Addition of Mean and variance

Example ① X and Y are two R.V with the following joint PMF :-

$x \backslash y$	-1	0	1
-1	$1/8$	$1/2$	0
1	0	$1/4$	$1/8$

ⓐ $E\{xy\} = ??$

$$E\{g(x,y)\} = \sum_{y=-\infty}^{\infty}\sum_{x=-\infty}^{\infty} g(x,y)\, P(X=x, Y=y)$$

$$= \sum_{y=-\infty}^{\infty}\sum_{x=-\infty}^{\infty} x\,y\, P(X=x, Y=y)$$

$$= (-1)(-1)P(X=-1, Y=-1) + \underset{x=-1}{(-1)}\underset{y=0}{(0)}\, P(X=-1, Y=0) + \dots$$

$$= (-1)(-1)(1/8) + (-1)\cancel{(0)}(1/2) + (-1)\cancel{(1)}(0) + (1)\cancel{(-1)}(0) + (1)\cancel{(0)}(1/4) + (1)(1)(1/8)$$

$$= \frac{1}{8} + \frac{1}{8} = \frac{2}{8}$$

ⓑ $E\{x^2 y\} = \sum_{y=-\infty}^{\infty}\sum_{x=-\infty}^{\infty} x^2 y\, P(X=x, Y=y)$

$$= (-1)^2(-1)(1/8) + 0 + 0 + 0 + 0 + (1)(1)(1/8)$$

$$= \cancel{-1/8} \; \cancel{+1/8}$$

$$= 0$$

ⓒ $E\{(x+1)y\} = ?$

$$= \sum_{y=-\infty}^{\infty}\sum_{x=-\infty}^{\infty} (x+1)\,y\, P(X=x, Y=y) = (0)\cancel{(-1)}\left(\tfrac{1}{8}\right) + 0 + 0$$

$$+ (1+1)\cancel{(-1)}(0) + (1+1)\cancel{(0)}\left(\tfrac{1}{4}\right) + (1+1)(1)(1/8)$$

$$= \frac{2}{8} = \frac{1}{4}$$

نحط ما دكال probability
بتع $(x=1, y=1)$
بدونه ال $x+1$ لل y

Example X and Y are two RVs with the following joint pdf :-

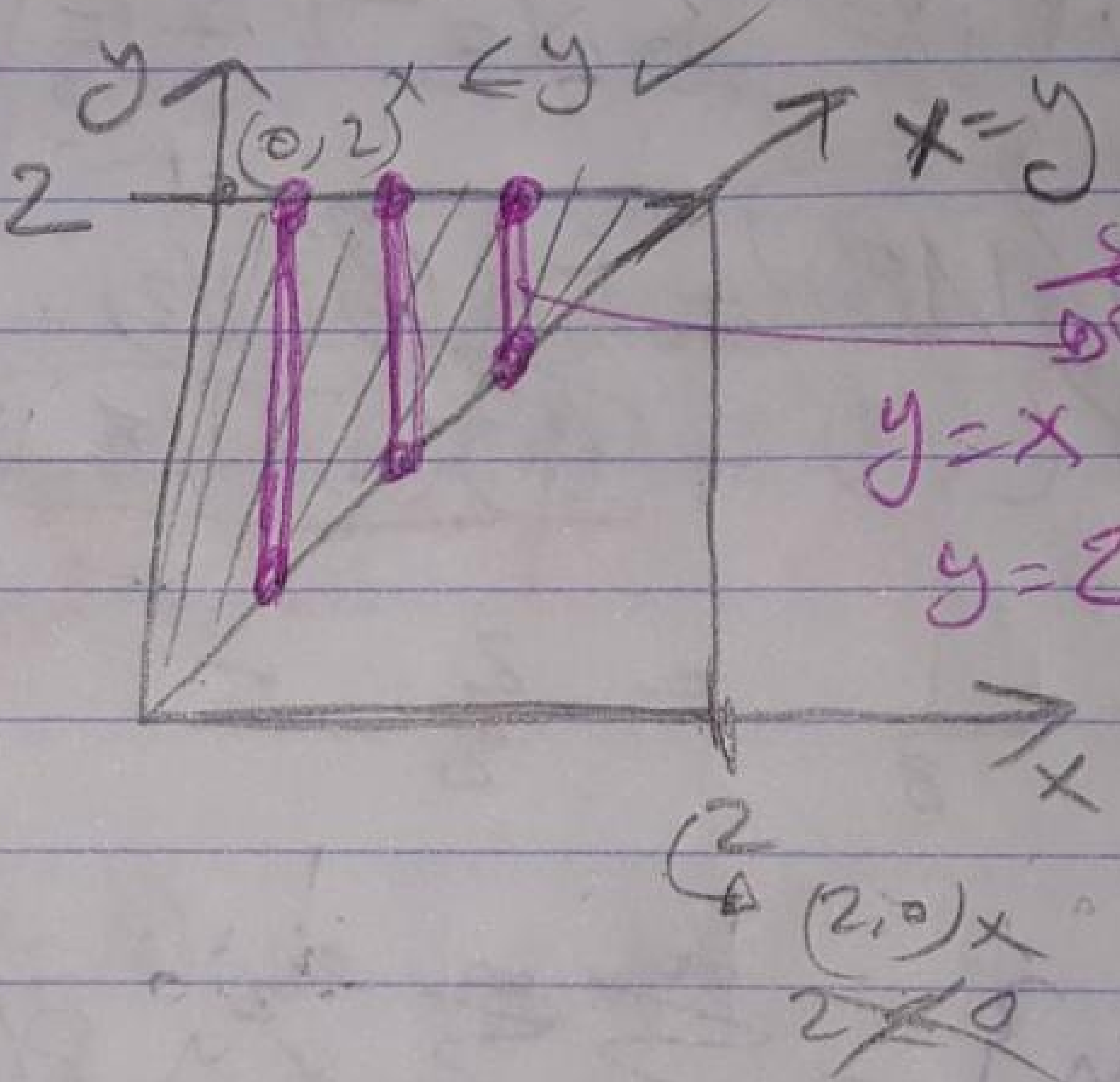
$$f_{x,y}(x,y) = \begin{cases} K x^2 y , & 0 \le x \le y \le 2 \\ 0 , & \text{o.w} \end{cases}$$

ⓐ Determine the value of the constant K.

معنى هاي الفترة :-

$$\boxed{\begin{array}{c} 0 \le x \le 2 \\ 0 \le y \le 2 \\ x \le y \end{array}}$$

