $(\bar{X}-S, \bar{X}+S)$ → $\boxed{K=1}$

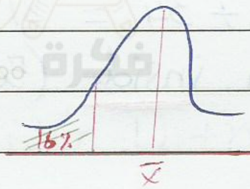
* في هذه الحالة بما أن الشكل متماثل والمسافات متساوية فتتوزع نسبة البيانات على الأطراف بالتساوي.

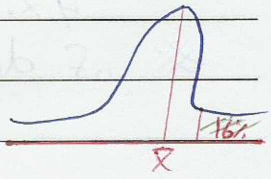
32% ← تتوزع 16% على كل ذيل



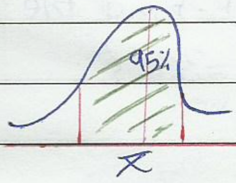


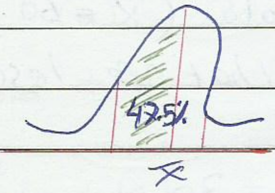


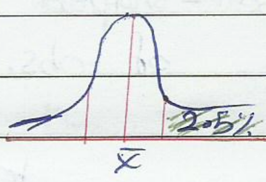


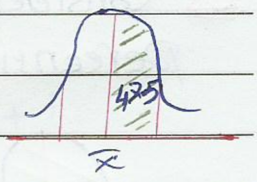


2- $\boxed{K=2}$ $\bar{X}$ بالمنتصف → Symmetrical

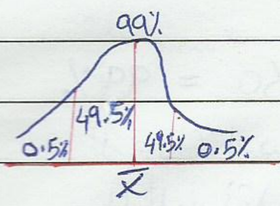




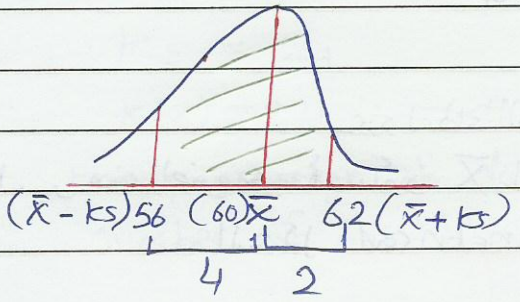




3- $\boxed{K=3}$



Ex Consider bell-shaped (symmetrical) data
Suppose $\bar{x} = 60$, $\delta = 2$, $n = 500$
How many observation are b/w 56, 62 ?



$ks = 4$
$k(2) = 4$
$\boxed{k=2}$

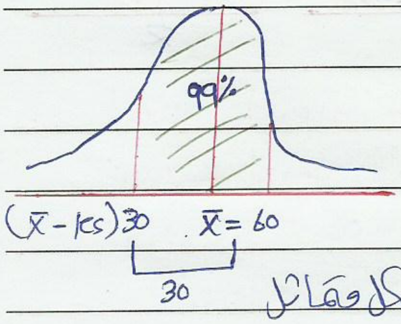
$ks = 2$
$k(2) = 2$
$\boxed{k=1}$

- Percentage of data b/w (56, 60)
$(\bar{x} - 2S , \bar{x}) = 47.5\% \Rightarrow \left(\frac{1}{2}(95\%)\right)$
- Percentage of data b/w (60, 62)
$(\bar{x} , \bar{x} + S) = 34\% \Rightarrow \left(\frac{1}{2}(68\%)\right)$
- Percentage of data b/w (56, 62)
$47.5 + 34 = 81.5\%$

※ of data inside $(56, 62) = \frac{81.5}{100}(500)$

$\therefore$ 407.5 round it up ↑
408 obs.

Ex Consider bell-shaped data $(\bar{x} = 60 , S = 10)$. Find the percentage of obs. that are less than 30 ?

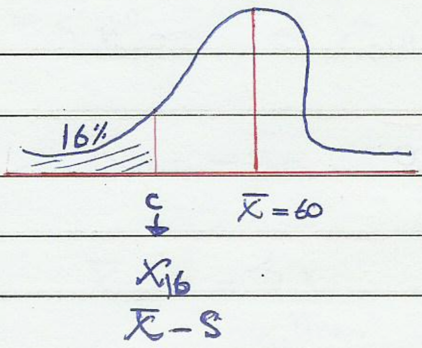


$ks = 30$
$k(10) = 30$
$k = 3$

Percentage b/w 30, 60 = 99%
يبقى ١٪ يتوزع على الذيلين بالتساوي لأن الشكل متماثل
المطلوب أقل من 30 جهة اليسار = $\boxed{0.5\%}$

ملاحظة : إذا لم يذكر السؤال كلمة bell-shaped أو Symmetrical فإن الحل يكون باستخدام Checheves.

Ex suppose the grades are bell-Shaped with mean $=60$ and Standard deviation $=10$. Find the 16th percentile of the grades ?



$K=1 \; (P_{16})$

$X_{16} = \bar{X} - S$

$X_{16} = 60 - 10$

$X_{16} = 50$

Updating Statistical Measures

① نستخرج ΣX من $\bar{X}$ كما يلي :

$$\bar{X} = \frac{\Sigma X}{n} \longrightarrow \boxed{\bar{X} n = \Sigma X}$$

② نستخرج ΣX^2 من S^2 كما يلي :

$$S^2 = \frac{\Sigma X^2 - n\bar{X}^2}{n-1} \longrightarrow \boxed{\Sigma X^2 = S^2(n-1) + n\bar{X}^2}$$

Adding , Deleting and Modification

1- If an observation (a) is added to the data.

$$\left.\begin{array}{l} \Sigma X_{new} = \Sigma X_{old} + a \\ \Sigma X^2_{new} = \Sigma X^2_{old} + (a)^2 \end{array}\right\} n_{new} = n_{old} + 1$$

2- If an obs. (D) is deleted.

$$\left.\begin{array}{l} \Sigma X_{new} = \Sigma X_{old} - d \\ \Sigma X^2_{new} = \Sigma X^2_{old} - (d)^2 \end{array}\right\} n_{new} = n_{old} - 1$$

3- If an obs. (c) is modified to (d)

$$\left.\begin{array}{l} \Sigma X_{new} = \Sigma X_{old} - c + d \\ \Sigma X^2_{new} = \Sigma X^2_{old} - (c)^2 + (d)^2 \end{array}\right\} n_{new} = n_{old}$$

Ex let $\bar{x} = 40$, $S^2 = 100$, $n = 30$

Find $\bar{x}$, S^2 in each of the following ?

1- The obs. 60 and 10 are added

a. $\sum x = n\bar{x}$

$\sum x_{old} = 30(40) = \boxed{1200}$

b. $\sum x^2_{old} = S^2(n-1) + n\bar{x}^2$

$\sum x^2_{old} = 100(29) + 30(40)^2 = \boxed{50900}$

c. $n_{old} = \boxed{30}$

$\boxed{a}$ $\sum x_{new} = \sum x_{old} + 60 + 10$

$\sum x_{new} = 1200 + 70 = \boxed{1270}$

$\boxed{b}$ $\sum x^2_{new} = \sum x^2_{old} + (60)^2 + (10)^2$

$\sum x^2_{new} = 50900 + 3600 + 100 = \boxed{54600}$

$\boxed{c}$ $n_{new} = \boxed{32}$

$$* \quad \bar{x}_{new} = \frac{\sum x_{new}}{n_{new}} = \frac{1270}{32} = 39.68$$

$$* \quad S^2_{new} = \frac{\sum x^2_{new} - n_{new}\bar{x}^2_{new}}{n_{new} - 1} = \frac{54600 - 32(39.68)^2}{32-1} = 135.38$$

2- Delete the obs. 32 from the data / sample ?

a. $\sum x_{new} = \sum x_{old} - 32$

$1200 - 32 = \boxed{1168}$

b. $\sum x^2_{new} = \sum x^2_{old} - (32)^2$

$= 50900 - 1024 = \boxed{49876}$

c. $n_{new} = n_{old} - 1 = 30 - 1 = \boxed{29}$

$$* \quad \bar{x}_{new} = \frac{\sum x_{new}}{n_{new}} = \frac{1168}{29} = \boxed{40.22}$$

$$* \quad S^2_{new} = \frac{\sum x^2_{new} - n_{new}\bar{x}^2_{new}}{n-1} = \frac{49876 - (29)(40.22)^2}{(29-1)} = \boxed{101.7}$$

3- The obs. 45 is modified to 60 ?

a. $\sum X_{new} = \sum X_{old} - 45 + 60 = \boxed{1215}$

b. $\sum X^2_{new} = \sum X^2_{old} - (45)^2 + (60)^2 = \boxed{52475}$

c. $n_{new} = \boxed{30}$

* $\bar{X}_{new} = \frac{\sum X_{new}}{n_{new}} = \frac{1215}{30} = 40.5$

* $S^2_{new} = \frac{\sum X^2_{new} - n\bar{X}^2_{new}}{n_{new}-1} = \frac{52475-(30)(40.5)^2}{30-1} = 112.67$

Pooled Variance & Combined mean

	n	mean	Variance
Sample1	n_1	$\bar{X}_1$	S_1^2
Sample2	n_2	$\bar{X}_2$	S_2^2

The two samples are combine to make one sample

① Combined mean ($\bar{X}_c$)

$$\boxed{\bar{X}_c = \frac{n_1\bar{X}_1 + n_2\bar{X}_2}{n_1+n_2}}$$

② Pooled variance

$$\boxed{S_p^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1-1+n_2-1}} \text{ OR } \boxed{S_p^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}}$$

Ex

	n	$\bar{x}$	sd
Sample1	30	40	8
Sample2	50	30	10

Compute
① Combined mean
② Pooled Variance

Solution :

① $\bar{X}_c = \dfrac{n_1\bar{X}_1 + n_2\bar{X}_2}{n_1+n_2}$

$$\frac{30(40)+50(30)}{30+50} = 33.75$$

② $S_p^2 = \dfrac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1-1+n_2-1}$

$$\frac{(30-1)(8)^2 + (50-1)(10)^2}{30-1+50-1} = 86.62$$

Linear Transformation

$Y = aX + b$

① $\bar{Y} = a\bar{X} + b$

② $Mode(y) = a\,(mode(x)) + b$

③ $Median\,(y) = a\,(Median(x)) + b$

④ $P_{100p}\,(y) = \begin{cases} a\,P_{100p}(x) + b, & a>0 \\ a\,P_{100-100p} + b, & a<0 \end{cases}$

Ex $y = 1 + 2X$ $\qquad\qquad y = 1 - 2X$

$P_{60}(y) = 1 + 2P_{60}(x) \qquad P_{60}(y) = 1 - 2P_{40}(X)$

⑤ $Var(y) = var\,(ax + b)$

$= Var\,(aX) + \overset{zero}{\cancel{Var}}(b)$

$= \dfrac{\sum(ax - a\bar{x})^2}{n-1}$

$= \dfrac{a^2\sum(x-\bar{x})^2}{n-1}$

$S_y^2 = a^2 S_x^2$

⑥ $S_y = \sqrt{a^2 S_x^2} = |a|\,S_x \qquad S^2 \geq 0 \;,\; S > 0$

⑦ $Range(y) = |a|\;Range(x)$

⑧ $IQR\,(y) = |a|\;IQR\,(x)$

Ex $\bar{X} = 3$

$S_x^2 = 4$

$Q_1(x) = 2$

$Q_3(x) = 2$

Range $(x) = 5$

median $(x) = 4$

Mode $(x) = 5$

$Y = 1 + 2X$

① $\bar{Y} = 1 + 2\bar{X} = 1 + 2(3) = 7$

② Mode$(y) = 1 + 2$ mode(x)

$1 + 2(5) = 11$

③ Med$(y) = 1 + 2$ med(x)

$1 + 2(4) = 9$

④ $Q_1(y) = 1 + 2Q_1(x)$

$1 + 2(2) = 5$

⑤ $Q_3(y) = 1 + 2Q_3(x)$

$1 + 2(7) = 15$

⑥ IQR$(y) = Q_3(y) - Q(y)$

$15 - 5 = 10$

أو $|a|$ IQR(x)

$2(Q_3(x) - Q_1(x))$

$2(7 - 2) = 10$

② $S_y = |a| S_x$

$2(2) = 4$

⑧ Range$(y) = |a|$ Range(x)

$2(5) = 10$

⑨ $S_y^2 = a^2 S_x^2$

$4(4) = 16$

$Y = 1 - 2X$

① $\bar{Y} = 1 - 2\bar{X} = 1 - 2(3) = -5$

② Mode$(y) = 1 - 2$ mode(x)

$1 - 2(5) = -9$

③ Med$(y) = 1 - 2$ med(x)

$1 - 2(4) = -7$

④ $Q_1(y) = 1 - 2Q_3(x)$

$1 - 2(7) = -13$

⑤ $Q_3(y) = 1 - 2Q_1(X)$

$1 - 2(2) = -3$

⑥ IQR$(y) = Q_3(y) - Q_1(y)$

$-3 - (-13)$

10

أو $|-2|(5)$

10

⑦ $S_y = |-2| S_x$

4

⑧ Range$(y) = |a|$ Range(x)

$|-2|(5) = 10$

⑨ $S_y^2 = a^2 S_x^2$

$4(4) = 16$

ملاحظة :

مقاييس التشتت لا تختلف إذا كان معامل X (a) موجب أو سالب

اعتبر

Ex Consider a bell-shaped data with $\bar{X} = 60$, $S = 14$

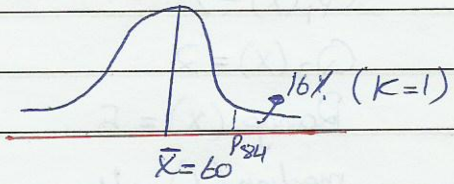
Let $y = 1 - 2X$ Find $P_{16}(y)$?

$Y = 1 - 2X$

$P_{16}(y) = 1 - 2P_{84}(x)$

$1 - 2(74)$

-147



$\bar{X} + KS = P_{84}$

$60 + 1(14) = P_{84}$

$74 = P_{84}$

Ex Let $\bar{y} = 4$, $y = 1 - 3X$, $Q_1(X) = -25$, $IQR(X) = 20$

Find $\bar{X}$, $Q_1(y)$?

* $\bar{y} = 1 - 3\bar{X}$

$4 = 1 - 3\bar{X}$

$3 = -3\bar{X}$

$\boxed{-1 = \bar{X}}$ ✓

* $IQR(X) = Q_3(X) - Q_1(X)$

$20 = Q_3 - (-25)$

$\boxed{Q_3 = -5}$

* $Q_1(y) = 1 - 3Q_3(X)$

$Q_1(y) = 1 - 3(-5)$

$\boxed{Q_1(y) = 16}$ ✓

Chapter 2

- Random Experiment (تجربة عشوائية)
- Sample space (S) (الفضاء العيني) The set of all possible out Comes
- Event (الحدث) < عنصر / مجموعة يُشترط في الحدث أن ينتمي إلى الفضاء العيني
- Impossible event (ϕ / $\{\}$) (الحدث المستحيل)
- Certain event (حدث مؤكد) $\overset{Ex}{=}$ Extreme Value
- Union - Intersection (الاتحاد - التقاطع)
- disjoint events (الأحداث المنفصلة) $A \cap B = \phi$ << لا تحدث في نفس اللحظة >>
- Complement events (الأحداث المكملة لبعضها)
- Independent events (الأحداث المستقلة) << تحدث في نفس اللحظة >>

Some notes :

① Events
  - Simple event حدث بسيط
  - Compound event حدث مركب (يحتوي شروط كثيرة)

② denote the Complement by $\bar{A}$ oR A^c , $A \cap \bar{A} = \phi$, $A \cup \bar{A} = S$

③ مثال على الأحداث المستقلة: رجل عطس في القاعة ، حجة توفيت في الصين ... ! ؟

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

$A = \{2, 4, 6\}$ = set of even #'s

$B = \{1, 3, 5\}$ = set of odd #'s

$C = \{2\}$ = * Less than 3 greater 1

$A \cup B$: all events that are in A or B or Both.

$A \cup B = \{2, 4, 6, 1, 3, 5\}$

$A \cap B$: Set of elements that are in A and B together at the Same time.

$A \cap B = \phi$

$A A C = \{2\}$

④ Sample Space : كل العناصر (فلا نميز المطلوب عن غيره)

event : عناصر معينة (تمييز المطلوب)

Ex

Blood Type	F	rf
A	3	$3/10 = 0.3$
B	2	$2/10 = 0.2$
O	1	$1/10 = 0.1$
AB	4	$4/10 = 0.4$
	10	1

$\frac{F}{n}$ أبسط قوانين الاحتمالات

Equally Likely

$$P(E) = \frac{\# E}{\# S}$$

Ex $S = \{1, 2, 3, 4, 5, 7\}$

$A = \{1, 5, 7\}$

$B = \{2, 4\}$

$C = \{7\}$

① $P(A) = \frac{\# A}{\# S} = \frac{3}{6} = \frac{1}{2}$

② $P(B) = \frac{\# B}{\# S} = \frac{2}{6} = \frac{1}{3}$

③ $P(C) = \frac{1}{6}$

④ $P(A \cap B)$

$A \cap B = \emptyset$

$P(A \cap B) = P(\emptyset) = \frac{\# \emptyset}{\# S}$ = Zero

⑤ $P(A \cup B)$

$A \cup B = \{1, 5, 7, 2, 4\}$

$P(A \cup B) = \frac{5}{6}$

Counting Methods

In solving problems , there are basically three Methods:
① Multiplication principle مبدأ الضرب
② Permutation principle التباديل
③ Combination principle التوافيق

Multiplication principle

IF event 1 Can happen in N ways , and event 2 Can happen in M ways , then the event together Can happen in $(N * M)$ ways.

The same principle applies for more than two events, just multiply the possible ways each event Can happen.

Ex① You are going to make a License plate by choosing 3 letters & then 3 numbers , How many different license plates are possible?
لاحظ أن التكرار هنا مسموح
في هذا المثال نريد تكوين نمرة سيارة بحيث نختار ٣ أحرف تليها ٣ أرقام ، كم طريقة مختلفة يمكن بها تكوين نمرة للسيارة ؟

number $(0-9)=10$, letters $(A-Z)=26$
$\#$ of ways $= (26)^3 \cdot (10)^3$

Ex Same as in Ex 1 except that No letters or numbers may be Used more than once ? نرى أن الاختلاف هو أن التكرار هنا غير مسموح
$\#$ of ways $= 26 \cdot 25 \cdot 24 \cdot 10 \cdot 9 \cdot 8$

Permutations principle قابلة للتمييز

A permutation of a set of distinct object in an arrangement of the objects in a definite order.
ترتيب معين

Permutations → distinguishable object (قابلة للتمييز)
→ indistinguishable object (غير قابلة للتمييز)

Ex 1

There are 12 children in aschool, they are going to stand (يصطف) in a Line (في طابور). How many different lines are possible? (قابلة للتمييز)

→ $12! = 12 \times 11 \times 10 \times 9 \times \dots \times 1$

(هذه أحداث مستقلة (نضربهم ببعض))

OR

→ $_{12}P_{12} = \frac{12!}{(12-12)!} = \frac{12!}{0!} = 12!$

Ex 2

Same as ex1 except, Five of them are going to stand in a line.

→ $12.11.10.9.8$

OR

→ $_{12}P_5 = \frac{12!}{(12-5)!} = \frac{12!}{7!} = \frac{12.11.10.9.\cancel{7!}}{\cancel{7!}}$

* The number of permutation of (n) distinct object taken (r) at atime is given by

$$\boxed{_nP_r = \frac{n!}{(n-r)!} \quad , \quad r \leq n}$$

$0! = 1$, $1! = 1$, المضروب دائماً يكون عدد موجب

① $_{12}P_{12} = \frac{12!}{(12-12)!} = 12!$

② $_{50}P_{20} = \frac{50!}{(50-20)!} = \frac{50!}{30!}$

Permutations involving indistinguishable object غيرقابلة للتميز

Ex1 ABO $\Rightarrow 3!$ OR $P(ABO) = \frac{3!}{1!1!1!} = 6$

ABO , BAO , OBA , OAB , AOB , BOA

Ex2 BOB

BOB , OBB , BBO

what are the different permutations for the word BoB ?

$$\frac{3!}{2!\,1!} = \frac{3!}{2!} = 3$$

* the number of permutation of indistinguishable object equal $= \frac{n!}{n_1!\,n_2!\,n_r!}$

Given a set of (n) , $n_1 =$ object are alike
$n_2 =$ object are alike
$n_3 =$ object are alike

$n = n_1 + n_2 + \dots + n_r$

Ex 3

In How many different ways we can rearrange the word (ATLANTA) ?

A = 3
T = 2
L = 1
N = 1

$$\frac{7!}{3!\,2!\,1!\,1!} = \boxed{\frac{7!}{3!\,2!}} = \frac{7 \cdot \cancel{6} \cdot 5 \cdot 4 \cdot 3 \cdot \cancel{2 \cdot 1}}{\cancel{3!} \cdot \cancel{2!}}$$

يقبل هكذا

$$\frac{7 \cdot 5 \cdot 4 \cdot 3}{1} = 420 \text{ ways}$$

Combinations لا نهتم بالترتيب

are arrangements of elements without any regard (اهتمام) to their order or position.

(في حالة كانت الكرات مرقمة و كنا مهتمين بهذا الترقيم (لا يصلح التوافيق لحلها

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

Ex 1 Find the number of ways a 12 men in aclub can choose 3 men ?

$$\boxed{\binom{12}{3}} = \frac{12!}{(9!)3!}$$

يقبل هكذا في الامتحان

Ex 2 A box contains 3 Black & 2 white balls . 2 balls are drawn from the box without replacement .

① P (Both balls are black) ?

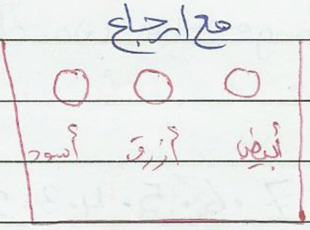
$$P(BB) = \frac{3}{5} \cdot \frac{2}{4}$$

② P(at least one ball is Black) ?

$$P(BW) + P(WB) + P(BB)$$

$$\frac{3}{5} \cdot \frac{2}{4} + \frac{2}{5} \cdot \frac{3}{4} + \frac{3}{5} \cdot \frac{2}{4}$$

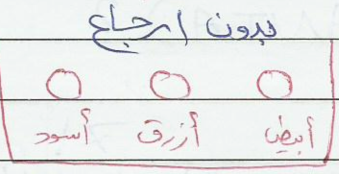
Sampling withreplacement



① P(white | Blue) = $\frac{1}{3}$

② P(white) = $\frac{1}{3}$

Sampling without replacement



① P(white) = $\frac{1}{3}$

② P(Blue white) = $\frac{1}{3} \cdot \frac{1}{2}$

Ex * Find the number of ways a 12 men in a club can choose expresident , president and asecretary

$${}_{12}P_3 = \frac{12!}{(12-3)!} = \frac{12!}{9!} = \frac{12 \cdot 11 \cdot 10 \cdot \cancel{9!}}{\cancel{9!}}$$

1320

Ex 3 We have $\left|\begin{matrix} 3\ W \\ 2\ R \end{matrix}\right|$

$$P(W) = \frac{\# W}{\# S} = \frac{3}{5}$$

$$P(R) = \frac{\# R}{\# S} = \frac{2}{5}$$

- Suppose we drew 2 balls find $P(RW)$?

1 with replacement

$$P(RW) = \frac{2}{5} \cdot \frac{3}{5} = \frac{6}{25}$$

2 without replacement

$$P(RW) = \frac{2}{5} \cdot \frac{3}{4} = \frac{6}{20} = \frac{3}{10}$$

- Suppose we drew 3 balls find $P(RRW)$?

1 with replacement

$$P(RRW) = \frac{2}{5} \cdot \frac{2}{5} \cdot \frac{3}{5} = \frac{12}{125}$$

2 without replacement

$$P(RRW) = \frac{2}{5} \cdot \frac{1}{4} \cdot \frac{\cancel{3}^{1}}{\cancel{3}} = \frac{2}{20} = \frac{1}{10}$$

بما أن الكرات التي تبقت هي بيضاء (جميع)
اذن احتمالية ظهور كرة بيضاء 1

- Suppose We drew 2 balls , then Find

1 P(both balls are of the same color) ?

$$P(RR) + P(WW)$$

$$\rightarrow \text{With rep.} = \frac{2}{5} \cdot \frac{2}{5} + \frac{3}{5} \cdot \frac{3}{5}$$

$$\rightarrow \text{Without rep.} = \frac{2}{5} \cdot \frac{1}{4} + \frac{3}{5} \cdot \frac{2}{4}$$

2] Prob. of (at least one ball is white)?

$P(WR) + P(RW) + P(WW)$ < with rep.
without repl.

OR Prob. (none is white)

$1 - P(RR)$

→ with rep. $= 1 - \left(\frac{2}{5} \cdot \frac{2}{5}\right) = \frac{21}{25}$

→ without rep. $= 1 - \left(\frac{2}{5} \cdot \frac{1}{4}\right) = \frac{18}{20} = \frac{9}{10} = 0.9$

Probability Rules

$(0 \leq P(E) \leq 1)$

1) $P(\bar{A}) = 1 - P(A)$ prob. of Complement

$\rightarrow$ $P(\bar{A} \cup A) = P(S)$

$\rightarrow$ $P(\bar{A}) + P(A) = 1$

2) $P(S) = 1$, ProP. (certain event) = 1

$P(\Phi) = 0$, ProP. (Imposible event) = 0

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

$\rightarrow$ if A, B are disjoint

i.e $A \cap B = \emptyset$

$P(A \cap B) = p(\emptyset) = 0$, then

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

$\rightarrow$ if A, B are independent

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

* حدوث A لايؤثر على حدوث B
وقد يحدثان في نفس الوقت

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

4) $P(A \Delta B) = P(A) + P(B) - 2P(A \cap B)$

Prob. of A or B but not both.

$A \Delta B$: element in A or B, but not both.

5) $P(A \setminus B) = P(A \cap \bar{B})$ تعني ماعدا ⇐ (\ )

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

Probability of A without B

6) $P(A|B) = \dfrac{P(A \cap B)}{P(B)}$, $P(B) \neq 0$ Conditional

* اذا حدث B أولاً فما احتمال حدوث A

Prob. of A condition on B

$\rightarrow$ $P(A \cap B) = P(A|B).P(B)$

OR $P(A \cap B) = P(B|A) \cdot P(A)$ from $P(B|A) = \dfrac{P(A \cap B)}{P(A)}$, P

$\rightarrow$ if A, B are disjoint $P(A|B) = 0$ and

$P(B|A) = 0$

→ if A, B are independent, then

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

$\to P(A|B) = \frac{P(A\cap B)}{P(B)} = \frac{P(A).\cancel{P(B)}}{\cancel{P(B)}} = P(A)$

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

$\to P(B|A) = \frac{P(B\cap A)}{P(A)} = \frac{P(B).\cancel{P(A)}}{\cancel{P(A)}} = P(B)$

⑨ $P(\overline{A}) = P(A)$

⑩ $P(\overline{A\cup B}) = 1 - P(A\cup B)$

$P(\overline{A}\cap\overline{B})$

⑪ $P(\overline{A\cap B}) = 1 - P(A\cap B)$

$P(\overline{A}\cup\overline{B})$

Ex 1 $\boxed{\begin{matrix}3B\\4W\end{matrix}}$ draw 2 balls without replacement

① P(2nd is B | 1st is B)

② P(both balls have the same color)

③ P(at least one black)

① $P(B_2|B_1) = \frac{P(B_1\cap B_2)}{P(B_1)} = \frac{P(B_1B_2)}{P(B_1)}$

$\frac{(\frac{3}{7})(\frac{2}{6})}{(\frac{3}{7})}$

$\frac{2}{6}$

$\frac{1}{3}$

② $P(BB \overset{\cup}{or} WW) = P(BB) + P(WW)$

$(\frac{3}{7})(\frac{2}{6}) + (\frac{4}{7})(\frac{3}{6}) = \frac{3}{7}$

③ $P(BW \text{ or } WB \text{ or } BB) = P(\text{at least one is } B)$

$1 - P(\text{none is } B)$

$1 - P(WW)$

$1 - (\frac{4}{7})(\frac{3}{6})$

$\frac{5}{7}$

Ex 2

$P(A) = 0.2$

$P(B) = 0.7$

$P(A \cap B) = 0.15$, Find

① $P(A \cap \overline{B}) = P(A) - P(A \cap B)$

$0.2 - 0.15 = 0.05$

② $P(\overline{A} \cap \overline{B}) = P(\overline{A \cup B})$

$1 - P(A \cup B)$

$1 - (P(A) + P(B) - P(A \cap B))$

$1 - (0.2 + 0.7 - 0.15) = 0.25$

③ $P(A | B) = \frac{P(A \cap B)}{P(B)} = \frac{0.15}{0.7}$

④ $P(A | \overline{B}) = \frac{P(A \cap \overline{B})}{P(\overline{B})} = \frac{0.05}{1 - P(B)} = \frac{0.05}{1 - 0.7} = \frac{0.05}{0.3}$

	A	$\overline{A}$	
B	$P(A \cap B)$	$P(\overline{A} \cap B)$	$P(B)$
$\overline{B}$	$P(\overline{B} \cap A)$	$P(\overline{A} \cap \overline{B})$	$P(\overline{B})$
	$P(A)$	$P(\overline{A})$	1

مجموع 1

مجموع 1

	A	$\overline{A}$	
B	0.15	0.55	0.7
$\overline{B}$	0.05	0.25	0.3
	0.2	0.8	1

① $0.55 + 0.15 = 0.7$

② $0.25 + 0.05 = 0.3$

③ $0.05 + 0.15 = 0.2$

④ $0.25 + 0.55 = 0.8$

⑤ $0.3 + 0.7 = 1$

⑥ $0.8 + 0.2 = 1$

Ex 3

Given $P(A) = 0.2$
$P(B) = 0.7$

A and B are disjoint , find $P(A \cap B)$

	B	$\overline{B}$	
A	0	0.2	0.2
$\overline{A}$	0.7	0.1	0.8
	0.7	0.3	

$P(A \cap B) = 0$

Ex 4

$\perp\!\!\!\perp$ independent

Given $P(A) = 0.2$
$P(B) = 0.7$

A and B are indepent

→ $P(A \cap B) = P(A) \cdot P(B)$

	A	$\overline{A}$	
B	0.14	0.56	0.7
$\overline{B}$	0.06	0.24	0.3
	0.2	0.8	

→ A and B are independent if
$P(A \cap B) = P(A) \cdot P(B)$

→ If A and B are independent $A \perp\!\!\!\perp B$, $\overline{A} \perp\!\!\!\perp B$
$\overline{A} \perp\!\!\!\perp \overline{B}$, $A \perp\!\!\!\perp \overline{B}$

Ex 5 مثال ال تايتانك

	M Men	W Woman	B Boys	G Girls	
Survival S	332	318	29	27	706
Died D	1360	104	35	18	1517
	1692	422	64	45	2223

→

Find, ① $P(\text{Men}) = \frac{1692}{2223}$

② $P(S) = \frac{706}{2223}$

③ $P(B \cap D) = \frac{35}{2223}$

④ $P(M|S) = \frac{P(M \cap S)}{P(S)} = \frac{\frac{332}{2223}}{\frac{706}{2223}} = \frac{332}{706}$

⑤ $P(\bar{B} \cap D) = \frac{360 + 104 + 18}{2223}$

OR $P(D \cap \bar{B}) = P(D) - P(D \cap B)$

⑥ $P(S|\bar{M}) = \frac{P(S \cap \bar{M})}{P(\bar{M})} = \frac{\frac{706-332}{2223}}{\frac{2223-1692}{2223}}$

$\frac{706 - 332}{2223 - 1692}$

⑦ $P(W|D) = \frac{P(W \cap D)}{P(D)} = \frac{\frac{104}{2223}}{\frac{1517}{2223}} = \frac{104}{1517}$

⑧ $P(B|S) = \frac{\frac{29}{2223}}{\frac{706}{2223}} = \frac{29}{706}$

⑨ $P(W) = \frac{422}{2223}$

⑩ $P(M \cap S) = \frac{332}{2223}$

Ex. 6 هام

لدينا خمسة مقاعد وخمسة أطفال ، نريد أن نجلس ٣ أطفال منهم بجانب بعضهم (لأنهم إخوان) في الكل

Sample Space : خمس أطفال جالسين على ٥ مقاعد

Event : يجلس ٣ اخوان بجانب بعضهم

$n = 5 \quad , \quad r = 5$

${}_5P_5 = \frac{5!}{(5-5)!} = 5! \quad , \quad \# S = 5!$

E : the 3 brothers sit next to each other.