خبرسم ال area اللي بتحددها هاي الشروط



* ملاحظة هامة : بوضع الرسمة بنجد الفترة بغض النظر عن ال function هاي

~~...~~ بيبدأ الشرائح من $y=x$ وبتنتهي عند $y=2$

إذاً حل السؤال :-

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{x,y}(x,y)\,dy\,dx = 1$$

$$\int_0^2 \left[\int_x^2 Kx^2 y\,dy\right] dx = 1 \Rightarrow \int_0^2 Kx^2 \frac{y^2}{2}\Big|_x^2 dx$$

$$= \int_0^2 \frac{Kx^2}{2}\left[4 - x^2\right] dx = \int_0^2 \frac{K}{2}\left[4x^2 - x^4\right] dx$$

$$= \frac{K}{2}\left[\frac{4x^3}{3} - \frac{x^5}{5}\right]_0^2 = 1 \Rightarrow \frac{K}{2}\left[\frac{32}{3} - \frac{32}{5}\right] = 1$$

$$= \frac{K}{2}\left[\frac{160}{3\times5} - \frac{96}{3\times5}\right] = \frac{K}{2} * \frac{64}{15} = \frac{32}{15}K = 1$$

$$\therefore \boxed{K = \frac{15}{32}}$$

Scanned by TapScanner

Interval → area

(b) $E\{x(y+1)\} = ??$

صوالكي الي موجود جوا

جيب بينضرب باد $f_{xy}^{(x)}$

$$E\{x(y+1)\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x(y+1)\, f_{xy}(x,y)\, dy\, dx$$

$$= \int_0^2\int_x^2 x(y+1)\, x^2 y\, dy\, dx = \frac{15}{32}\int_0^2\int_x^2 (x^3y^2 + x^3y)\, dy\, dx$$

$$= \frac{15}{32}\int_0^2 \left(\frac{x^3y^3}{3} + \frac{x^3y^2}{2}\right)\Big|_x^2 dx = \frac{15}{32}\int_0^2 x^3\left[\left(\frac{8}{3}+2\right) - \left(\frac{x^3}{3} + \frac{x^2}{2}\right)\right]$$

$$= \frac{15}{32}\left(\frac{8}{3}\frac{x^4}{4} + \frac{2x^4}{4} - \frac{x^7}{21} - \frac{x^6}{12}\right)\Big|_0^2$$

$= \ldots\ldots\ldots$

Notes:- (I) $E\{x+y\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x+y)\, f_{xy}(x,y)\, dy\, dx$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x\, f_{xy}(x,y)\, dy\, dx + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y\, f_{xy}(x,y)\, dy\, dx$$

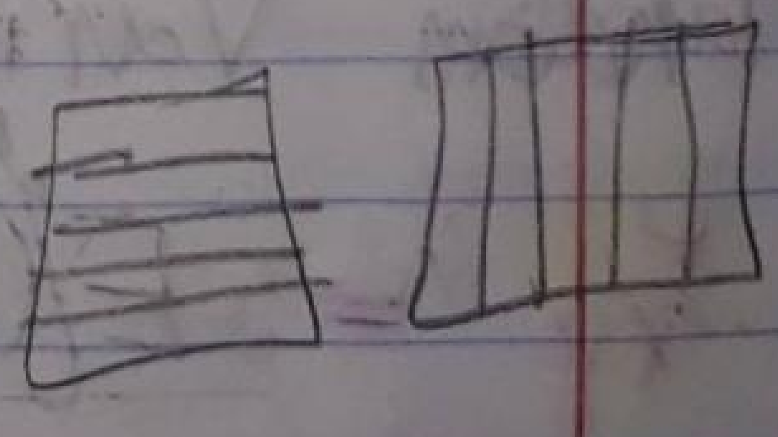
$$= \int_{-\infty}^{\infty} x \int_{-\infty}^{\infty} f_{(x,y)}(x,y)\, dy\, dx + \int_{-\infty}^{\infty} y \int_{-\infty}^{\infty} f_{xy}(x,y)\, dx\, dy$$

عملنا تبديل بين dy وdx عشان الحدود ضلها زي ماهي يعني سواء أخذنا الslices بالطول (dx) أو بالعرض (dy) كل إشي رح يضل زي ماهو

$$= \int_{-\infty}^{\infty} x\, f_x(x) + \int_{-\infty}^{\infty} y\, f_y(y)\, dy$$

$$= E\{x\} + E\{y\}$$

وإذا مضروبين بأي constant عادي كل إشي بينزل زي ماهو



II) $E\{a_1 xy\} = ??$ Only if (x) and (y) are Statistically Independent

$$E\{a_1 xy\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} a_1 xy\, f_{x,y}(x,y)\, dy\, dx$$

⇓ If x & y are SI :-

$$f_{x,y}(x,y) = f_x(x)\, f_y(y)$$

$$= \int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} a_1 xy\, f_x(x)\, f_y(y)\, dy\right] dx$$

$$= \int a_1 x\, f_x(x) \left[\int_{-\infty}^{\infty} y f_y(y)\, dy\right] dx$$

$$= \int_{-\infty}^{\infty} a_1 x f_x(x)\, \underline{\mu_y}\, dx$$

constant عشان هيك اطلعت برا

$$= a_1 \mu_y \underbrace{\int_{-\infty}^{\infty} x f_x(x)\, dx}_{\mu_x} = a_1 \mu_y \mu_x = a_1 E\{x\} E\{y\}$$

$$\therefore E\{a_1 xy\} \underset{\text{only if X \& Y are SI}}{=\!=} a_1 E\{x\} E\{y\}$$

Theorem3- Addition of Variances

Definition:- The correlation coefficient between two random variables (X) and (y) is:-

(correlation: علاقة ، coefficient: معامل)

$$\rho_{xy} = \frac{E\{(X-\mu_x)(Y-\mu_y)\}}{\sigma_x \sigma_y} = \frac{\mu_{xy}}{\sigma_x \sigma_y}$$

⇒

مشترك : Co
توزيع شرطي : Variance

where μ_{xy} is called the covariance.

$\rho_{x,y}$ is bounded between $-1 \le \rho_{x,y} \le 1$

ما في بينهم علاقة ، يعني إذا عرفت قيمة x بعرفش أحدد قيمة y من خلال قيمة x

when $\rho_{xy} = 0$, X and Y are said to be Uncorrelated

when $\rho_{x,y} = \pm 1$, x and y are said to be fully correlated

يعني إذا بعرف قيمة x بقدر مباشرة أقول كم قيمة y .

* إذا كان $\rho_{x,y}$ سالب معناه إنو العلاقة بينهما عكسية ، إذا زاد x يقل y ، أو إذا قل x يزداد واي .

* أما إذا كان $\rho_{x,y}$ موجب معناه إنو العلاقة بينهما طردية إذا زاد (x) يزداد y ، وإذا قل (x) يقل y .

Talking about the covariance :-

$$\mu_{xy} = E\{(x-\mu_x)(y-\mu_y)\}$$

$$= E\{xy - x\mu_y - y\mu_x + \mu_x\mu_y\}$$

$$= E\{xy\} - \mu_y E\{x\} - \mu_x \cancel{E\{y\}} + \cancel{\mu_x\mu_y}$$

$$\therefore \boxed{\mu_{xy} = E\{xy\} - \mu_x\mu_y}$$

* إذا عرفت ال covariance بقدر بكل سهولة أجيب ال correlation coefficient

Example: X and Y are two RVs (Random variables) with the following joint PMF:

x \ y	-1	1
-1	1/4	1/4
1	1/4	1/4

a Determine $\rho_{x,y} = ??$

$$\rho_{xy} = \frac{\mu_{xy}}{\sigma_x \sigma_y}, \quad \mu_{xy} = E\{(x-\mu_x)(y-\mu_y)\}$$

$$P(X=x) = \begin{cases} 1/2 & x=-1 \\ 1/2 & x=1 \\ 0 & o.w \end{cases}, \quad P(Y=y) = \begin{cases} 1/2 & y=-1 \\ 1/2 & y=1 \\ 0 & o.w \end{cases}$$

$$\therefore \mu_x = \sum_{-\infty}^{\infty} x\, P(X=x) = (-1)(1/2) + 1*(1/2) = 0$$

$$\mu_y = \sum_{-\infty}^{\infty} y\, P(Y=y) = (-1)(1/2) + (1)(1/2) = 0$$

$$\mu_{xy} = E\{(x-\cancel{\mu_x}^{0})(y-\cancel{\mu_y}^{0})\} = E\{xy\}$$

$$= \sum_{y=-\infty}^{\infty}\sum_{x=-\infty}^{\infty} x\, y\, P(X=x, Y=y)$$

$$= (-1)(-1)(1/4) + \cancel{(-1)(1)(1/4)} + \cancel{(1)(-1)(1/4)} + \cancel{(1)(1)(1/4)}$$

$$= 0$$

$$\therefore \rho_{xy} = \frac{0}{\sigma_x \sigma_y} = 0 \leftarrow \therefore \text{X and Y are Uncorrelated.}$$

b Are X and Y Statistically Independent?

$$P(X=x, Y=y) \overset{?}{=} P(X=x)\,P(Y=y) \text{ for all values of } x \,\&\, y$$

$\langle x=-1, Y=1\rangle \rightarrow \frac{1}{4} \overset{\checkmark}{=} \frac{1}{2} * \frac{1}{2}$

$\langle x=-1, Y=1\rangle \rightarrow \frac{1}{4} \overset{\checkmark}{=} \frac{1}{2} * \frac{1}{2}$

$\langle x=1, Y=-1\rangle \rightarrow \frac{1}{4} \overset{\checkmark}{=} \frac{1}{2} * \frac{1}{2}$

$\langle x=1, Y=1\rangle \rightarrow \frac{1}{4} \overset{\checkmark}{=} \frac{1}{2} * \frac{1}{2}$

So X and y are Statistically Independent

* If x and y are Independent, then they are Uncorrelated ($\rho_{x,y}=0$), But the converse is NOT Necessarily true.

As :- $\mu_{xy} = E\{(x-\mu_x)(Y-\mu_Y)\}$

$= E\{xy\} - \mu_x \mu_Y$

$\downarrow$ if x and y are SI:-

$= E\{x\} E\{Y\} - \mu_x \mu_y$

$= 0$

يعني إذا مايلي سؤال بالأول هل x و y ستاستكلي independent

وطلع معي yes

والشي التاني سألني شو الـ correlation coefficient بقوله

طول بحكيله إنو = 0 .

$\boxed{\text{Example}}$ let x be a R.V with $\boxed{\mu_x = 1}$ and $\boxed{\sigma_x^2 = 4}$. Y is another R.V with $\boxed{\mu_y = -1}$ and $\boxed{\sigma_y^2 = 9}$, $\boxed{R = 2x - y}$, and $\boxed{\rho_{xy} = 0.5}$

(a) $\mu_R = ??$

$\mu_R = 2\mu_x - \mu_y$

$= 2(1) + 1$

$= 3$

تابع →

Scanned by TapScanner

(b) $Var\{R\} = ??$ $\quad R = 2x - y \Rightarrow a_1 x + a_2 y$ $\quad (a_1 = 2, \; a_2 = -1)$

$$Var\{R\} = \sigma_R^2 = (a_1)^2 \sigma_x^2 + (a_2)^2 \sigma_y^2 + 2 a_1 a_2 \sigma_x \sigma_y \rho_{xy}$$

$$= (2)^2(4) + (-1)^2(9) + 2(2)(-1)(\sqrt{4})(\sqrt{9})(0.5)$$

$$= 16 + 4 + (-12) = 13$$

* Theorem : let $Y = a_1 X_1 + a_2 X_2$, then

$$\sigma_y^2 = a_1^2 \sigma_{x_1}^2 + a_2^2 \sigma_{x_2}^2 + 2 a_1 a_2 \sigma_{x_1} \sigma_{x_2} \rho_{xy}$$

إذا كانوا x, y ← SI ← إذاً $\boxed{\rho_{xy} = 0}$ فكل الجزء هاد بيشطب

Functions of Random Variables.

Example :-

Consider the joint pdf shown in the table,

Let $Z = X + Y$

[1] Find the probability mass function of Z, $P(Z=z)$.

x \ y	1	2	3	4
1	0.1	0	0.1	0
2	0.3	0	0.1	0.2
3	0	0.2	0	0

شلنا الي عم لأنو احتمالية حدوثهم = صفر فخلي د لا تكتب الاحتمالات صفر

مجموعها
$x+y$

X	Y	Z	$P(Z=z)$
1	1	2	$P(X=1, Y=1) = 0.1$
~~1~~	~~2~~		0
1	3	4	$P(X=1, Y=3) = 0.1$
~~1~~	~~4~~		
2	1	3	$P(X=2, Y=1) = 0.3$
~~2~~	~~2~~		
2	3	5	$P(Z=5) = P(X=2, Y=3) + P(X=3, Y=2) = 0.1 + 0.2 = 0.3$
2	4	6	$P(X=2, Y=4) = 0.2$
~~3~~	~~1~~		
3	2	5	
~~3~~	~~3~~		
~~3~~	~~4~~		

لازم أنتبه إذا في شغلات مكررة

$$\therefore P(Z=z) = \begin{cases} 0.1 , & z=2 \\ 0.3 , & z=3 \\ 0.1 , & z=4 \\ 0.3 , & z=5 \\ 0.2 , & z=6 \\ 0 , & o.w \end{cases}$$

هذا الجدول الكامل الجدول عبارة عن كل زوج مع التاني tupples ٣×٤

2 Find $P(X=Y)=?$

$= P(X=1, Y=1) + P(X=2, Y=2) + P(X=3, Y=3)$

$= 0.1 + 0 + 0$

$= 0.1$

3 $E\{X+Y\}=?$

$E\{X+Y\} = E\{Z\} = \sum_{-\infty}^{\infty} z\,P(Z=z)$

0.2   1.5   0.9   0.4   1.5   1.2

$= 2(0.1) + 3(0.3) + 4(0.1) + 5(0.3) + 6(0.2)$

$= 4.2$

بيجيك سؤال إذا كان طالب $E\{x+y\}$ بدون ما يطلب Z ، فهي داي بالأول
أطلع Z ، إنما تحلّه عادي زي ما نحل دايماً.

The Continuous R.V case

Note: X and Y are two continuous R.Vs. $f_{X,Y}(x,y)$ is the joint Pdf of X and Y. A new R.V $\boxed{Z = X+Y}$ is defined. Determine the Pdf of Z.

In such a question, we can't find the pdf directly.
Thus, we know that :- $f_Z(z) = \dfrac{d F_Z(z)}{dz}$ ← ال Pdf هو تفاضلة ال cdf