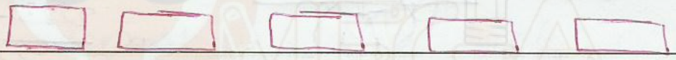
$\# E = (3!)(2!) + (3!)(2!) + (3!)(2!)$

$3(3!\,2!)$

$n = 3$

$r = 3$

${}_nP_r = {}_3P_3 = 3!$

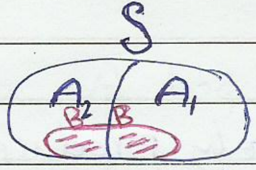


$\overset{*}{3} \; . \; \overset{*}{2} \; . \; \overset{*}{1} \; . \; 2 \; . \; 1$

$2 \; . \; \overset{*}{3} \; . \; \overset{*}{2} \; . \; \overset{*}{1} \; . \; 1$

$2 \; . \; 1 \; . \; \overset{*}{3} \; . \; \overset{*}{2} \; . \; \overset{*}{1}$

$\therefore P(E) = \frac{3(3!\,2!)}{5!}$



Bayes

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

$P(B) = P(B \cap A_1) + P(B \cap A_2) - 0$

$= P(B | A_1) P(A_1) + P(B | A_2) P(A_2)$

OR $= P(A_1 | B) P(B) + P(A_2 | B) P(B)$ → لانستعملها الا اذا بتغير

Ex 1

Drawer 1 Contains 5 pennies of 3 Dimes, drawer 2 Contains 3 pennies of 7 Dimes, A drawer is selected at Random and a coin is selected at $\overset{\text{equaly likly}}{\text{random}}$ from that drawer.

Find ① the Prob. of selecting adime?

② Suppose adime is selected. what is the prob. that it came from drawer 2?

Suppose / Knowing that / Given that / Conditional on

I: 3D, 5P

II: 7D, 3P

$D \cap I = (D, I) = D$ and I

$P(I) = \frac{1}{2}$ $\qquad$ $P(II) = \frac{1}{2}$

$P(D|I) = \frac{3}{8}$ $\qquad$ $P(D|II) = \frac{7}{10}$

1] $P(D) = P(D \cap I) + P(D \cap II)$

$P(D|I) P(I) + P(D|II) P(II)$

$\frac{1}{2}\left(\frac{3}{8}\right) + \frac{1}{2}\left(\frac{7}{10}\right)$

2] $P(II|D) = \frac{P(II \cap D)}{P(D)} = \frac{P(D|II) P(II)}{P(D|I) P(I) + P(D|II) P(II)} = \frac{\frac{7}{10}}{\frac{3}{8} + \frac{7}{10}}$

Ex 2 (not random)

In abolt factory machines (1, 2, 3) respectively بالترتيب produce 20%, 30% and 50% of the total out pub الناتج الكلي. the percentage of defectives are 5%, 3% and 2% are respectively A bolt is selected at random.

1) What is the Prob that it is defective.

2) Given that it is defective, what is the prob. that it came from machine I.

I	II	III
$P(I)=0.2$	$P(II)=0.3$	$P(III)=0.5$
$P(D\|I)=0.05$	$P(D\|II)=0.03$	$P(D\|III)=0.02$

D: Defected Jbs

① $D=(D,I)\cup(D,II)\cup(D,III)$

$P(D) = P(D,I)+P(D,II)+P(D,III)$

$P(D|I)P(I)+P(D|II)P(II)+P(D|III)P(III)$

$\frac{5}{100}\left(\frac{2}{10}\right)+\frac{3}{100}\left(\frac{3}{10}\right)+\frac{2}{100}\left(\frac{5}{10}\right)=\frac{29}{1000}$

② $P(I|D)=\frac{P(I\cap D)}{P(D)}$

$\frac{P(D|I)P(I)}{P(D)}$

$\frac{P(D|I)\,P(I)}{P(D|I)P(I)+P(D|II)P(II)+P(D|III)\,P(III)}$

$\frac{\frac{5}{100}\left(\frac{2}{10}\right)}{\frac{29}{1000}}$

$\frac{10}{29}$

Ex 3 A bag 1 contains 4 red and 2 black balls. A bag 2 contains 3 red and 3 black balls. A fair die will be thrown, if the outcome of the thrown die is 6 then one ball is drawn at random from bag 1, while if the outcome of the thrown die is anything else the ball is drawn from bag 2.

① Find the prob. that the drawn ball is red ?

② Given that the ball is red, Find the prob. that it come from the bag 2 ?

I	II
4R 2B	3R 3B

$P(I) = \frac{1}{6}$ $\qquad$ $P(II) = \frac{5}{6}$

$P(R|I) = \frac{4}{6}$ $\qquad$ $P(R|II) = \frac{3}{6}$

1) $P(R) = \frac{1}{6}\left(\frac{4}{6}\right) + \frac{5}{6}\left(\frac{3}{6}\right) = \frac{19}{36}$

2) $P(II|R) = \dfrac{\frac{5}{6}\left(\frac{3}{6}\right)}{\frac{19}{36}} = \frac{15}{19}$

Discrete Random Variable المتغير العشوائي

Ex 1

Toss a coin 2 times, $S = \{TT, HH, TH, HT\}$, let $X = \#$ of Heads, $X = 0, 1, 2$. Find.

1) $P(X=0) = P(\text{No Heads}) = P(TT) = \frac{1}{4}$

2) $P(X=1) = P(TH \text{ or } HT) = \frac{2}{4} = \frac{1}{2}$

3) $P(X=2) = P(HH) = \frac{1}{4}$

Ex 2

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

⇒ we call this a probability function (Prob. fcn.) اقتران عشوائي

because

1) $0 \leq P(x) \leq 1$ } axioms of probability
2) $\sum P(x) = 1$

Ex 3

X	1	2	10
P(x)	$\frac{1}{3}$	$\frac{2}{3}$	$\frac{1}{6}$

⇒ Is this a prob. fct?
No, $\sum P(x) \neq 1$
$\left(\frac{1}{3} + \frac{2}{3} + \frac{1}{6} = \frac{7}{6} > 1\right)$

Ex 4

X	-1	2	3	4
P(x)	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	0

⇒ Is this a prob. fct?
yes, $\sum P(x) = 1$ &
$0 \leq P(x) \leq 1$

Ex 5

X	-1	1
P(x)	$\frac{2}{4}$	$\frac{-1}{4}$

⇒ Is this a prob. fct?
No, $P(x) < 0$ &
$\sum P(x) \neq 1$

Ex 6

X	0	1	2	3	4
P(x)	$\frac{2}{12}$	$\frac{3}{12}$	C	$\frac{5}{12}$	$\frac{2}{12}$

⇒ Find c, P(x) represents a prob. fct?

$$\frac{2}{12} + \frac{3}{12} + c + \frac{5}{12} + \frac{2}{12} = 1$$

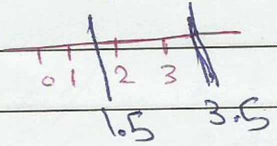
$$\frac{12}{12} + c = 1$$

$$C = 0$$

Ex 7

x	0	1	2	3
$P(x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

1) $P(X=0) = \frac{1}{8}$
2) $P(X \leq 0) = P(X=0) = \frac{1}{8}$
3) $P(0 \leq X < 1) = P(X=0) = \frac{1}{8}$
4) $P(0 < X < 1) = 0$
5) $P(0 \leq X \leq 1) = P(X=0) + P(X=1) = \frac{4}{8} = \frac{1}{2}$
6) $P(1.5 < X < 3.5) = \frac{4}{8} = \frac{1}{2}$



7) $P(X > 0) = 1 - P(X \leq 0)$

$1 - P(X=0)$

$1 - \frac{1}{8}$

$\frac{7}{8}$

Ex 8

$x \backslash y$	1	2	3
-2	0.2	0.1	0.1
-1	0.1	0.1	0.1
0	0.1	0.1	0.1

Find $P(x+y=1)$?

$= P(x=-2, y=3) + P(x=-1, y=2) +$
$P(X=0, y=1)$
$= 0.1 + 0.1 + 0.1$
$= \boxed{0.3}$

$* \; Ex = \sum x P(x)$
$Ex^2 = \sum X^2 P(x)$
$Ex^3 = \sum X^3 P(x)$

Ex * $E(x) = 1$, $Ex^2 = \frac{3}{2}$, Find $E(3X^2 + 2x + 1)$?

1 $E(3X^2 + 2X + 1) = \sum (3X^2 + 2x + 1) P(x)$
$= (3(0)^2 + 2(0) + 1) P(0) + \dots$

OR

2 $E(3X^2 + 2X + 1) = E3X^2 + E2X + E1$
$3EX^2 + 2EX + 1$
$3\left(\frac{3}{2}\right) + 2(1) + 1$
$\frac{15}{2}$

Ex 9 Given that

x	0	1	2
P(x)	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

① Find the mean?

μ, average, expected value Ex / E(x)

$Ex = \sum x\,P(x)$

$= 0p(0) + 1p(1) + 2p(2)$

$= 0 + \frac{1}{2} + \frac{1}{2}$

$= 1$

② Find $Ex^2 = \sum x^2 P(x)$

$= 0^2 p(0) + 1^2 P(1) + 2^2 P(2)$

$= 0 + \frac{1}{2} + 1$

$= \frac{3}{2}$

③ Find $E(2x+1) = 2Ex + 1$

$= 2(1) + 1$

$= 3$

* Variance of x

$Var(x) = \sigma^2 = E(x - Ex)^2 \geq 0$

$Ex^2 - (Ex)^2$

$Ex^2 \geq (Ex)^2$ * لأن التباين دائماً موجب

- $Var(ax+b) = a^2\,Var(x)$
- $Var(2x+1) = (2)^2\,Var(x) = 4(\frac{1}{2}) = 2$ $\rightsquigarrow$ if $var(x) = \frac{1}{2}$
- $Var(3-5x) = (-5)^2\,var(x)$

Ex 10

x	0	1	2
P(x)	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

⇒ Find the var(x)? then Find S.d? $\sigma^2 = Ex^2 - (Ex)^2$

$= \frac{3}{2} - (1)^2 = \frac{1}{2}$

$\sigma = \sqrt{\frac{1}{2}}$

Ex 11

x	0	1	2	3
$P(x)$	c	c	d	$2d$

$\Rightarrow$ Find c and d to make the fct a prob. fct. given $Ex=2$

- $Ex = \sum x\,P(x) = 2$

$Ex = 0(c) + 1(c) + 2(d) + 3(2d) = 2$

$2 = c + 8d \ldots ①$

- $\sum P(x) = 1$

$c + c + d + 2d = 1$

$2c + 3d = 1 \ldots ②$

نحل المعادلتين

* $-2(c + 8d = 2) \rightarrow -2\cancel{c} - 16d = -4$

$2c + 3d = 1$ $\quad$ $2c + 3d = 1$

$-13d = -3$

$\boxed{d = \frac{3}{13}}$

* نعوض d في ①

$c + 8d = 2$

$c + 8\left(\frac{3}{13}\right) = 2$

$c = 2 - \frac{24}{13}$

$c = \frac{26-24}{13} = \frac{2}{13}$, $\boxed{c = \frac{2}{13}}$

Discrete Joint prob. Fct.

We say that $p(x,y)$ is $\overset{\text{Joint}}{\uparrow}$ prob. fct. if

$0 \leq P(x,y) \leq 1$

$\sum\sum P(x,y) = 1$

x \ y	1	3	
0	$P(x=0, y=1)$	$P(x=0, y=3)$	$P(x=0)$
	Joint Prob. Fct.		
1	$p(x=1, y=1)$	$P(x=1, y=3)$	$P(x=1)$
	$P(y=1)$	$P(y=3)$	1

Ex 1

x \ y	1	3	
~~0~~	~~0.25~~	~~0.3~~	~~0.55~~
1	0.1	0.35	0.45
	0.35	0.65	1

فعلياً قيمة أجوبته صفر

1) Find $P(x=0, y=1) = 0.25$

2) $P(x>0, y \geq 3) = 0.35$

3) Compute $Exy = \sum\sum xy\, P(x,y)$

$= \cancel{0(1)(0.25)} + \cancel{0(5)(0.3)} + 1(1)(0.1 + 1(3)(0.35)$

$= 1.15$

4) Compute $Ex^2y^3 = \sum\sum x^2y^3 P(x,y)$

$= \cancel{0^2(1)^3(0.25)} + \cancel{0^2(3)^3(0.3)} + 1^2(1)^3(0.1) + (1)^2(3)^3(0.35)$

Ex 2

x \ y	1	2	3	
1	0.1	0.05	0.02	0.17
2	0.1	0.35	0.05	0.5
3	0.03	0.1	0.2	0.33
	0.23	0.5	0.27	1

* إذا طلب السؤال إيجاد
الـ Ex / Var نجد
أولاً الـ Marginals ثم
أجد Var/Ex

Find ① $P(X=1, y=2) = 0.05$

② $P(X \geq 2, y \geq 3) = 0.05 + 0.2 = 0.25$

③ Marginals of x & y

X	1	2	3
P(X)	0.17	0.5	0.33

y	1	2	3
P(y)	0.23	0.5	0.27

④ $P(X>1 \mid y=1) = \dfrac{P(X>1 \cap y=1)}{P(y=1)}$

$\dfrac{0.1 + 0.03}{0.23}$

⑤ $P(X<2 \mid y>1) = \dfrac{0.05 + 0.02}{0.5 + 0.27} = \dfrac{0.07}{0.77}$

⑥ $P(X>1 \mid y \geq 3) = \dfrac{P(X>1, y \geq 3)}{P(X \geq 3)} = \dfrac{0.2 + 0.05}{0.27}$

$\dfrac{0.25}{0.27} = \dfrac{25}{27}$

⑦ $P(X \geq y \mid y>2) = \dfrac{P(X \geq y, y>2)}{P(y>2)} = \dfrac{0.2}{0.27}$

⑧ $Exy = \sum\sum xy\, P(x,y)$

$= 1(1)(0.1) + 1(2)(0.05) + 1(3)(0.02) + \dots = 4.5$

⑨ $Ex = \sum x P(x) = 2.16$

⑩ $Ey = \sum y P(y) = 2.04$

على الدرس الجديد ⑪ $Cov(X,y) = Exy - ExEy$

$4.5 - (2.16)(2.04)$

Ex 3

x \ y	-1	0	1	
-1	0.1	0.1	0.1	0.3
0	0.05	0.05	0.1	0.2
1	0.2	0.2	0.1	0.5
	0.35	0.35	0.3	

Find the Mean?

$Exy = \sum\sum xy\,p(x,y)$

$= (-1)(-1)(0.1) + (-1)(1)(0.1) + (1)(-1)(0.2) + (1)(1)(0.1)$

$= -0.1$

Covariance b/w X and Y

عند وجود X وY بيسمى الـ Var بـ CoV

$E(X-EX)(Y-EY)$

$Exy - ExEy$ لـيـس من المضمون أن تكون موجبة

$Cov(X,Y) = Exy - \underset{①}{Ex}\,\underset{②}{Ey}$

من ① و ② نعلم بأننا بحاجة للـ Marginals

Ex use Ex 3 to Find Cov(X, Y)?

- $Exy = -0.1$

x	-1	0	1
P(x)	0.3	0.2	0.5

y	-1	0	1
P(y)	0.35	0.35	0.3

- $Ex = -1(0.3) + 1(0.5)$
  $= 0.2$
- $Ey = -1(0.35) + 1(0.3)$
  $= -0.05$
- $Cov(X,Y) = Exy - ExEy$

$-0.1 - (0.2)(\bar{0}.05)$

$-0.1 + 0.01$

-0.09

Correlation b/w X and y

$$\text{Corr}(X,y) = \frac{\text{Cov}(X,y)}{\sqrt{\text{Var}(x)\,\text{Var}(y)}}$$

$-1 \le e \le 1$ $\quad$ e : rho

هنا الإشارة لا تعنى سوى معنى الإتجاه

Ex Compute Corr(X, y).

Given

x \ y	1	2	
-1	0.1	0.5	0.6
1	0.3	0.1	0.4
	0.4	0.6	1

$\text{Cov}(x,y) = Exy - ExEy$

$\text{Corr}(x,y) = \frac{\text{Cov}(x,y)}{\sqrt{\sigma_x^2\,\sigma_y^2}}$

x	-1	1
P(x)	0.6	0.4

y	1	2
P(y)	0.4	0.6

$Ex = -0.6 + 0.4 = -0.2$ $\quad$ $Ey = 0.4 + 1.2 = 1.6$

$Ex^2 = 1$ $\quad$ $Ey^2 = 2.8$

$\text{Var}\,x = 0.96$ $\quad$ $\sigma_y^2 = 0.24$

$Exy = -0.1 + (-2)(0.5) + 0.3 + 2(0.1) = -0.6$

$\text{Cov}(x,y) = Exy - ExEy$

$-0.6 - (-0.2)(1.6)$

-0.28

$$\text{corr}(x,y) = \frac{\text{Cov}(x,y)}{\sqrt{\sigma_x^2\,\sigma_y^2}} = \frac{-0.28}{\sqrt{(0.96)(0.24)}} = -0.58$$

$-1 \le e \le 1$

ProPerties

① $Cov(x, a) = 0$

$$Exa - Ex\,Ea$$

$$aEx - aEx = 0$$

② $Cov(x, x) = Var(x) = \sigma_x^2$

$$= EXX - EXEX$$

$$= Ex^2 - (EX)^2$$

$$= \text{Variance}(X)$$

③ $Cov(ax+b, cy+d) = Cov(ax, cy)$

$$ac\,Cov(X, y)$$

④ $Cov(aX, X) = a\,Var(X)$

⑤ $Corr(aX, X) = \begin{cases} 1, & a>0 \\ -1, & a<0 \end{cases}$

$$Corr(aX, X) = \frac{Cov(aX, X)}{\sqrt{Var(aX)\,Var(X)}}$$

$$\frac{a\,\sigma_x^2}{\sqrt{a^2 Var(X)\,Var(X)}}$$

$$\frac{a\,\sigma_x^2}{|a|\,\sigma_x^2}$$

$$\frac{a}{|a|} = \begin{cases} 1, & a>0 \\ -1, & a<0 \end{cases}$$

$Corr(X, X) = 1$

$Corr(y, y) = 1$

$Corr(X, y) = -1 \leq \rho \leq 1$

⑥ $Corr(ax+b, cy+d) = \begin{cases} Corr(x, y), & ac>0 \\ -Corr(x, y), & ac<0 \end{cases}$

Ex1

x\y	1	2	
1	0.42	0.28	0.7
2	0.18	0.12	0.3
	0.6	0.4	

1) Are x and y ind.?

2) Find corr(x, y) ?

$\boxed{1}$

- $P(X=1)=0.7$, $P(Y=1)=0.6$
  $P(X=1)\cdot P(Y=1)=0.42$
  and $P(X=1, Y=1)=0.42$
- $P(X=1)=0.7$, $P(Y=2)=0.4$
  $P(X=1)\cdot P(Y=2)=0.28$
  and $P(X=1, Y=2)=0.28$
- $P(X=2)=0.3$, $P(Y=1)=0.6$
  $P(X=2)\cdot P(Y=1)=0.18$
  and $P(X=2, Y=1)=0.18$
- $P(X=2)=0.3$, $P(Y=2)=0.4$
  $P(X=2)\cdot P(Y=2)=0.12$
  and $P(X=2, Y=2)=0.12$

Yes, they are independent.

$\boxed{2}$ نجد الـ Corr بالطريقة المعتادة

The result is Zero.

Ex 2

$Ex=1$, $Exy=1$, $\sigma_x^2=4$, $Ey^2=10$, $Ey=3$

① $Cov(X, Y)$

② $Cov(2X+1, 1-Y)$

③ $Corr(2X+1, 1-Y)$

④ $Var(-2X+3Y)$

$\boxed{1}$ $Cov(X,Y) = Exy - ExEy$
$= 1 - 1(3) = -2$

$\boxed{2}$ $Cov(2X+1, 1-Y) = Cov(2X, -Y)$

* نهتم بالعوامل ولا نهتم بالثوابت

$= -2\ Cov(X,Y)$
$= -2(-2)$
$= 4$

$\boxed{3}$ $Corr(2X+1, 1-Y) = \dfrac{Cov(2X+1, 1-Y)}{\sqrt{\sigma^2_{(2x)}\ \sigma^2_{(-y)}}}$

$\dfrac{4}{\sqrt{4\ \sigma^2_{(x)}\ \sigma^2_{(y)}}}$

$\dfrac{4}{\sqrt{16}} = 1$

OR $Corr(2X+1, 1-y) = -Corr(X, y)$

يهمها اشارة x , y فقط ولانهتم
بالعوامل ولا بالثوابت

$$-\left[\frac{Cov(X,y)}{\sqrt{\sigma_x^2 \sigma_y^2}}\right]$$

$$-\left[\frac{(-2)}{\sqrt{4(1)}}\right]$$

$$1$$

$$\sigma_y^2 = Ey^2 - (Ey)^2 = 10 - 9 = 1$$

* Corr ⇒ نهتم باشارة معامل X و y فقط
حيث ؛

+ Corr إذا كانوا متشابهين بالإشارة

− Corr و = = مختلفين بالإشارة

④ على الدرس الجديد

$$Var(-2X+3y) = Var(-2X) + Var(3y) + 2\,Cov(-2X, 3y)$$

$$(-2)^2 Var(x) + 9\,Var(y) + 2(-2)(3)\,Cov(X, y)$$

$$4(4) + 9(1) - 12(-2)$$

$$49$$

Ex 3

X \ y	1	2	3	
-1	0.1	0.1	0.1	0.3
0	0.2	0.1	0.1	0.4
1	0.1	0.1	0.1	0.3
	0.4	0.3	0.3	

- Find $cov(X, y)$?
- Are X and y indp.?

$$Cov(X,y) = Exy - ExEy$$

$$Exy = \sum\sum xy\,P(x,y) = 0$$

x	-1	0	1
PX	0.3	0.4	0.3

y	1	2	3
P(y)	0.4	0.3	0.3

$Ex = 0$ $Ey = 1.9$

So $Cov(X,y) = 0 - 0(1.9) = 0$

* لا يكفي أن يكون $cov(x,y)=0$ للحكم على أن $X \perp\!\!\!\perp y$

$$P(X \cdot y) \overset{?}{=} P(X)\,P(y)$$

$$P(x=1, y=1) \overset{?}{=} P(X=1)\,P(y=1)$$

$$0.1 \neq (0.3)(0.4)$$

$$0.1 \neq 0.12$$

So $X \not\perp\!\!\!\perp y$
Not indp.

For any Random Variable هام جداً

- $Var(X+y) = Var(x) + var(y) + 2Cov(X,y)$
- $Var(X-y) = Var(X) + Var(-y) + 2\,Cov(X,-y)$

$Var(X+(-y)) = Var(x) + Var(y) - 2Cov(X,y)$

$\Rightarrow$ $X \perp\!\!\!\perp y$ then $\rightarrow Cov(x,y) = Zero$

$Var(X+y) = Var(X) + Var(y)$

$Var(X-y) = Var(x) + Var(y)$

Ex 1 $Var(2Z - W)$

$= Var(2Z) + Var(-W) + 2Cov(2Z,-W)$

$= 4\,Var(Z) + Var(W) - 4\,Cov(Z,W)$

Ex 2

$Ex = 1,\ Ey = 3,\ Exy = 1,\ Ey^2 = 10,\ Var(X) = 4$

- $Var(1 - 2X + 3y + 4)$?

$= Var(-2X) + Var(3y) + 2\,Cov(-2X, 3y)$

$= 4\,Var(X) + 9\,Var(y) - 12\,Cov(X,y)$

$= 4(4) + 9(1) - 12(-2) = 49$

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

$10 - (3)^2 = 1$

- $Cov(X,y) = Exy - ExEy$

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

EX 3 سؤال امتحان

$Ex = 12,\ Ey = 18,\ \sigma_x^2 = 9,\ \sigma_y^2 = 16,\ Corr(X,y) = 0.5$, Find $Var(2X + 3y)$?

$= Var(2X) + Var(3y) + 2\,Cov(2X, 3y)$

$= 4\,Var(x) + 9\,Var(y) + \underbrace{12\,Cov(X,y)}$ خدها من Corr.

$Corr(X,y) = \dfrac{Cov(X,y)}{\sqrt{\sigma_x^2\,\sigma_y^2}}$

$0.5 = \dfrac{Cov(X,y)}{\sqrt{9(16)}}$

$0.5(3)(4) = Cov(X,y)$

$Cov(X,y) = 6$

$= 4\,var(x) + 9\,var(y) + 12(6$

$= 4(9) + 9(16) + 12(6)$

if X and Y are indp then

1) $P(X, Y) = P(X) P(Y)$
2) $EXY = EX EY$
3) $Cov(X, Y) = 0$
4) $Corr(X, Y) = 0$

- Binomial distribution
- Poisson =
- Geometric =
- Hyper Geometric =

1) Binomial Distribution

خيارين فقط

$X \sim \text{Binomial}(n, P)$ iff

$F(x) = \binom{n}{x} P^{x} q^{(n-x)}$ $\{q = 1-P\}$

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

$\sim$: distributed as or distribtanted as

n : # of trials عدد المحاولات $\quad Bi(\underbrace{n, P})$

P : Prob. of success احتمال النجاح $\quad$ Parameter معلمات

q : $(1-P)$ Prob. of Failer احتمال الفشل

X : $0, 1, \dots, n$; # of correct answer.

Ex1 $X \sim Bi(5, \frac{1}{3})$

$F(x) = \binom{5}{x} (\frac{1}{3})^{x} (1-\frac{1}{3})^{5-x}$, Find $P(x=2)$?

$P(X=2) = F(2) = \binom{5}{2} (\frac{1}{3})^{2} (\frac{2}{3})^{3}$

$P(X=0) = F(0) = \binom{5}{0} (\frac{1}{3})^{0} (\frac{2}{3})^{5}$

Ex 2 Toss a Coin 3 times

- Find the Prob of $\underline{\text{getting } \overset{P(x=3)}{3} \text{ heads}}$?

X = # of head

$X \sim Bi(3, \frac{1}{2})$

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

$= \binom{3}{x} (\frac{1}{2})^{x} (1-\frac{1}{2})^{3-x}$

$P(X=3) = F(3) = \binom{3}{3} (\frac{1}{2})^{3} (\frac{1}{2})^{0} = \frac{1}{8}$

اذا ظهر H ← نجاح (Success) احتماله $\frac{1}{2}$

اذا ظهر T ← فشل

Rules

① * $F(c) = P(X \leq c)$

$F(5) = P(X \leq 5)$

$F(1) = P(X \leq 1)$

* $F^{-}(c) = P(X < C)$ أقل بدرجة (من اليسار) ; (-)

② $P(a < X \leq b) = F(b) - F(a)$

$P(a < x < b) = \bar{F}(b) - F(a)$

$P(a \leq x \leq b) = F(b) - \bar{F}(a)$

$P(a \leq x < b) = \bar{F}(b) - \bar{F}(a)$

③ $P(X > a) = 1 - P(X \leq a)$

$1 - F(a)$

$P(X \geq a) = 1 - P(X < a)$

$1 - \bar{F}(a)$

Ex 3

$X \sim Bi(10, 0.2)$ use Binomial tables to find:

① $P(2 < X \leq 5) = F(5) - F(2)$

$0.994 - 0.678$

② $P(2 \leq X \leq 5) = F(5) - \bar{F}(2)$

$F(5) - F(1)$

$0.994 - 0.376$

③ $P(2 < X < 5) = \bar{F}(5) - F(2)$

$F(4) - F(2)$

$0.967 - 0.678$

④ $P(2 \leq X < 5) = \bar{F}(5) - \bar{F}(2)$

$F(4) - F(1)$

$0.967 - 0.376$

⑤ $P(X \leq 4) = F(4) = 0.967$

⑥ $P(X > 3) = 1 - P(X \leq 3)$

$1 - F(3)$

$1 - (0.879)$

⑦ $P(X<4) = F^{-}(4)$

$F(3)$

0.879

⑧ $P(X=4) = P(3 < X < 5) = F^{-}(5) - F(3)$

$F(4) - F(3)$

$0.967 - 0.879$

<u>Ex 4</u>

if $X \sim Bi(n, p)$ then

$EX = np$

$\sigma_x^2 = npq$ $\quad$ $q = 1-p$

$X \sim Bi(10, 0.2)$

$EX = np$

$10(0.2) = 2$

$\sigma_x^2 = npq = 10(0.2)(0.8) = 1.6$

* إسئلة امتحانات *

1. $X \sim Bi(25, 0.2)$ Find $P(\mu - \sigma < X \leq \mu + \sigma)$

$\mu = np$

$25(0.2) = 5$

$\sigma^2 = \overset{n}{25}(\overset{p}{0.2})(\overset{q}{0.8}) = 4$

$\sigma = 2$

$P(5-2 < X \leq 5+2)$

$P(3 < X \leq 7)$

$F(7) - F(3)$

$0.891 - 0.234$

$\boxed{2}$ $X \sim Bi(\text{with } \mu = 2 , \sigma^2 = 1.6)$

Find $P(X \leq 1) = F(1)$?

- $\mu = np$ $\rightsquigarrow$ $\mu = np$
- $\sigma^2 = npq$ $\quad 2 = n(0.2)$

$\sigma^2 = \mu q$ $\quad n = 10$

$1.6 = 2q$

$0.8 = q$

$P = 0.2$

- $X \sim Bi(10, 0.2)$

$P(X \leq 1) = F(1) = 0.3226$

② Poisson distribution

$X \sim Po(\mu)$ iff

$F(X) = \frac{e^{-\mu} \mu^{X}}{X!}$ $\quad X = 0, 1, 2, \ldots$

$\left.\begin{array}{l} EX = \mu \\ Var(X) = \mu \end{array}\right\}$ هنا الـ μ دائماً موجب

$Bi(n, p) \xrightarrow{n \text{ large} ,\ p < 0.01} Po(\mu)$

Per time / every / average( ) / Ex

Per distance

typo errors Per page

Ex 1

IF X = # of deaths Per month and the average # of deaths per month is 2. Find the distribution of X.

$X \sim Po(2)$

Ex 2

If $y = \#$ of accidents per day and the average $\#$ of accidents per week is 7. Find the distribution of Y.

$Y \sim Po(1)$

Ex 3

If $Z = \#$ of typo error per page and the average $\#$ of typo error in a book consists of 100 page is 200. Find the distribution of Z.

$Z \sim Po(2)$

at most / at least $\rightsquigarrow$ متضمنة المساواة

Ex 4 Suppose $X \sim Po(2)$, find

① $P(X \leq 3) = F(3) = 0.857$

② $P(X = 3) = P(2 < X < 4) = \bar{F}(4) - F(2) = F(3) - F(2)$

$= 0.857 - 0.677$

③ $P(1 \leq X \leq 4) = F(4) - \bar{F}(1) = F(4) - F(0)$

$= 0.947 - 0.135$

④ $P(1 \leq X < 4) = \bar{F}(4) - \bar{F}(1) = F(3) - F(0)$

$= 0.857 - 0.135$

⑤ $P(X \text{ is at most } 3) = P(X \leq 3) = F(3) = 0.857$

⑥ $P(X \text{ is at least } 3) = P(X \geq 3) = 1 - P(X < 3)$

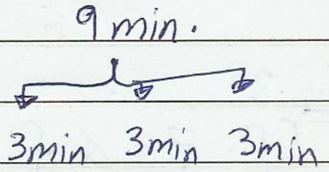
$1 - \bar{F}(3)$

$1 - F(2)$

$1 - 0.677$

Ex 5

Suppose a fast food restaurant can expect two Customers every 3 minutes, on average. What is the prob. that at most four Customers will enter the restaurant in a 9 minute Period.