$F(z) = P(\underset{\text{Random Variable}}{Z} \le \underset{\text{value}}{z})$

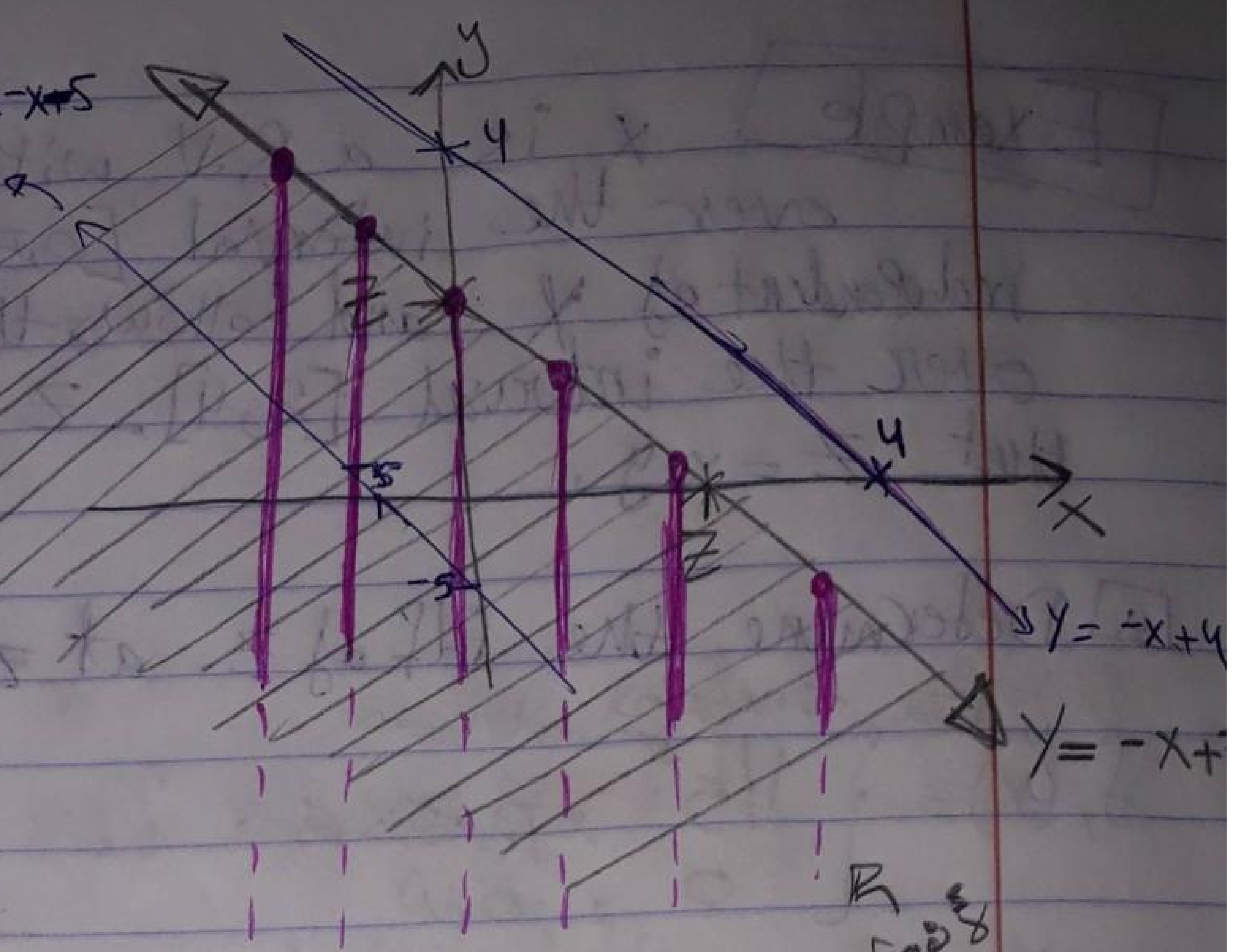
$= P(X+Y \le z)$

$\therefore F(z) = P(Y \le -x+z)$

لتحديد حدود التكامل :-

$\rightsquigarrow y = -x+5$ هي نقاط التقاطع مع x و y تُمثّل z .

محوري



$$= \int_{-\infty}^{\infty} \int_{-\infty}^{-x+z} f_{x,y}(x,y)\,dy\,dx$$

If x & y are S.I

$$F(z) = \int_{-\infty}^{\infty} \left[\int_{-\infty}^{-x+z} f_x(x) f_y(y)\,dy \right] dx$$

$$= \int_{-\infty}^{\infty} f_x(x) \underbrace{\int_{-\infty}^{-x+z} f_y(y)\,dy}_{F_y(y)}\,dx$$

$$= \int_{-\infty}^{\infty} f_x(x) \left[F_Y(y) \Big|_{-\infty}^{-x+z} \right] dx$$

$$= \int_{-\infty}^{\infty} f_x(x) \quad F_Y(-x+z) - \cancel{F_Y(-\infty)}\,dx$$

$$= \int_{-\infty}^{\infty} f_x(x)\, F_y(-x+z)\,dx$$

$$f_Z(z) = \frac{d\,F_Z(z)}{dz} = \int_{-\infty}^{\infty} f_x(x) \frac{d\,F_Y(-x+z)}{dz}\,dx$$

دخلنا الاشتقاق هون مش على كل التكامل عشان الاشتقاق بالنسبة إلى z مش إلى x .

$$= \int_{-\infty}^{\infty} f_x(x)\, f_y(-x+\overset{\text{مشتقة } z = 1}{z})\,dx$$

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x)\, f_Y(z-x)\,dx \; \rightsquigarrow \text{Only If } X \& Y \text{ are SI .}$$

$$= \int_{-\infty}^{\infty} f_Y(y)\, f_X(z-y)\,dy$$

هدول الصيغتين اسمهم : Convolutional integral

لأنو مش محددة فترة فهي بتبدأ من (-∞) و ظللنا المنطقة هاي عشان فيها بتكون الـ $-x+z \geq y$ إذا أخذنا فيها أي زوج مرتب على فرض القيمة x مثلاً 4 كنا نشوف مقارنتي y و $-x+z$ فوقه المنحنى تبع $y=-x+z$ أو تحته ووين بتحقق المعادلة بتبقى $\boxed{Y \leq -x+z}$

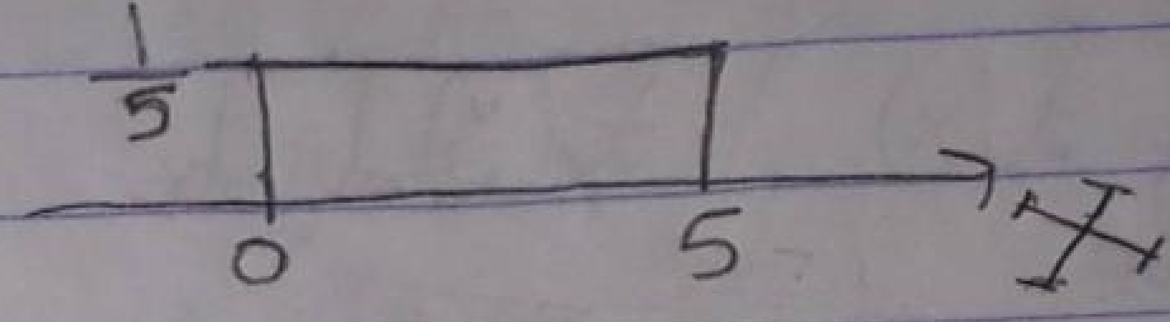
هدول طبعاً خاصين بال $Z = X+Y$ مش كل ال continuous

Example X is a R.V with **uniform distribution** over the interval [0,5], Y is another R.V independent of X and follows the **uniform distribution** over the interval [2,4]. Z is a new R.V such that $Z = X + Y$.

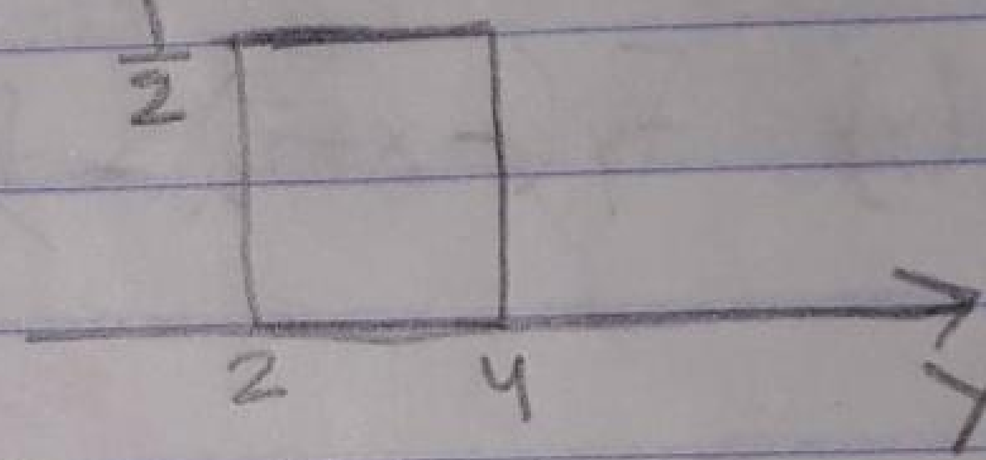
a Determine the Pdf of Z at $Z = 4$.