$X \sim Po(6)$

$P(X \leq 4) = F(4) = 0.285$

$2 + 2 + 2 = 6$

لو كانت قيمة μ غير موجودة في الجدول لاستخدمنا القانون

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

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

Ex 6

$X \sim Bi(3, 0.5)$ } $\perp\!\!\!\perp$ (they are indp.)
$y \sim Po(0.1)$

$X = 0, 1, 2, 3$
$y = 0, 1, 2, 3, 4, \cdots$

① $P(X + y = 2)$

$P(x+y=2) = P(x=0, y=2) + P(x=1, y=1) + P(x=2, y=0)$

$= P(x=0)P(y=2) + P(x=1)P(y=1) + P(x=2)P(y=0)$

Po. & Bin. كلهم من جداول $= (0.125)(1) + (0.5)(0.995) + (0.875)(0.905)$

② $Exy = Ex\,Ey$

$Ex = \mu = np = 3(0.5) = 1.5$

$Ey = \mu = 0.1$

$ExEy = (1.5)(0.1)$

(0.15)

③ $Var(3X - 5y + 4) = Var(3X - 5y)$

$= 9\,Var(X) + 25\,Var(y) + 2(3)(-5)\,Cov(x, y)$

$= 9\,Var(X) + 25\,Var(y) - 30\,\underbrace{\cancel{Cov}(X, y)}_{zero}$

$Var(x) = \sigma_x^2 = npq = (1.5)(0.5) = (0.75)$

$Var(y) = \mu = 0.1$

$= 9(0.75) + 25(0.1)$

$= 6.75 + 2.5$

$= 9.25$

Ex 7

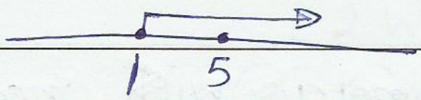
$X \sim Bi(5, 0.1)$, find

$P(X=5 | X \geq 1)$?

$$\frac{P(x=5, X \geq 1)}{P(X \geq 1)}$$

$$\frac{P(X=5)}{P(X \geq 1)}$$

$$\frac{P(4<X<6)}{1-P(X<1)}$$



$$\frac{F^{-}(6)-F(4)}{1-F^{-}(1)}$$

$$\frac{F(5)-F(4)}{1-F(0)}$$

$$\frac{1-1}{1-0.951} = \text{Zero}$$

3) Geometric distribution, $G(P)$.

- إذا كان عدد المحاولات ثابتة ، والنجاحات متغير (Bino.)
- أما إذا كان عدد المحاولات غير ثابتة ، والنجاحات ثابتة (Geo)

- IF X = # of trials required to obtain the 1st success.

$X \sim Geo(P)$

$F(x) = q^{(x-1)} p \qquad X = 1, 2, 3, \dots$

$EX = \frac{1}{p}$, $Var(x) = \frac{q}{p^2}$

Keyword

until, Keep on, with replace.

Ex 1

استمر يخالف قواعد السير

A car driver will Keep on Violating traffic rules until he gets caught by police. If the prob. that the driver will be caught is 0.3, Find the prob. that he will be caught on his 2nd violation? $P = 0.3$, $q = 0.7$

$P(X=2) = F(2) = q^{(2-1)}p = qp = (0.7)(0.3) = 0.21$

Ex 2

IF X~Geometric with mean = 4 , Find P(x = 3)?

$\mu = 4$

$Ex = \frac{1}{p} = 4$, $P = \frac{1}{4}$

$Geo(\frac{1}{4})$

$P(X=3) = q^2 P = (\frac{3}{4})^2(\frac{1}{4})$

$\sigma^2 = \frac{q}{p^2} = \frac{\frac{3}{4}}{\frac{1}{(4)}} = \frac{3}{4}(16) = 12$

$Ex^2 = \sigma^2 + (Ex)^2 \Longleftarrow \sigma^2 = Ex^2 - (Ex)^2$

$12 + (4)^2$

28

* لتمييز السؤال ...

لـ Bino. ← يجب أن يذكر عدد n

لـ Pois. ← من كلمة Per

لـ Geo ← من كلمة

* قد تكون Geo. أو Bin. ونميز بينهما من حيث عدد المحاولات ثابت أولاً + وجود until
تعني Geo

EX 3

A box contains 3 black and 2 white balls we $\boxed{\text{keep on}}$ drawing balls at a time from the box $\boxed{\text{with replacement}}$, $\boxed{\text{until}}$ the 1st black ball shows up. what is the prob. that the first black ball will show up in the 3rd trial.

with replacement $\Rightarrow$ Geo.

P(success)

$P(B) = \frac{3}{5}$

$P(X=3) = qqp$

$\frac{2}{5} \cdot \frac{2}{5} \cdot \frac{3}{5} = \frac{12}{125}$

| 3 B |
| 2 W |

qqp

$P(X=3) = q^{(3-1)} p$

$q^2 p$

qqp

4) Hypergeometric distribution

$\text{Hyp}(n, N, M)$ → total # of balls

$$F(x) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$$

# of R

sample size

$$\text{Max}(0, n-N+M) \leq X \leq \text{Min}(n, M)$$

Keyword without replacement عملية سحب بدون ارجاع

Ex 1 Suppose that 5 balls that draw without replacement from a box contains 4 red & 11 black balls?

① Find prob. of getting exactly 2 red balls.

② = = = = at least 4 red balls.

$$\boxed{1}\ P(X=2) = \frac{\overset{R}{\binom{4}{2}}\ \overset{NR}{\binom{11}{3}}}{\binom{15}{5}}$$

تكمل هنا

$\boxed{2}$ $P(X \geq 4) = P(X=4) + \cancel{P(X=5)}$

$$\frac{\overset{R}{\binom{4}{4}}\overset{NR}{\binom{11}{1}}}{\binom{15}{5}} + \cancel{\frac{\binom{4}{5}\binom{11}{0}}{\binom{15}{5}}}$$

$\binom{n}{r}, n \geq r$

$\text{Max}(0, n-N+M) \leq X \leq \text{Min}(n, M)$

$\text{Max}(0, 5-15+4) \leq X \leq \text{Min}(5, 4)$

$\text{Max}(0, -6) \leq X \leq 4$

$0 \leq X \leq 4$

Ex 2

6R
9B

5 balls نريد أن نسحب

$0 \leq X \leq 5 \Longrightarrow$ من القانون

$P(X \geq 4) = P(X=4) + P(X=5) + \cancel{P(X=6)}$

Ex 3

a box Contains 10 electric bulbs among them, there are 4 defectives if 3 bulbs are drawn without replacement From the box, What is the prob. of getting.

1) exactly one defective.
2) = two = .
3) Not more than 2 defectives. (at most)

4D
6ND

3bulbs

ND → Non defective (٤ تالفة والباقي غير تالف)

10 electric bulbs

① $P(X=1) = \frac{\overset{D}{\binom{4}{1}}\overset{ND}{\binom{6}{2}}}{\binom{10}{3}}$

② $P(X=2) = \frac{\binom{4}{2}\binom{6}{1}}{\binom{10}{3}}$

③ $P(X \leqslant 2) = P(X=0) + P(X=1) + P(X=2)$

$P(X=0) = \frac{\binom{4}{0}\binom{6}{3}}{\binom{10}{3}} +$

$P(X=1) = \frac{\binom{4}{1}\binom{6}{2}}{\binom{10}{3}} +$

$P(X=2) = \frac{\binom{4}{2}\binom{6}{1}}{\binom{10}{3}} .$

$\mathrm{Max}(0, 3-10+4) \leq X \leq \mathrm{Min}(3, 4)$

$0 \leq X \leq 3$

Ex 4

$X \sim \mathrm{Geo}\left(\frac{1}{4}\right)$, Find $P(\overset{\text{سيجما}}{\sigma} < x < M)$?

$EX = \frac{1}{p} = \frac{1}{\frac{1}{4}} = 4$

$\sigma^2 = \frac{q}{p^2} = \frac{3/4}{1/16} = 3(4) = 12$, $\sigma = \sqrt{12}$

$P(\sqrt{12} < X < 4) = 0$

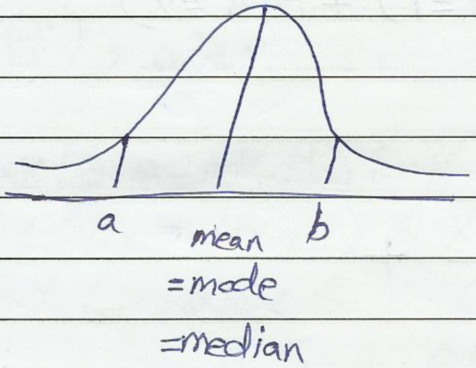
← بسبب وجود أعداد غير حقيقية

Chapter 4

Normal distribution

Continous random variable متغير متصل

discrete r.v متغير منفصل



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

عدم حفظ

$X \sim N(\mu, \sigma^2)$

$P(a < x \leq b) = \text{Area}$

So $P(x = a) = \text{Zero}$

Standared Normal

$\mu = 0 \quad , \sigma^2 = 1$

$Z \sim N(0, 1)$

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

عدم حفظ

Ex 1

① $P(Z<1.79) = F(1.79) = 0.9633$

② $P(Z<-3.39) = F(-3.39) = 0.0003$

③ $P(Z<0) = F(0) = 0.5$

④ $P(Z<-0.73) = F(-0.73) = 0.2327$

⑤ $P(-1<Z<1) = F(1) - F(-1) = 0.8413 - 0.1587$

⑥ $P(2.9<Z<2.99) = F(2.99) - F(2.9) = 0.9986 - 0.9981$

يمنع منعاً باتاً التقريب عند استخدام الجدول $= 0.0005$

⑦ $P(3.74<Z<4.35) = F(4.35) - F(3.74) = 1 - 1 = \text{Zero}$

لـه أكبر من أكبر قيمة في الجدول   أكبر من أكبر قيمة هنا في الجدول

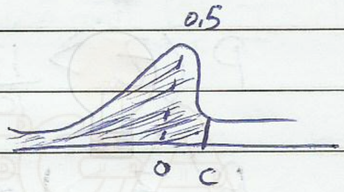
⑧ $P(Z<-5) = F(-5) = 0$ أصغر من أصغر قيمة في الجدول

Ex 2 هام جداً

Find c such that :

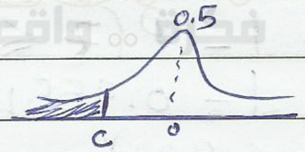
① $P(Z<c) = 0.6772$

$c = 0.46$



② $P(Z<c) = 0.0051$

$c = -2.57$



③ $P(Z<c) = 0.9$

$P(Z<c) = 0.8997$ * نأخذ القيمة الأصغر تقليل من المطلوب

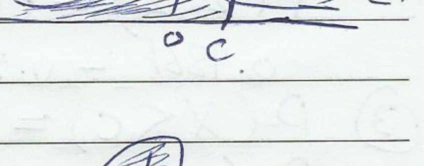
$c = 1.28$ لأن القيمة المطلوبة غير موجودة في الجدول

④ Find the upper 10th percentile of Z

$P(Z<c) = 0.9$

$\boxed{c = 1.28}$ Normal-Table

$c = 1.282$ T-table.



⑤ Find the 95th percentile of Z

$P(Z<c) = 0.95$

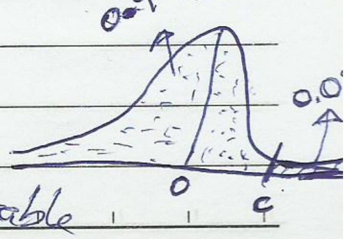
$\boxed{c = 1.64}$ Normal-Table

$c = 1.645$ T-Table

⑥ Find c such that $P(Z>c) = 0.025$

$P(Z<c) = 0.975$

$\boxed{c = 1.96}$ Normal and T table



$X \sim N(\mu, \sigma^2)$

$\mu \neq 0, \quad \sigma^2 \neq 1$

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

$$E\left(\frac{X-\mu}{\sigma}\right) = 0$$

$$\mathrm{Var}\left(\frac{X-\mu}{\sigma}\right) = 1$$

Ex 3 $\quad X \sim N(68, 100)$, Find

$\boxed{\begin{matrix}\mu = 68 \\ \sigma = 10\end{matrix}}$

① $P(X > 85) = P\left(\frac{X-\mu}{\sigma} > \frac{85-68}{10}\right)$

$P\left(Z > \frac{17}{10}\right)$

$P(Z > 1.7)$

$1 - P(Z \leq 1.7)$

$1 - 0.9554$

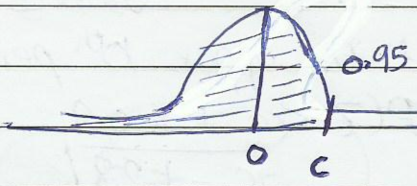
② $P(60 < X < 90)$

$P\left(\frac{60-68}{10} < \frac{X-\mu}{\sigma} < \frac{90-68}{10}\right)$

$P(-0.8 < Z < 2.2)$

$F(2.2) - F(-0.8)$

$0.9861 - 0.2119$



③ $P(X < c) = 0.95$ .. (Find 95th percentile of X)

$P\left(\frac{X-\mu}{\sigma} < \frac{c-68}{10}\right) = 0.95$

$P(Z < c^*) = 0.95$

$c^* = 1.64$

$\frac{c-68}{10} = 1.64$

$c = 68 + 1.64(10) = 84.4$

OR

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

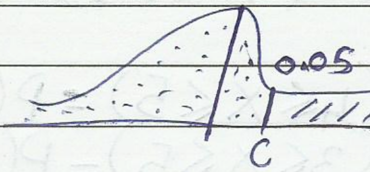
$$c = \mu_x + \sigma_x Z_{1-\alpha} \rightarrow$$ ألفا

$$68 + 10(Z_{0.95})$$

$$68 + 10(1.64)$$

④ Find upper 5th percentile of X

Same as ③



ملحوظة هامة جداً

نذهب إلى جدول T إذا القيمة لم تتواجد في جدول الـ Normal

Normal approximation to the Binomial

$$\text{Binomial}(n, p) \xrightarrow[nq \geq 5]{np \geq 5} N(\mu, \sigma^2)$$

$np \leftarrow \mu \qquad \sigma^2 \rightarrow npq$

$X \sim Bi(100, 0.2)$

Find $P(X \geq 30)$? لا يمكن حلها بالجدول

$P(x=0) = P(-0.5 < X < 0.5)$

$P(X=2) = P(1.5 < X < 2.5)$

$P(3 < x \leq 5) = P(4 \leq X \leq 5) = P(3.5 < X < 5.5)$

$P(3 \leq X \leq 5) = P(3 \leq X \leq 5) = P(2.5 < X < 5.5)$

$P(3 < X < 5) = P(X=4) = P(3.5 < X < 4.5)$

$P(3 \leq X < 5) = P(3 \leq X \leq 4) = P(2.5 < X < 4.5)$

<u>Ex 1</u> $X \sim Bi(100, 0.2)$, Find

① $P(X < 26)$ إذا ذكر approximation مباشرة بنفس الطريقة دون

② $P(X = 26)$ تشييك

③ $P(18 < X < 26)$

④ $P(18 \leq X < 26)$

$np = 100(0.2) = 20 \geq 5$ ✓ الاختبار

$nq = 100(0.8) = 80 \geq 5$ ✓

$\mu = np = 20 \rightsquigarrow \mu = 20$

$\sigma^2 = npq = 16 \rightsquigarrow \sigma = 4$

<u>الخطوات</u>

① Equality

② Continuity Conenta ∓ 0.5

③ standrization $X \rightarrow Z$

④ use table

solve ① $P(X < 26) = P(X \leq 25)$

$P(X < 25.5)$

$P\left(\frac{X-\mu}{\sigma} < \frac{25.5-20}{4}\right)$

$P(Z \leq 1.375)$

$P(Z \leq 1.38)$

0.9162

② $P(X = 26) = P(25.5 < X < 26.5)$

$P\left(\frac{25.5-20}{4} < Z < \frac{26.5-20}{4}\right)$

$P(1.375 < Z < 1.625)$

$P(1.38 < Z < 1.63)$

③ $P(18 < X < 26) = P(19 \leq X \leq 25)$

$P(18.5 < X < 25.5)$

$P\left(\frac{18.5-20}{4} < Z < \frac{25.5-20}{4}\right)$

④ $P(18 \leq X < 26)$

$P(18 \leq X \leq 25)$

$P(17.5 < X < 25.5)$

$P\left(\frac{17.5-20}{4} < Z < \frac{25.5-20}{4}\right)$

ملاحظة

$P(1.713 < Z < 2.257)$

$P(1.71 < Z < 2.26)$

تقريب عادي لمنزلتين لأن جدول الـ Norma L ليس فيه
سوى منزلتين

⑤ $P(X > 26) = P(X \geq 27)$

$= P(27 \leq X)$

$= P(26.5 \leq X)$

$= P\left(\frac{26.5-20}{4} \leq Z\right)$

$= P(Z \geq 1.63) = 1 - P(Z < 1.63)$

← نتيجة لاختلاف صيغة السؤال عن الترتيب المتعارف عليه نعيد ترتيبها ومن ثم نبدأ الحل المعتاد

Ex 2

Suppose that 10% of heavy smokers will suffer from Lung Cancer after the age of 40, In a Sample of 100 heavy smoker, what is the prob. that. (Find Approximate prob.)

① At Least 12 will have a Lung Cancer.
② not more than 12 have a Lung cancer.
③ Exactly 12 have Lung Cancer.

$X \sim Bi(100, 0.1) \longrightarrow X \sim N(10, 9)$

① $P(X \geq 12)$
② $P(X \leq 12)$
③ $P(X = 12)$

Like the previons ex $\mu = 10$
$\sigma = 3$

Central Limit Theorem (C.L.T)

Let X represents apopulation having mean μ and Variance σ^2. If we Collected $\bar{X}$'s for millions of Samples of size n from X, then the distribution of $\bar{X}$'s would be approximatEly Normal with mean μ and s.d $\frac{\sigma}{\sqrt{n}}$

Population $N = 3 \rightarrow \{1, 2, 3\}$

n → Sample size. Select sample of size $\boxed{n=2}$ with replacement.

Sample	$\bar{X}$
(1,1)	1
(1,2)	1.5
(1,3)	2
(2,1)	1.5
(2,2)	2
(2,3)	2.5
(3,1)	2
(3,2)	2.5
(3,3)	3

$\sigma^2 =$

$\mu = 2$

$E\bar{X} = \mu$

$Var(\bar{X}) = \frac{\sigma^2}{n}$

$E(\bar{X}) = \frac{18}{9} = 2$

$E\bar{X} = \mu$ if n Large

$\sigma_{\bar{X}}^2 = \frac{\sigma^2}{n}$

$$\underset{\substack{\downarrow \\ \text{approximately}}}{\bar{X} \sim} N(\mu, \frac{\sigma^2}{n})$$

Ex 1

$X_1, X_2, \dots, X_{16} \sim N(\overset{\mu}{10}, \overset{\sigma^2}{9})$

1) Find $p(\bar{x} < 12)$

2) Find the 95th percentile of $\bar{x}$.

For Larg n=16 , $\bar{x} \sim N(\mu, \frac{\sigma^2}{n})$

① $P(\bar{x} < 12) = P((Z < \frac{12-10}{3/4})$

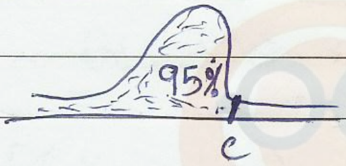
$P(Z < 8/3)$

$P(Z < 2.666\dots)$

$P(Z < 2.67)$

0.9962

②



$P(\bar{x} < c) = 0.95$ $\quad c = \mu_{\bar{x}} + \sigma_{\bar{x}} Z_{0.95}$

$P(Z < c^*) = 0.95$ $\quad 10 + \frac{3}{4}(1.64)$

$c^* = 1.64$

$\frac{c-10}{3/4} = 1.64$

$c = 11.23$

Ex 2

The average age of a car 96 months. assume the S.d = 16 months if a random sample of 36 cars is selected. Find the Prob. that the mean of their age is b/w 90 and 100 months. (S.d : standerd diveation)

$\mu = 96 \quad \mu_{\bar{x}} = 96$

$n = 36 \quad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{16}{6} = \frac{8}{3}$

$\sigma = 16$

$P(90 < \bar{x} < 100)$

$P(\frac{90-96}{8/3} < Z < \frac{100-96}{8/3})$

Note

$$\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$$

$$\frac{\bar{X} - \mu}{\sqrt{\frac{\sigma^2}{n}}} = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \sim N(0, 1)$$

Ex 3 the average number of pounds of meat aperson consumes ayear is 218.4 pounds. Assume that the S.d is 25 pounds and the dist. is approximately Normal.

1) Find the Prob. that aperson selected at random consumes less than 224 Pound / Year.

2) If asample of 40 individuals is selected. Find the Prob. that the mean of the sample will be less than 224 pound / year

Solve 1) $P(X < 224) = P(\frac{x - \mu}{\sigma} < \frac{224 - 218.4}{25})$

2) $P(\bar{X} < 224) = P(\frac{\bar{X} - \mu}{\sigma / \sqrt{n}} < \frac{224 - 218.4}{25 / \sqrt{40}})$

عند ذكر كلمة Sample فكل المعطيات التي تليها تكون لـ Sample
فليست لل Population وما عدا ذلك لل Population

Ex 4

Suppose that the weight of orange boxes are randomly distributed with mean 10 kgs and S.d = 1.5 kg. If anumber of boxes will be Loaded in aCar with threshole 1000 kgs. Find the ~~x~~ of boxes that will be Loaded so that their total weight doesn't exceed the threshole with Prob. 0.95

$P(\sum X \leq 1000) = 0.95$ $\mu = 10$

$P\left(\frac{\sum X}{n} \leq \frac{1000}{n}\right) = 0.95$ $S.d = \sigma = 1.5$

$P\left(\bar{X} \leq \frac{1000}{n}\right) = 0.95$

$c = \mu_{\bar{x}} + \sigma_{\bar{x}} \, z_{(0.95)}$ → بشرط أن تكون $\bar{X} \leq c$

$\frac{1000}{n} = 10 + \frac{1.5}{\sqrt{n}}(1.64)$

Solve for n $\Rightarrow$ $n = 97.57 \simeq 98$

Ex 5

Let X has apoisson distribution such that $P(X=0) = 5\,P(X=1)$, then $P(X=0)$?

$X \sim Po(\mu)$

$P(X=0) = 5\,P(X=1)$ طبقاً للقانون

$\frac{e^{-\mu}\mu^{0}}{0!} = 5\,\frac{e^{-\mu}\mu^{1}}{1!}$

$\cancel{e^{-\mu}} = 5\,\cancel{e^{-\mu}}\,\mu$

$1 = 5\mu$

$\boxed{\mu = \frac{1}{5}}$ نعود ونطبق

$P(X=0) = \frac{e^{-\frac{1}{5}}\left(\frac{1}{5}\right)^{0}}{0!} = \boxed{e^{-\frac{1}{5}}}$

Ex 6

Suppose $X \sim N(10, 16)$ and $y \sim Geo.(P)$ where $P = P(X < 10)$, then $P(y=5)$

$P = P(Z < \frac{10-10}{4})$

$\mu_x = 10$

$\sigma_x = 4$

$P = P(Z < 0)$

$P = F(0) = 0.5 = \frac{1}{2}$ So

$y \sim Geo(\frac{1}{2})$

$P(y=5) = q^{y-1} p$

$(\frac{1}{2})^4 (\frac{1}{2})$

$\frac{1}{2^{(5)}}$

$\boxed{\frac{1}{32}}$

Ex 7

$X \sim Geo.\overset{p}{(0.4)}$. Find $P(X>3)$

$F(X) = q^{x-1} p$

$P(X>3) = 1 - P(X \leq 3)$

$1 - \{ P(x=3) + P(X=2) + P(X=1) \}$

$1 - \{ (0.6)^2 (0.4) + (0.6)^1 (0.4) + (0.6)^0 (0.4) \}$

$1 - 0.784$

0.216

Ex 8

let X be arandom variable with Possible value 0, 1, 2 . If $E(X)=1$ and $E(X^2)=0.75$, then $P(X=2)$?

$0f(0) + (1)F(1) + (2)F(2) = 1$

$F(1) + 2F(2) = 1$ --- ①

$F(1) + 4F(2) = 0.75$ --- ②

بحل المعادلتين ينتج

$F(2) = -0.125$

	Population	Sample
Size	N	n
Mean	μ	$\bar{X}$
Variance	σ^2	S^2
proportion	P	$\hat{P}$ → P hat

Sampling distribution of $\bar{X}$

$$\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$$

$$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1) \quad \sigma \text{ is unknown}$$

$$\frac{\bar{X}-\mu}{S/\sqrt{n}} \sim t_{n-1}$$

$t_{0.1,10} = 1.372$, $t_{0.1,15} = 1.341$

d.f = degrees of freedom

n = Sample size

t_α → المساحة

∞ → Normal distribution

Q1

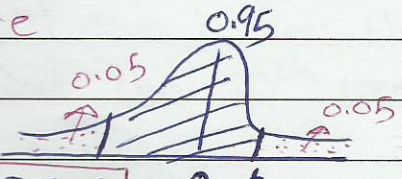
① Find c such that $P(t_{20} > c) = 0.95$

② = = = = $P(t_{20} < c) = 0.01$

③ = = = = $P(t_{19} > c) = 0.99$

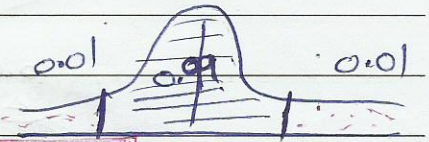
Solve

①



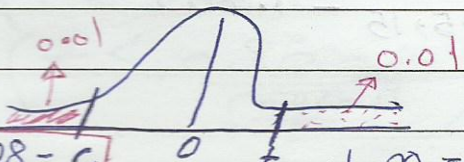
$\boxed{-1.725 = c}$ 0 $t_{0.05,20} = 1.725$

③



$\boxed{-2.539 = c}$ 0 $t_{0.01,19} = 2$

②



$\boxed{-2.528 = c}$ 0 $t_{0.01,20} = 2.528$

Q2

$\mu = 250$

$n = 16$

$S = 25$, Find ① $P(\bar{X} < 230)$

② Find the 95th percentile of $\bar{X}$

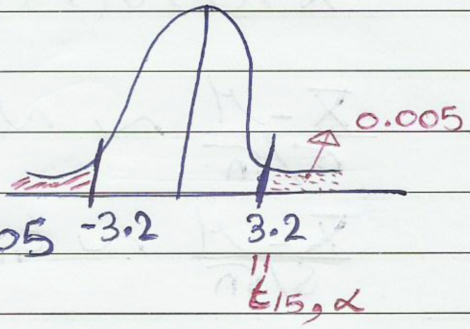
③ = = 5th percentile of $\bar{X}$

Solve ① $P\left(\frac{\bar{X}-\mu}{\cancel{\sigma}/\sqrt{n}} < \frac{230-250}{25/\sqrt{16}}\right)$

$P\left(\frac{\bar{X}-\mu}{S/\sqrt{n}} < -3.2\right)$

$P(t_{n-1} < -3.2)$

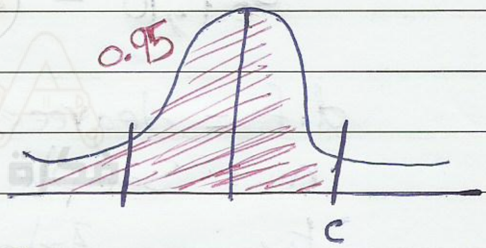
$P(t_{15} < -3.2) = 0.005$



② $P(\bar{X} < c) = 0.95$

$c = \mu_{\bar{x}} + \sigma_{\bar{x}} Z_{\alpha}$

$\mu + \frac{S}{\sqrt{n}} t_{\alpha,n-1}$

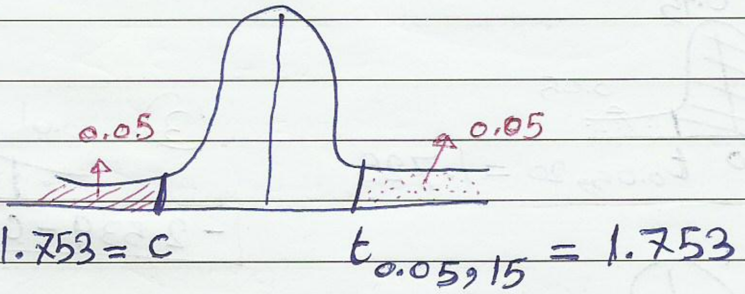


$250 + \frac{25}{4} t_{0.05,15}$

$250 + \frac{25}{4}(1.753)$

③ $P(\bar{x} < c) = 0.05$

$c = 250 + \frac{25}{4}(-1.753)$



Q3

Find ① $P(t_9 < 1.834) = 1 - P(t_9 \geq 1.834)$

$1 - 0.05$

0.95

② $P(t_9 < 2.12) = 1 - P(t_9 \geq 2.12)$

$1 - 0.025$

0.975

Sampling distribution of $\hat{P}$

$X \sim Bi(n, P)$

$\hat{P} = \frac{X}{n}$

$E\hat{P} = \frac{1}{n} Ex = \frac{1}{\cancel{n}} \cancel{n} P = P$

$Var(\hat{P}) = Var\left(\frac{X}{n}\right) = \frac{1}{n^2} Var(X) = \frac{1}{n^2} npq \xrightarrow{\text{For Large } n}$

$\hat{P} \sim N\left(P, \frac{Pq}{n}\right)$

$\frac{\hat{P} - P}{\sqrt{\frac{Pq}{n}}} \sim N(0,1)$

Ex 1

Suppose that 90% of the university students pass Calculus 101, In asample of 200 students taking Calculus 101.

① What is the prop. that the proporation of these who will pass is less than 85%

② Find the 90th percentile of $\hat{P}$.

Solve

$P = 0.9$, $n = 200$, $\hat{P} \sim N\left(0.9, \frac{0.9(0.1)}{200}\right)$

1) $P(\hat{P} < 0.85)$

$P\left(Z < \frac{0.85 - 0.9}{\sqrt{\frac{0.9(0.1)}{200}}}\right) = P(Z < -2.357) = 0.0091$

② $P(\hat{P} < c) = 0.9$

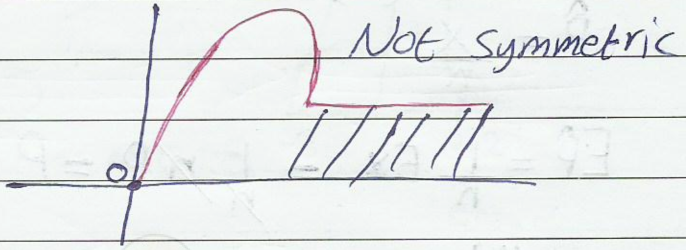
$$c = \mu_{\hat{P}} + \sigma_{\hat{P}} Z_{0.9} = 0.9 + \sqrt{\frac{0.9(0.1)}{200}}\left(\underline{1.282}\right)$$

انتبهوا أن الإشارة أصغر (<) وفي جدول الـ T يجب أن تكون الإشارة أكبر (>)

Sampling distribution of S^2

Let $X_1, \dots, X_4 \sim N(\mu, \sigma^2)$

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$$



$\chi^2_{0.05,5} = 11.07050$

$\chi^2_{0.95,5} = 1.145476$

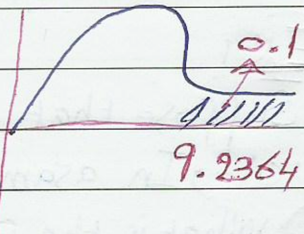
Ex1

If asample of size $n=6$ is drawn from a pop with $\sigma^2 = 10$, find :

① $P(S^2 > 18.4727)$.