حسب قانون الـ uniform :- $f_X(x) = \frac{1}{b-a}$

$$f_X(x) = \begin{cases} 1/5 & , 0 \le X \le 5 \\ 0 & , o.w \end{cases}$$



$$f_Y(y) = \begin{cases} 1/(4-2) & , 2 \le y \le 4 \\ 0 & , o.w \end{cases}$$



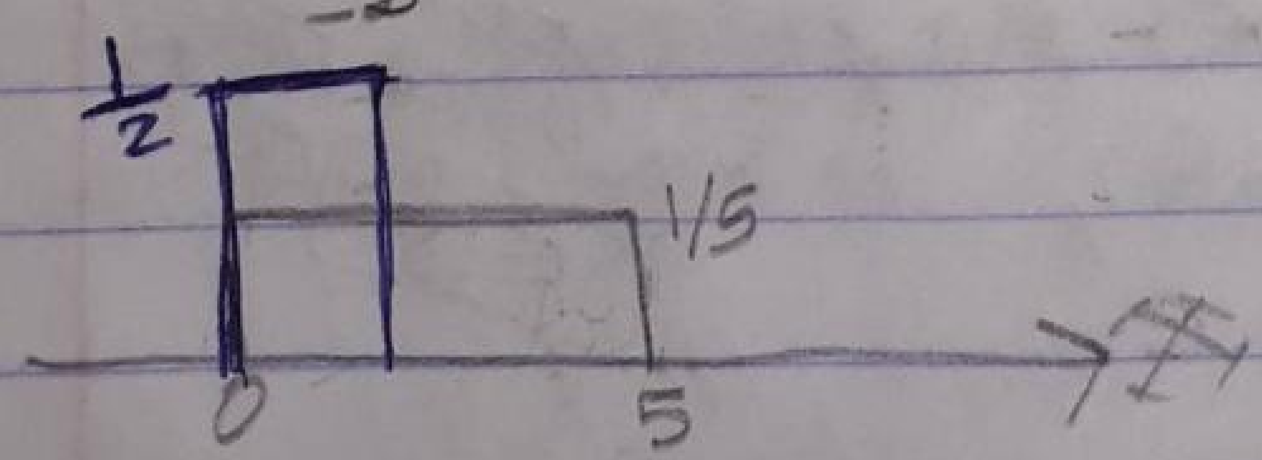
$$= \begin{cases} 1/2 & , 2 \le y \le 4 \\ 0 & , o.w \end{cases}$$

حسب القانون اللي أخذته بالصفحة اللي فاتت :-

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x)\, f_Y(z-x)\, dx$$

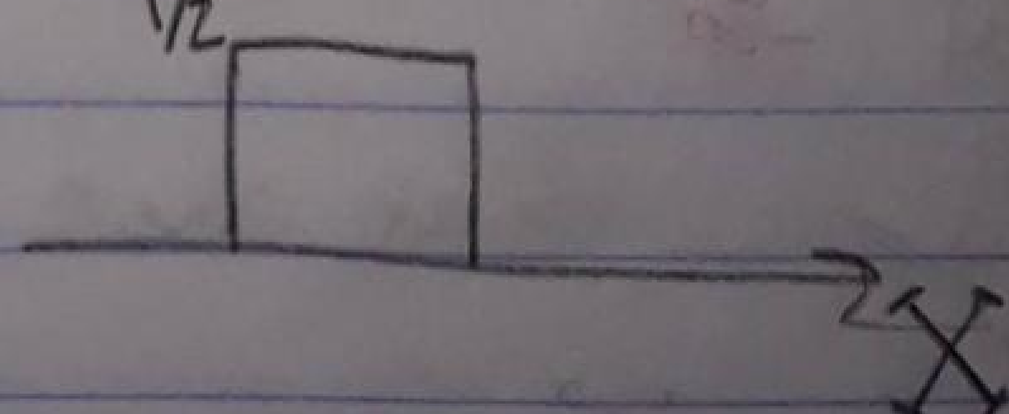
$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x)\, f_Y(4-x)\, dx \xRightarrow{\text{Thus}} f_Y(4-x) = \begin{cases} 1/2 & , 2 \le 4-x \le 4 \\ 0 & , o.w \end{cases}$$

$$= \begin{cases} 1/2 & , -2 \le -x \le 0 \\ 0 \end{cases}$$

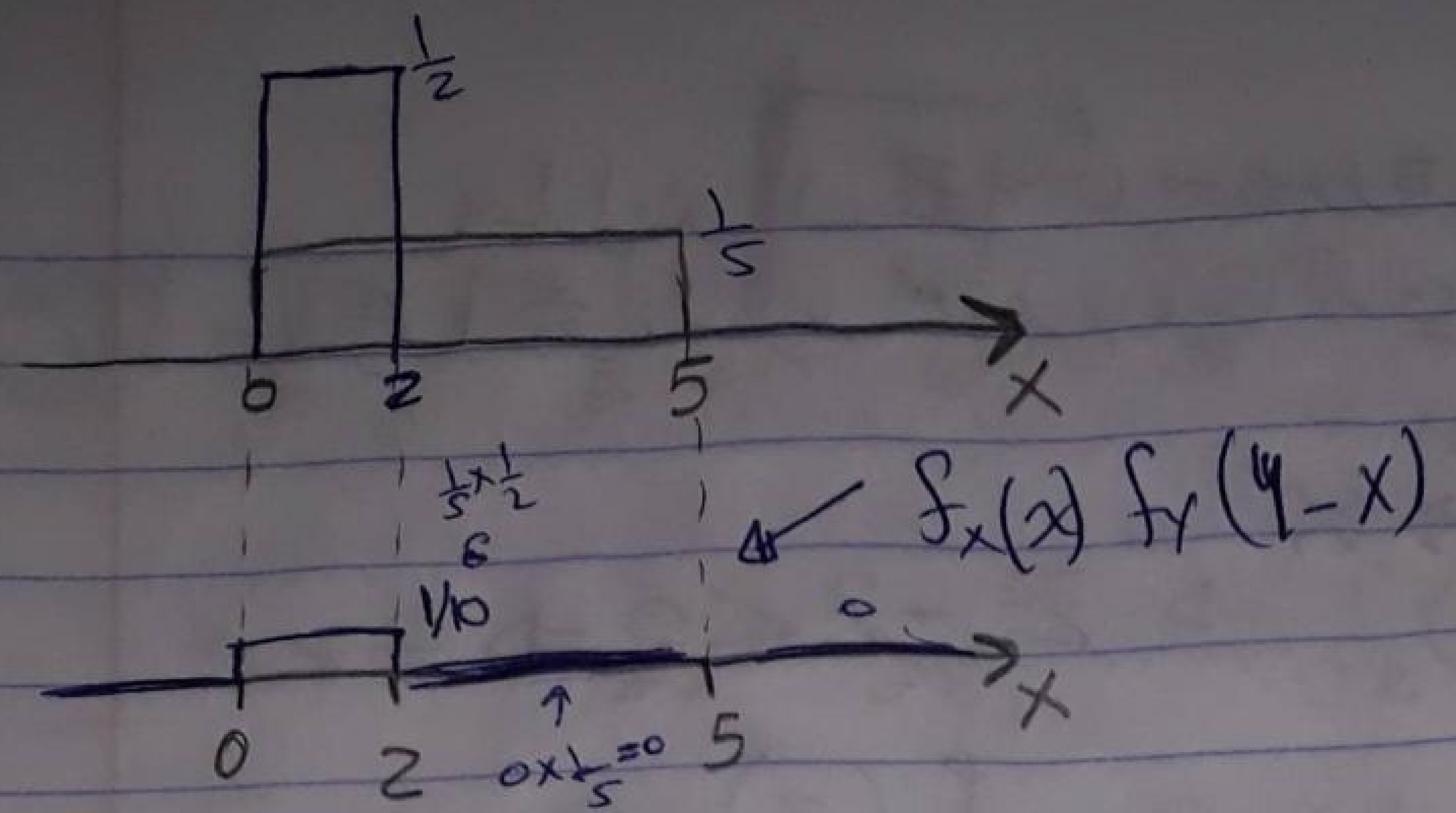


حولنا الـ y لـ x فصرنا نقدر نتعامل معها كـ x ، وحطيناها على نفس خط الأعداد الـ x .

$$\therefore f_Y(4-x) = \begin{cases} 1/2 & , 0 \le x \le 2 \\ 0 & , o.w \end{cases}$$



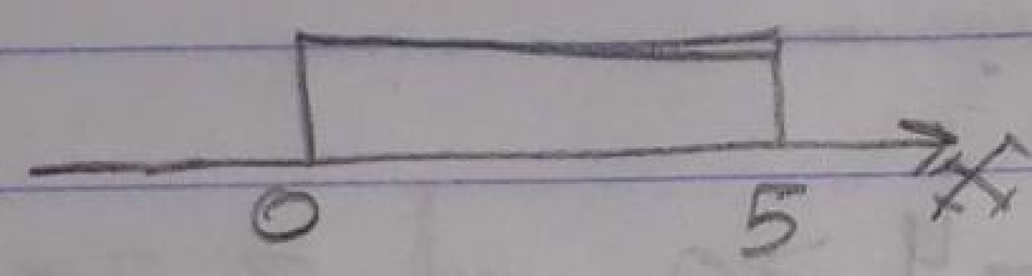
تكملة

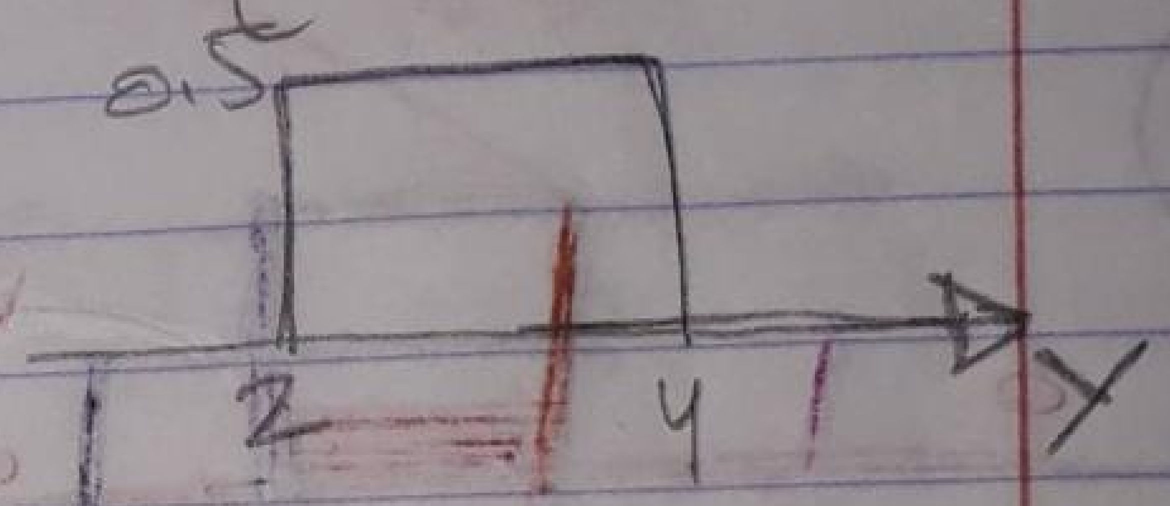


$$\therefore f_z(4) = \int_0^2 \frac{1}{10}\, dx = \frac{2}{10} = \frac{1}{5}$$

$\boxed{b}$ Determine the pdf of z ?

$$f_x(x) = \begin{cases} 1/5 , & 0 \le x \le 5 \\ 0 , & o.w \end{cases} \quad , \quad f_y(y) = \begin{cases} 1/2 , & 2 \le y \le 4 \\ 0 , & o.w \end{cases}$$



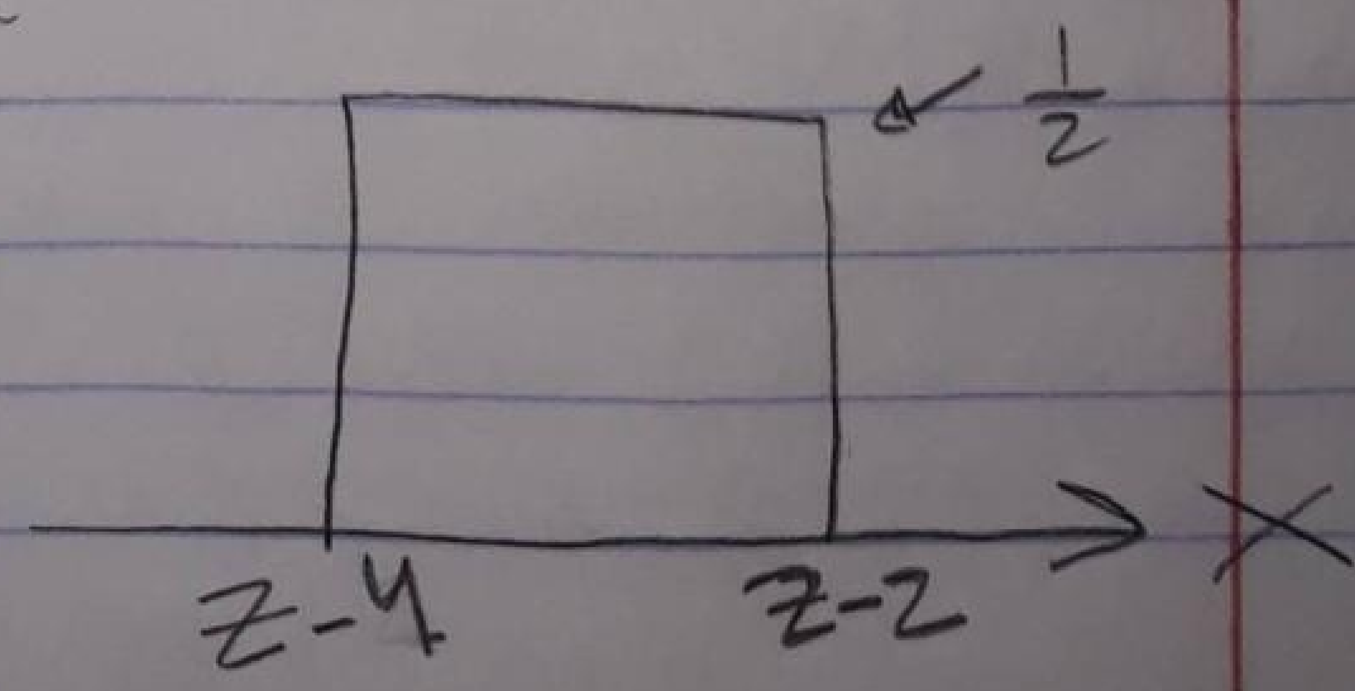


$$f_z(z) = \int_{-\infty}^{\infty} f_x(x)\, f_y(y)\, dx = \int_{-\infty}^{\infty} f_x(x)\, f_y(z-x)\, dx$$