② Find the 95th percentile of S^2.

solve 1 $P\left(\frac{(n-1)}{\sigma^2}S^2 > \frac{(6-1)}{10}(18.4727)\right)$



$P(\chi^2_{n-1} > 9.2364)$

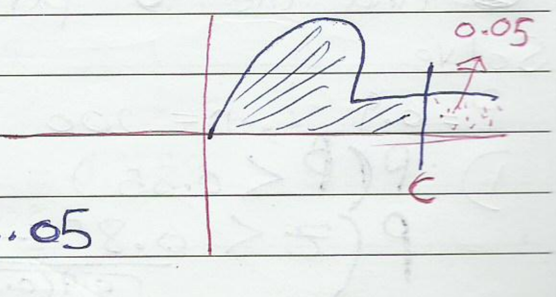
$P(\chi^2_5 > 9.2364) = 0.1$

2

$P(S^2 < c) = 0.95$

$1 - P(S^2 \geq c) = 0.95$

$P(S^2 \geq c) = 0.05$



$P\left(\frac{(n-1)}{\sigma^2}S^2 \geq \frac{5}{100}(c)\right) = 0.05$

$P\left(\chi^2_5 \geq \frac{c}{2}\right) = 0.05$

$\frac{c}{2} = 11.0705$

$c = 2(11.0705)$

Sampling distribution for $\bar{X}-\bar{Y}$

عندما تغير أحد الــ (Parameter) يتغير نوع التوزيع

Let $X_1, \dots, X_{n_1} \sim N(\mu_1, \sigma_1^2)$ } $\perp\!\!\!\perp$
$\quad\; y_1, \dots, y_{n_2} \sim N(\mu_2, \sigma_2^2)$ } (they are independent)

$$\left.\begin{array}{l}\bar{X} \sim N\left(\mu_1, \dfrac{\sigma_1^2}{n_1}\right) \\ \bar{y} \sim N\left(\mu_2, \dfrac{\sigma_2^2}{n_2}\right)\end{array}\right\} \perp\!\!\!\perp \text{ (they are independent)}$$

$N \Longleftarrow N \,\&\, N$ حاصل جمع
$N \Longleftarrow N \,\&\, N$ حاصل طرح

- $E(\bar{X}-\bar{y}) = E\bar{X} - E\bar{y}$
  $\qquad\qquad\quad \mu_1 - \mu_2$
- $\mathrm{Var}(\bar{X}-\bar{y}) = \mathrm{Var}(\bar{X}) + \mathrm{Var}(\bar{y}) + \overset{\bar{X} \perp\!\!\!\perp \bar{y}}{\overset{\uparrow}{0}}$
  $\qquad\qquad\qquad \dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}$
- $\bar{X}-\bar{y} \sim N\left(\mu_1-\mu_2, \dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}\right)$

$$\frac{(\bar{X}-\bar{y})-(\mu_1-\mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} \sim N(0,1)$$

What if σ_1 and σ_2 are unknown.

σ_1 & σ_2 unknown

$$\frac{(\bar{X}-\bar{y})-(\mu_1-\mu_2)}{\sqrt{\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}}} \sim t_{n_1+n_2-2}$$

σ_1 & σ_2 are equal but unknown

$$S_p^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}$$

$$\frac{(\bar{X}-\bar{y})-(\mu_1-\mu_2)}{\sqrt{\dfrac{S_p^2}{n_1} + \dfrac{S_p^2}{n_2}}} \sim t_{n_1+n_2}$$

Ex 1

I

$\mu_1 = 70$

$S_1 = 8$

$n_1 = 10$

II

$\mu_2 = 65$

$S_2 = 10$

$n_2 = 15$

① $P(\bar{X} > \bar{Y})$ if $\sigma_1 = \sigma_2$

② Find 95th percentile of $\bar{X} - \bar{Y}$, $\sigma_1 = \sigma_2$

Solve

كل المتغيرات على طرف والثوابت على طرف آخر

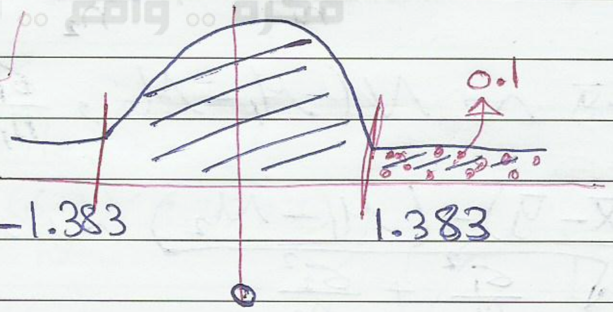
1] $P(\bar{X} > \bar{Y}) = P(\bar{X} - \bar{Y} > 0)$

$$P\left(\frac{\bar{x} - \bar{y} - (\mu_1 - \mu_2)}{\sqrt{\frac{S_p^2}{n_1} + \frac{S_p^2}{n_2}}} > \frac{0-5}{\sqrt{85.913043\left(\frac{1}{10} + \frac{1}{15}\right)}} \right)$$

$$P\left(t_{n_1+n_2-2} > -1.3235 \right)$$

$$P\left(t_{23} > -1.3235 \right) = 0.9$$

$$\boxed{S_p^2 = 85.913043}$$



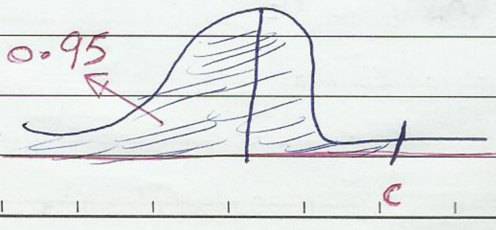
2] $P(\bar{X} - \bar{Y} < c) = 0.95$

$$c = \mu_{\bar{x}-\bar{y}} + \sigma_{\bar{x}-\bar{y}} z$$

$$= 5 + \sqrt{\frac{S_p^2}{n_1} + \frac{S_p^2}{n_2}} \; t_{0.05,23} \rightarrow 1.714$$

$$= 5 + \sqrt{85.913043\left(\frac{1}{10} + \frac{1}{15}\right)} \; t_{0.05,23}$$

$$= 11.48581616$$



Ex 2

Suppose that the grades of Female and male students in Calculus101 are normally distributed with means 70 and 65 respectively and S.d 8 & 10 respectivly. In a sample of 15 Female and 20 male students, Find the prob. that the females will have an average more than the male average.

I	II
$\mu_1 = 70$	$\mu_2 = 65$
$\sigma_1 = 8$	$\sigma_2 = 10$
$n_1 = 15$	$n_2 = 20$

$$P(\overline{X}_1 > \overline{X}_2) = P(\overline{X}_1 - \overline{X}_2 > 0)$$

$$P\left(Z > \frac{0-5}{\sqrt{\frac{8^2}{15}+\frac{10^2}{20}}}\right)$$

$$1 - P(Z \leq -1.64)$$

$$0.9495$$

Sampling distribution for $\hat{P}_1 - \hat{P}_2$

$$\hat{P}_1 \sim N\left(P_1, \frac{P_1 q_1}{n_1}\right)$$

$$\hat{P}_2 \sim N\left(P_2, \frac{P_2 q_2}{n_2}\right)$$

$$\hat{P}_1 - \hat{P}_2 \sim N\left(P_1 - P_2, \frac{P_1 q_1}{n_1} + \frac{P_2 q_2}{n_2}\right)$$

$$\frac{(\hat{P}_1 - \hat{P}_2) - (P_1 - P_2)}{\sqrt{\frac{P_1 q_1}{n_1} + \frac{P_2 q_2}{n_2}}} \sim N(0, 1)$$

Ex 1

Suppose that 50% of population A own cars & 35% of population B own cars. A sample from pop. A with size 100 and from B = 80 are drawn from A and B respectively.

1) Find the prob. that the difference b/w the sample proporations will be b/w 0.1 and 0.2

2) Find the 95th percentile for the differences b/w the sample proportion.

I

$P_1 = 0.5$

$n_1 = 100$

II

$P_2 = 0.35$

$n_2 = 80$

1) $P(0.1 < \hat{P}_1 - \hat{P}_2 < 0.2)$

$$\hat{P}_1 - \hat{P}_2 \sim N\left(P_1 - P_2 , \frac{P_1 q_1}{n_1} + \frac{P_2 q_2}{n_2}\right)$$

$$N\left(0.15 , \frac{0.5(0.5)}{100} + \frac{0.35(0.65)}{80}\right)$$

$$P\left(\frac{0.1 - 0.15}{\sqrt{\frac{0.25}{100} + \frac{0.35(0.65)}{80}}} < Z < \frac{0.2 - 0.15}{\sqrt{\frac{0.25}{100} + \frac{0.35(0.65)}{80}}}\right)$$

2) $P(\hat{P}_1 - \hat{P}_2 < c) = 0.95$

$$C = \mu_{\hat{P}_1 - \hat{P}_2} + \sigma_{\hat{P}_1 - \hat{P}_2} Z_{0.95}$$

$$= 0.15 + \sqrt{\frac{0.25}{100} + \frac{0.35(0.65)}{80}} \, (1.645)$$

Sampling distribution of S_1^2/S_2^2

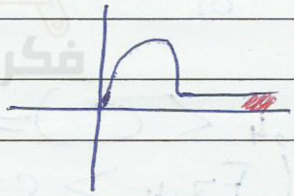
if $X_1, X_2, \dots, X_{n_1} \sim N(\mu_1, \sigma_1^2)$ } $\perp\!\!\!\perp$
$y_1, y_2, \dots, y_{n_2} \sim N(\mu_2, \sigma_2^2)$]

$\frac{(n_1-1)S_1^2}{\sigma_1^2} \sim \chi^2_{n_1-1}$ } indep.
$\frac{(n_2-1)S_2^2}{\sigma_1^2} \sim \chi^2_{n_2-1}$]

~~$n_1=n_2$~~

$$\frac{\frac{(n_1-1)S_1^2}{\sigma_1^2}}{\frac{(n_2-1)S_2^2}{\sigma_2^2}} \sim \frac{\chi^2_{n_1-1}}{\chi^2_{n_2-1}}$$

$$\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \sim F_{n_1-1, n_2-1}$$



$$\frac{S_2^2}{S_1^2} \cdot \frac{\sigma_1^2}{\sigma_2^2} \sim F_{n_2-1, n_1-1}$$

$$\frac{S_1^2}{S_2^2} \cdot \frac{\sigma_2^2}{\sigma_1^2} \sim F_{n_1-1, n_2-1}$$

* Important *_^

1) $F_{3,4,0.1} = 4.19$

2) $F_{7,10,0.05} = 3.14$

3) $F_{20,11,0.01} = 4.1$

*4) $F_{11,20,0.99} = \frac{1}{F_{20,11,0.01}} = \frac{1}{4.1}$

*5) $F_{4,3,0.9} = \frac{1}{F_{3,4,0.1}} = \frac{1}{4.19}$

Ex1 $\sigma_1^2 = 50$ $\sigma_2^2 = 20$

$n_1 = 5$ $n_2 = 8$

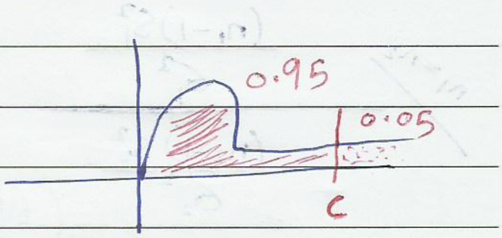
① Find $P\left(\frac{S_1^2}{S_2^2} > 7.4\right)$?

$$P\left(\frac{\sigma_2^2}{\sigma_1^2} \cdot \frac{S_1^2}{S_2^2} > \frac{20}{50}(7.4)\right)$$

$$P(F_{n_1-1, n_2-1} > 2.96)$$

$$P(F_{4,7} > 2.96) = 0.1$$

② Find the 95th percentile of $\frac{S_2^2}{S_1^2}$



$$P\left(\frac{S_2^2}{S_1^2} < c\right) = 0.95$$

$$P\left(\frac{S_2^2}{S_1^2} \cdot \frac{\sigma_1^2}{\sigma_2^2} < c^*\right) = 0.95$$

$$P(F_{7,4} < c^*) = 0.95$$

$$1 - P(F_{7,4} \geq c^*) = 0.95$$

$$P\left(F_{n_2-1, n_1-1} \geq \frac{50}{20} c\right) = 0.05$$

$$P\left(F_{7,4} \geq \frac{5c}{2}\right) = 0.05$$

$$\frac{5}{2} c = 6.09$$

$$c = \frac{2(6.09)}{5}$$

$\mu \quad \bar{X}$ → Point estimate تقدير نقطي

$\sigma^2 \quad S^2$

$p \quad \hat{p}$

Confidence interval

<CI>

$\bar{X} \mp$ error

$(\bar{X} -$ error $, \bar{X} +$ error $)$

$$X_1 \dots X \sim N(\mu, \sigma^2)$$

$$\bar{X} \sim N(\mu, \sigma^2/n)$$

$\frac{\sigma}{\sqrt{n}}$ $Sd(\bar{X})$ standard error

EX1

	Size	Sample mean	Sample variance
Sample 1	5	100	400
Sample 2	10	90	200

Assume $\sigma_1 = \sigma_2$

① Find Point estimate for $\mu_1 - \mu_2$

② Find standerd error for the Point estimate.

③ Construct 95% CI for $\mu_1 - \mu_2$

1) $\bar{X}_1 - \bar{X}_2 = 100 - 90 = 10$

2) $S.d(\bar{X}_1 - \bar{X}_2) = \sqrt{\frac{S_p^2}{n_1} + \frac{S_p^2}{n_2}}$

$$S_p^2 = \frac{(n_1 - 1) S_1^2 + (n_2 - 1) S_2^2}{n_1 + n_2 - 2} = 261.54$$

3) $1 - \alpha = 0.95 \Rightarrow \alpha = 0.05 \Rightarrow \alpha/2 = 0.025$

$$\bar{X}_1 - \bar{X}_2 \pm \sqrt{\frac{S_p^2}{n_1} + \frac{S_p^2}{n_2}}\, t_{\alpha/2, n_1 + n_2 - 2}$$

$$10 \pm \sqrt{261.54\left(\frac{1}{5} + \frac{1}{10}\right)}\, t_{0.025, 13}$$

$(-9.13, 29.13)$

Ex 2 Sample of 60 students gave mean score = 48.2 the population s.d = 9.6.
Find 95% CI for μ?

$n = 60$
$\bar{x} = 48.2$
$\sigma = 9.6$

$1-\alpha = 0.95 \Rightarrow \alpha = 0.05 \Rightarrow \frac{\alpha}{2} = 0.025$

$\bar{x} \pm \sqrt{\frac{\sigma^2}{n}}\ z_{\alpha/2}$

$\bar{x} \pm \frac{\sigma}{\sqrt{n}}\ z_{\alpha/2}$

$48.2 \pm \frac{9.6}{\sqrt{60}}\ \overset{1.96}{\boxed{z_{0.025}}} \longrightarrow$ جدول الـ T أسهل
$\vdots$
∞

$(45.77, 50.63)$

Ex 3 كان يعتقد

It was belived in the Arab world that 50% of persons are smoking. During the Year 2000, a sample of 1000 persons showed that the number of smokers is 620.
1) Construct 95% CI for the proportion of smokers.
2) Can you conclude that the proportion of smokers is different from 50%.