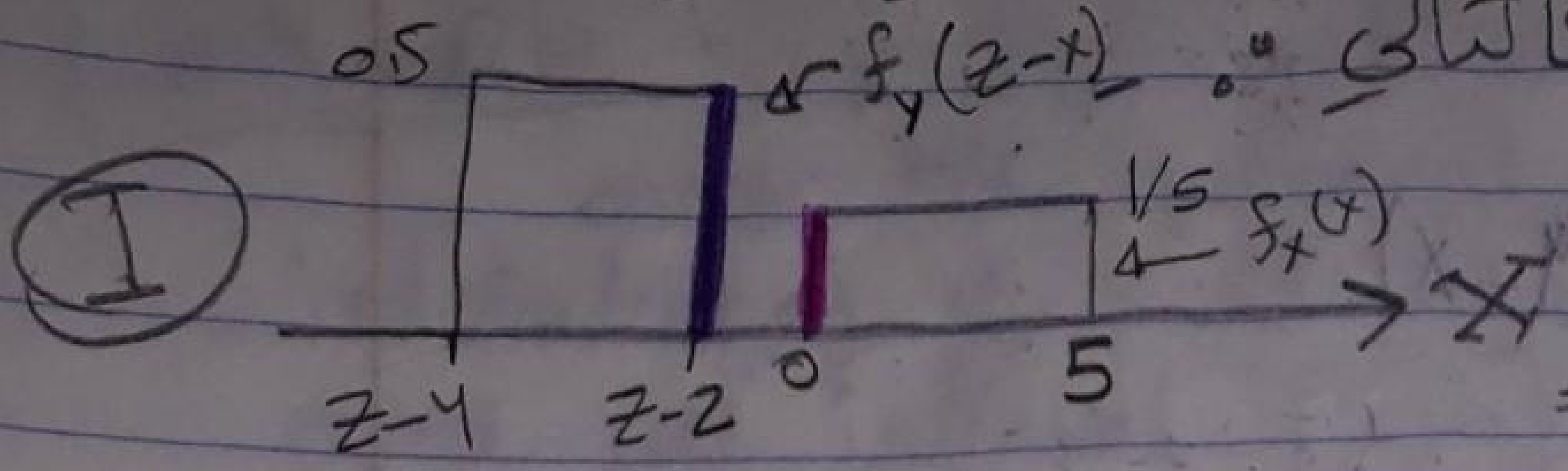
but $f_y(z-x) = \begin{cases} 1/2 , & 2 \le z-x \le 4 \\ 0 , & o.w \end{cases}$

$$= \begin{cases} 1/2 , & -z+2 \le -x \le -z+4 \\ 0 , & o.w \end{cases}$$

$$f_y(z-x) = \begin{cases} 1/2 , & z-4 \le x \le z-2 \\ 0 \end{cases}$$

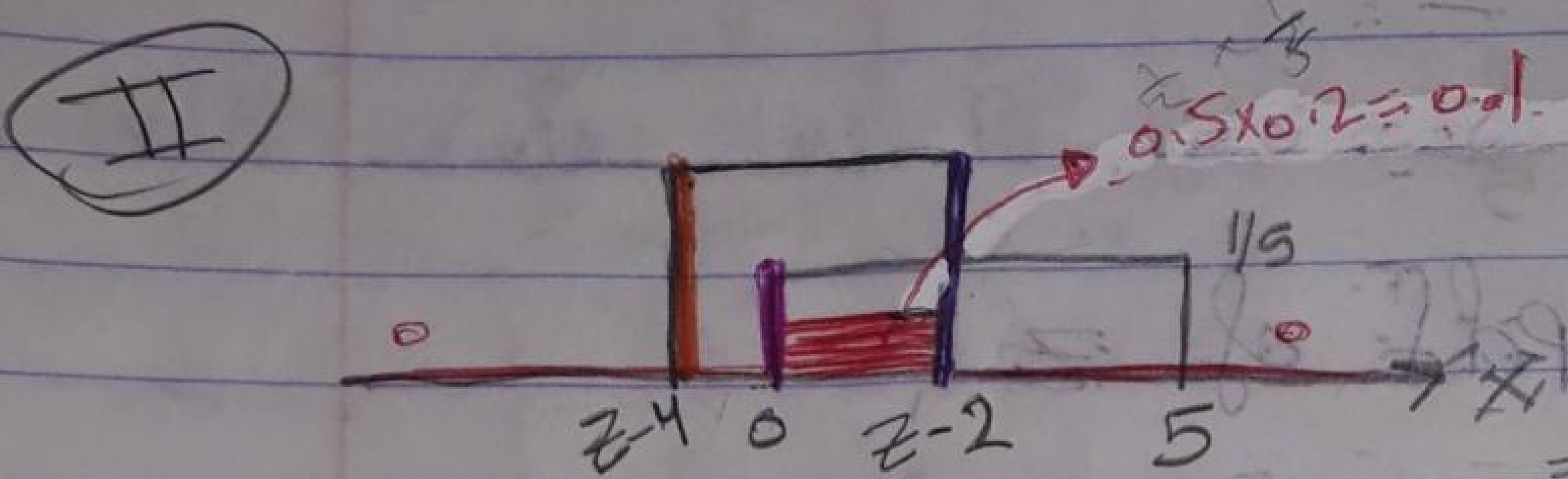


بما إنو الـ Z مش محصورة بفترة معينة إذاً في أكثر من Case ، كالتالي :