$\left.\begin{array}{l} n = 1000 \\ x = 620 \end{array}\right\} \hat{p} = \frac{x}{n} = 0.62$

$1-\alpha = 0.95 \Rightarrow \alpha = 0.05 \Rightarrow \alpha/2 = 0.025$

Solve
① $\hat{p} \pm \sqrt{\frac{\hat{p}\hat{q}}{n}}\ z_{\alpha/2}$

$0.62 \pm \sqrt{\frac{0.62(1-0.62)}{1000}}\ \overset{1.96}{\boxed{z_{0.025}}}$

$(0.59, 0.65)$

② Yes, b/c $0.5 \notin$ CI

Ex 4

	Males	Females
# of smokers	45	25
# of smokers	35	45
	80	70

P_1 : Proportion of smokers / Males $\left(\hat{P}_1 = \frac{45}{80}\right)$
P_2 : Proportion of smokers / females $\left(\hat{P}_2 = \frac{25}{70}\right)$

Construct 95% CI for $P_1 - P_2$?

$$\hat{P}_1 - \hat{P}_2 \pm \sqrt{\frac{\hat{P}_1 \hat{q}_1}{n_1} + \frac{\hat{P}_2 \hat{q}_2}{n_2}}\ Z_{\alpha/2}$$

$$(0.56 - 0.36) \pm \sqrt{\frac{0.56(1-0.56)}{80} + \frac{0.36(1-0.36)}{70}}\ \overset{1.96}{\boxed{Z_{0.025}}}$$

Sample size needed to Construct C.I

$(1-\alpha)\%$ CI for μ

$$\bar{x} \pm \frac{\sigma}{\sqrt{n}} Z_{\alpha/2}$$

$$\text{Error} = \frac{\sigma}{\sqrt{n}} Z_{\alpha/2}$$

Ex 1 Find the sample size needed to Construct 95% CI for μ Given $E = 0.1$
$\sigma = 0.4$

$$\text{Error} = \frac{\sigma}{\sqrt{n}} Z_{\alpha/2}$$

$$0.1 = \frac{0.4}{\sqrt{n}} \overset{1.96}{\boxed{Z_{0.025}}}$$

$1 - \alpha = 0.95$
$\alpha = 0.05$
$\alpha/2 = 0.025$

$n = 61.4656$

$n \approx 62$ ⟿ Round up ↑ لحجم العينة ككل

$(1-\alpha)\%$ CI for P

$$\hat{P} \pm \sqrt{\frac{pq}{n}}\, z_{\alpha/2}$$

$$E = \sqrt{\frac{pq}{n}}\, z_{\alpha/2}$$

P

No prior info.
$P = \frac{1}{2}$

Previous studies
مثل بحث مع احصائية

Ex 1 Assume that it is requires to estimate P using 95% C.I if Error = 0.1
1) No Prior information
2) Previous study showed $P \approx 0.2$

① $E = \sqrt{\frac{pq}{n}}\, \overset{1.96}{z_{\alpha/2}}$

$0.1 = \sqrt{\frac{(0.5)^2}{n}}\,(1.96)$

$n = 96.04$
$n \approx 97$

$E = \sqrt{\frac{pq}{n}}\, z_{\alpha/2}$

$0.1 = \sqrt{\frac{0.2(0.8)}{n}}\,(1.96)$

$n = 61.4656$
$n \approx 62$

Chaptter 6

Testing Hypothesis

$H_0 \quad \mu \leq 200 \quad Vs \quad H_1 \quad \mu > 200$

$H_0 \quad \mu \geq 250 \quad Vs \quad H_1 \quad \mu < 250$

$H_0 \quad \mu = 0.05\ mg \quad Vs \quad H_1 \quad \mu \neq 0.05\ mg$

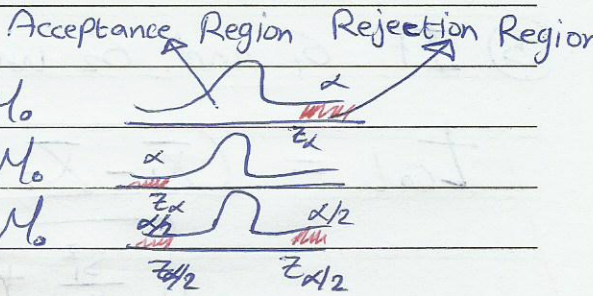
$$Z_{cal} = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}$$

① * Test about μ

$H_0 : \mu = \mu_0$ معطى → (ثابت) $\quad Vs \quad H_1 : \mu > \mu_0$

$H_0 : \mu = \mu_0 \quad Vs \quad H_1 : \mu < \mu_0$

$H_0 : \mu = \mu_0 \quad Vs \quad H_1 : \mu \neq \mu_0$



Test stat

$$\boxed{Z_{cal} = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}}$$

② if σ is unknown

$$\boxed{t_{cal} = \frac{\bar{X} - \mu_0}{S / \sqrt{n}} \sim t_{n-1}}$$

③ if $H_0 : P = P_0 \quad Vs \quad H_1 : P > P_0$

$P < P_0$

$P \neq P_0$

Test stat

$$\boxed{Z_{cal} = \frac{\hat{P} - P_0}{\sqrt{\frac{P_0 q_0}{n}}} \sim N(0,1)}$$

④ $H_0 : \mu_1 = \mu_2$ Vs $H_1 : \mu_1 > \mu_2$ $\quad \mu_1 - \mu_2 > 0$

$\mu_1 < \mu_2$ $\quad \mu_1 - \mu_2 < 0$

$\mu_1 \neq \mu_2$ $\quad \mu_1 - \mu_2 \neq 0$

Test stat :

$$\boxed{Z_{cal} = \frac{(\bar{X}_1 - \bar{X}_2) - \overset{\text{Zero}}{\cancel{(\mu_1 - \mu_2)}}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1)}$$

⑤ If σ_1 and σ_2 unknown

$$\boxed{t_{cal} = \frac{(\bar{X}_1 - \bar{X}_2) - \overset{\text{Zero}}{\cancel{(\mu_1 - \mu_2)}}}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \sim t_{n_1+n_2-2}}$$

⑥ σ_1 and σ_2 are unknown and equal

$$\boxed{t_{cal} = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S^2p}{n_1} + \frac{S^2p}{n_2}}} \sim t_{n_1+n_2-2}}$$

⑦ $H_0 : P_1 = P_2$ Vs $H_1 : P_1 > P_2 = P_1 - P_2 > 0$

$P_1 - P_2 = 0$ $\quad P_1 < P_2 = P_1 - P_2 < 0$

Test stat $\quad P_1 \neq P_2 = P_1 - P_2 \neq 0$

$$\boxed{Z_{cal} = \frac{(\hat{P}_1 - \hat{P}_2) - \overset{\text{Zero}}{\cancel{(P_1 - P_2)}}}{\sqrt{\frac{\bar{p}\bar{q}}{n_1} + \frac{\bar{p}\bar{q}}{n_2}}} \sim N(0,1)}$$

$$\bar{P} = \frac{n_1 \hat{P}_1 + n_2 \hat{P}_2}{n_1 + n_2}$$

$\underbrace{\bar{p} \qquad \bar{q}}$ → مشتركة (S^2p مثل)

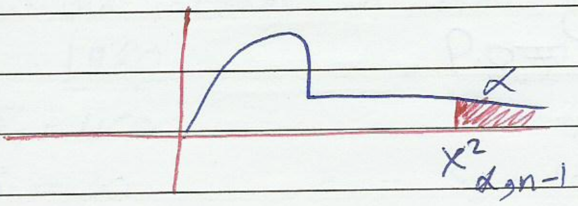
حيث $P_1 = P_2$

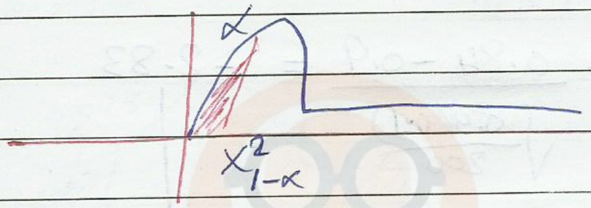
حروق ⑧ $H_0: \sigma^2 = \sigma_0^2$ Vs $H_1: \sigma^2 > \sigma_0^2$

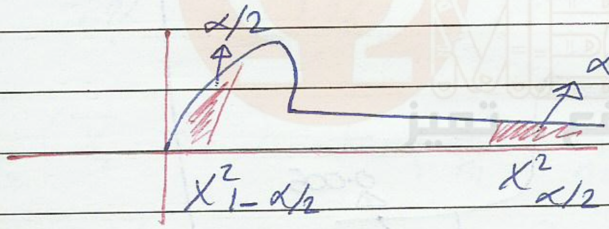
$\sigma^2 < \sigma_0^2$

Test stat $\sigma^2 \neq \sigma_0^2$

$$\boxed{\chi^2_{cal} = \frac{(n-1)S^2}{\sigma_0^2} \sim \chi^2_{n-1}}$$





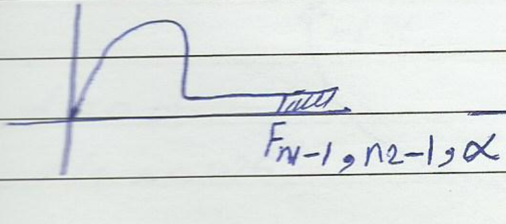


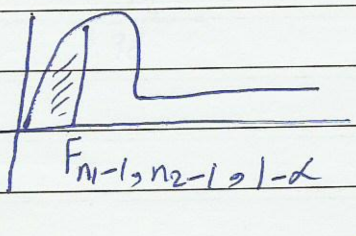
حروق ⑨ $H_0: \sigma_1^2 = \sigma_2^2$ Vs $H_1: \sigma_1^2 > \sigma_2^2$ | $\frac{\sigma_1^2}{\sigma_2^2} > 1$

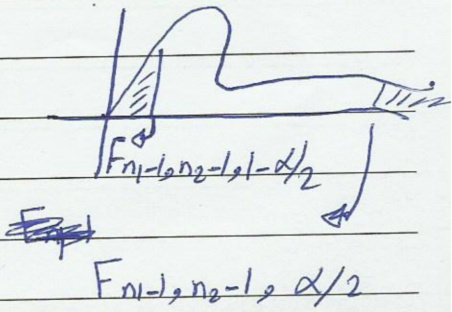
$\sigma_1^2 < \sigma_2^2$ | $\frac{\sigma_1^2}{\sigma_2^2} < 1$

$\sigma_1^2 \neq \sigma_2^2$ | $\frac{\sigma_1^2}{\sigma_2^2} \neq 1$

$$F_{cal} = \frac{S_1^2}{S_2^2} \cdot \frac{\sigma_2^2}{\sigma_1^2} \sim F_{n_1-1,n_2-1}$$







Ex

يدّعي

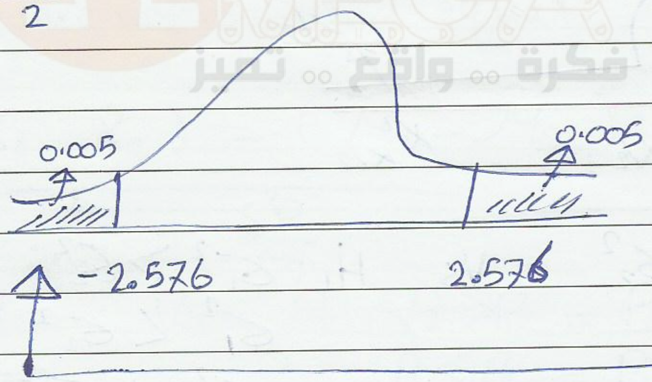
The manufacturel of amedicin claims it is 90% effective in
حساسية
relieving allergy Symptoms for aperiod of 8 hours. What Conclusion would you draw if from a r.S of 200 people with
حساسية
allergy, the medicin provided relief for 168 People use Significance Level ($\alpha = 0.01$)

$H_0 : P = 0.9$ Vs $H_1 : P \neq 0.9$

$$\hat{P} = \frac{168}{200} = 0.84$$

Test Stat $Z_{Cal} = \frac{\hat{P} - P_0}{\sqrt{\frac{P_0 q_0}{n}}} = \frac{0.84 - 0.9}{\sqrt{\frac{0.9(0.1)}{200}}} = -2.83$

$$\frac{\alpha}{2} = \frac{0.01}{2} = 0.005$$



منطقة رفض

Reject H_0

مش وافقين انه الكلام المفروض بالسؤال صحيح (بعبارة أخرى)
الكلام المفروض بالسؤال ليس صحيح

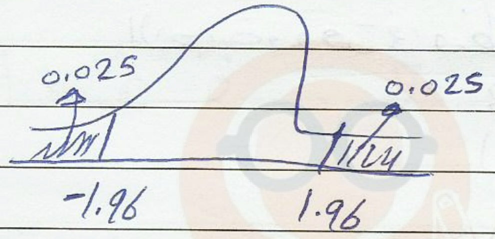
Ex The mean and s.d of the income of Jordan Families were 4000 and 300 JD in 1980 and 4500 and 400 JD in 1990 Assume these Values are basses on sample size of 100 families for 1980 and 150 Families for 1990.

① Construct 95% CI for $\mu_1 - \mu_2$

② Test at 5% Significance level if these is any difference b/w the incomes in 1980 and 1990.

1980	1990
$\bar{X}_1 = 4000$	$\bar{X}_2 = 4500$
$S_1 = 300$	$S_2 = 400$
$n_1 = 100$	$n_2 = 150$

①



$1-\alpha = 0.95$

$\alpha = 0.05$

$\alpha/2 = 0.025$

$$(\bar{X}_1 - \bar{X}_2) \pm \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}} \; t_{\alpha/2, n_1+n_2-2}$$

$(-558.9 , -441.1)$

② $0 \notin (-558.9 , -441.3)$

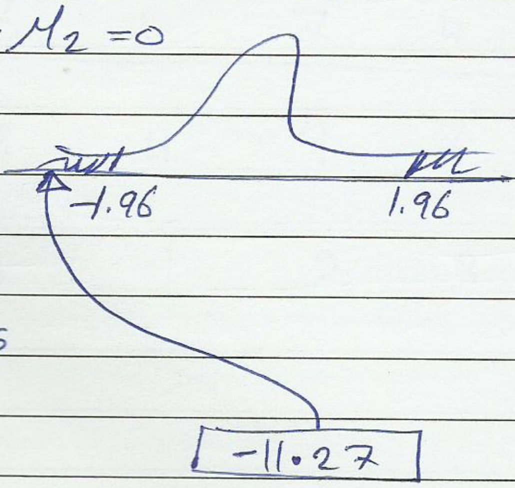
Reject

نستعمل هذه الطريقة عندما تكون
$H_1 : \mu_1 \neq \mu_2$ فقط

(طريقة أخرى)

$H_0 : \mu_1 = \mu_2$ Vs $H_1 : \mu_1 \neq \mu_2$

$\mu_1 - \mu_2 = 0$ $\mu_1 - \mu_2 \neq 0$



$\alpha = 0.05$

$\alpha/2 = 0.025$

Test Stat

$$t_{cal} = \frac{(\bar{X}_1 - \bar{X}_2) - \overset{zero}{\cancel{(\mu_1 - \mu_2)}}}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

$$\frac{4000 - 4500}{\sqrt{\frac{(300)^2}{100} + \frac{(400)^2}{150}}} = \boxed{-11.27}$$

الاجابة وقعت في منطقة رفض

$n_1 + n_2 - 2 = 250 - 2 = 248$

له آخر سطر بجدول t

Reject H_0

Paired Samples

Before	After	d = Before − After
X_1	y_1	$d_1 = X_1 - y_1$
⋮	⋮	$d_2 = X_2 - y_2$
		⋮
X_n	y_n	$d_n = X_n - y_n$

$(1-\alpha)\%$ CI for μ_D

$$\bar{d} \pm \frac{sd}{\sqrt{n}} t_{\alpha/2, n-}$$

$d_1, d_2, \ldots, d_n \Rightarrow$ Find $\bar{d}$ and S_d

$H_0: \mu_D = 0$ Vs $H_1: \mu_D > 0 \Rightarrow$ before > After

$\mu_D < 0$

$\mu_D \neq 0$

Test stat:

$$t_{cal} = \frac{\bar{d} - \mu_D}{sd/\sqrt{n}} \sim t_{\alpha, n-1}$$

key

Before , After and جدول

Ex

Before(B)	6	5	7	4	5
After(A)	9	4	9	7	6
D	3	-1	2	3	1

After > Before

d= After − Before

$\mu_D > 0$

→

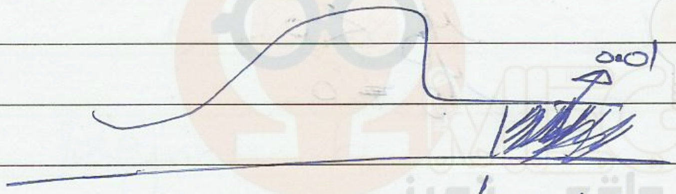
Can You Conclude that the medicin is effective in increasing the sleep hours ? $\alpha = 0.01$

$H_0 : \mu_D = 0$ Vs $H_1 : \mu_D > 0$

Test Stat

$$t_{cal} = \frac{\bar{d} - \mu_D}{sd/\sqrt{n}}$$

$$\frac{1.6}{1.67/\sqrt{5}} = 2.13$$



$t_{0.01,4} = 3.747$

$\bar{d} = 1.6 \left(\frac{8}{5}\right)$

$sd = 1.67$

Accept H_0

Ex

A random sample of 15 observation
10.7 , 11.9 , 10.4 , 11.5 , 9.3 , 10. 4 , 11.6 , 9.7 ,
10.1 , 10.5 , 9.1 , 10.4 , 10. 3 , 9.8 , 10.9

1) Construct 95% CI for the population (Mean)
2) at significance level $\alpha = 0.05$. Can you conclude the μ is different from 10.

① $\bar{X} = 10.44$ $S = 0.8052$ $n = 15$

$1 - \alpha = 0.95 \Rightarrow \alpha = 0.05 \Rightarrow \alpha/2 = 0.025$

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

$$\bar{X} \pm \frac{S}{\sqrt{n}} t_{\alpha/2, n-1}$$

$$10.44 \pm \frac{0.8052}{\sqrt{15}} \overset{2.145}{\boxed{t_{0.025, 14}}}$$

$(9.994 , 10.886)$

② $10 \in C.I$
Accept H_0

$H_0 : \mu = 10$ Vs $H_1 : \mu \neq 10$