$\Rightarrow Z - 2 < 0 \Rightarrow Z < 2$

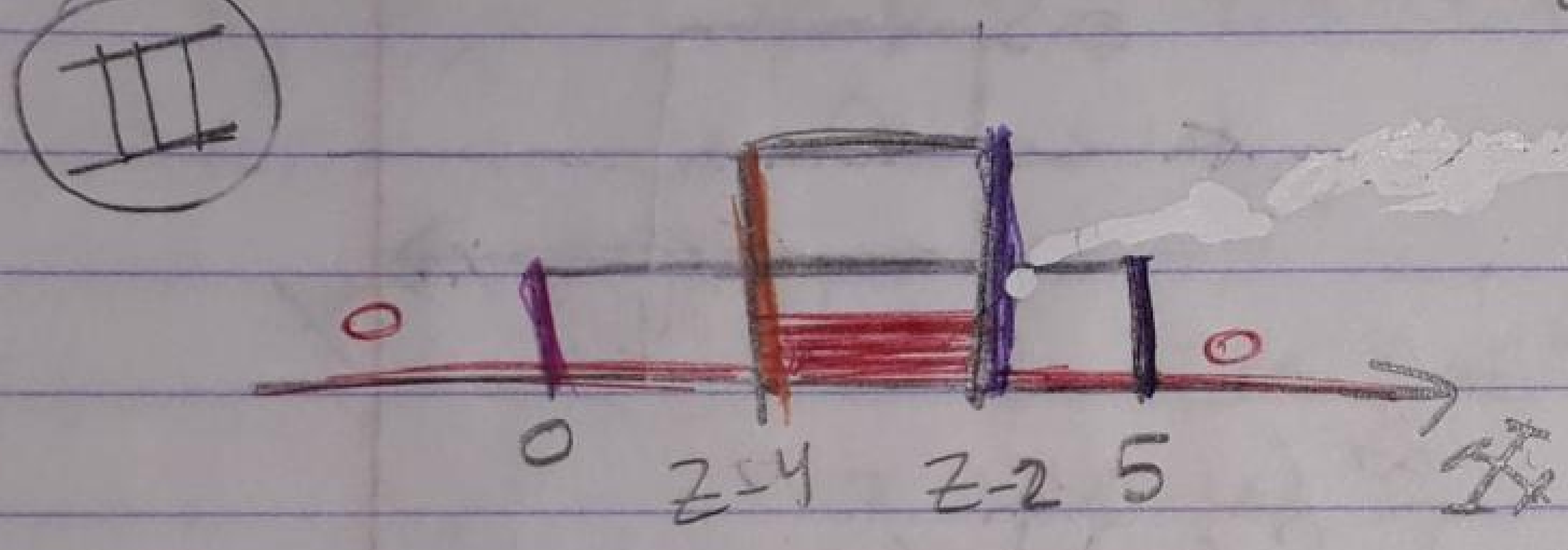
$f_Z(z) = 0$



$\Rightarrow Z-2 > 0$ and $Z-4 < 0$

$Z > 2$ and $Z < 4$

$\therefore f_Z(z) = \int_0^{z-2} \frac{1}{10}\, dx = \frac{1}{10}(z-2)$

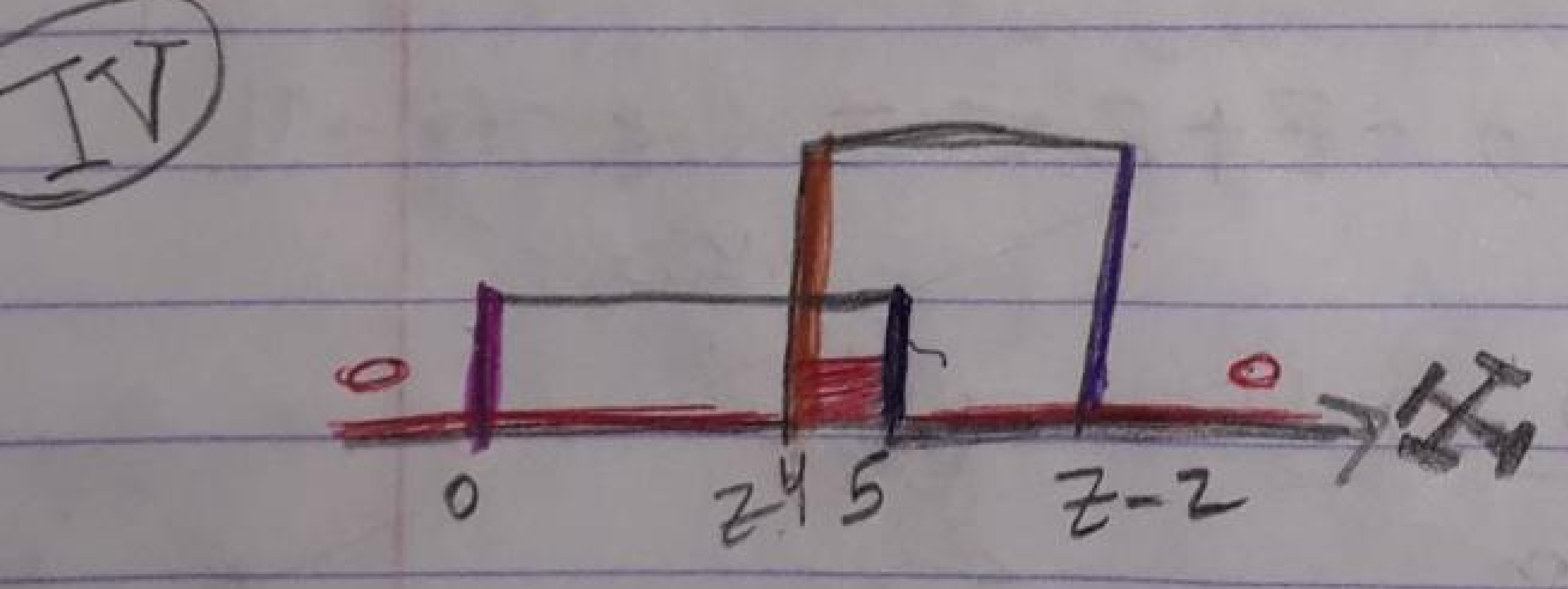


$\Rightarrow Z-4 > 0$ and $Z-2 < 5$

$Z > 4$ and $Z < 7$

$f_Z(z) = \int_{z-4}^{z-2} \frac{1}{10}\, dx = \frac{1}{10} \times \Big|_{z-4}^{z-2}$

$= \frac{1}{10}[\cancel{z} - 2 - \cancel{z} + 4] = \frac{1}{10}[2] = \frac{1}{5}$



$Z-2 > 5$ and $Z-4 < 5$

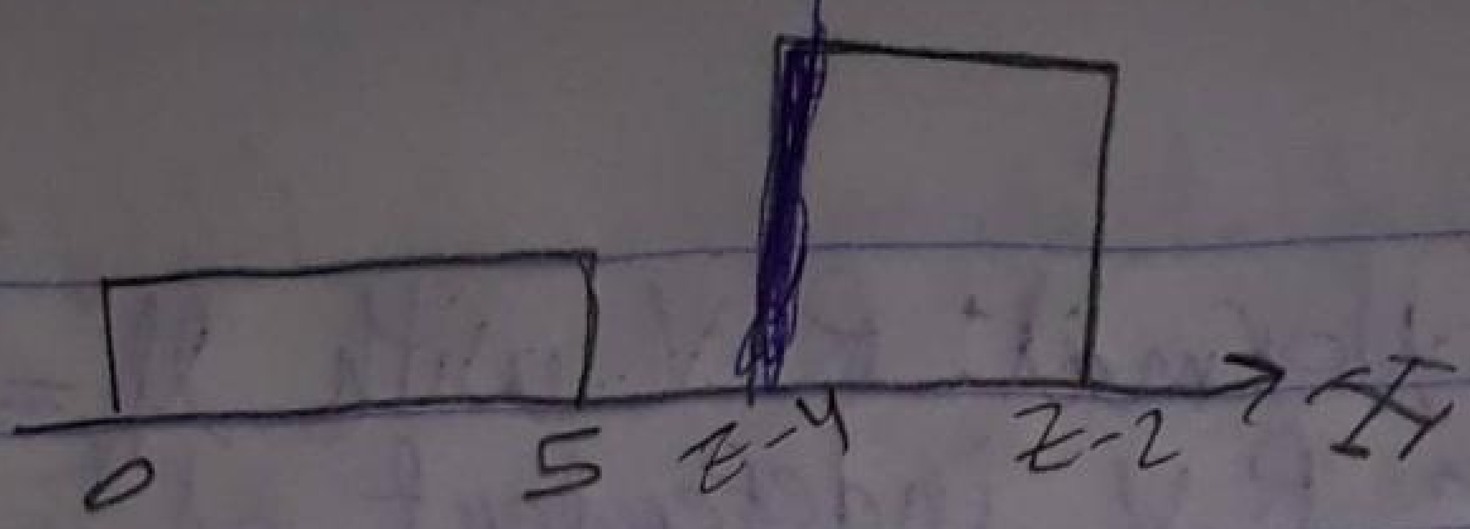
$Z > 7$ and $Z < 9$

$\therefore f_Z(z) = \int_{z-4}^{5} \frac{1}{10}\, dx = \frac{1}{10}[5 - z + 4]$

$= \frac{1}{10}[9 - z]$.

Scanned by TapScanner

(V)



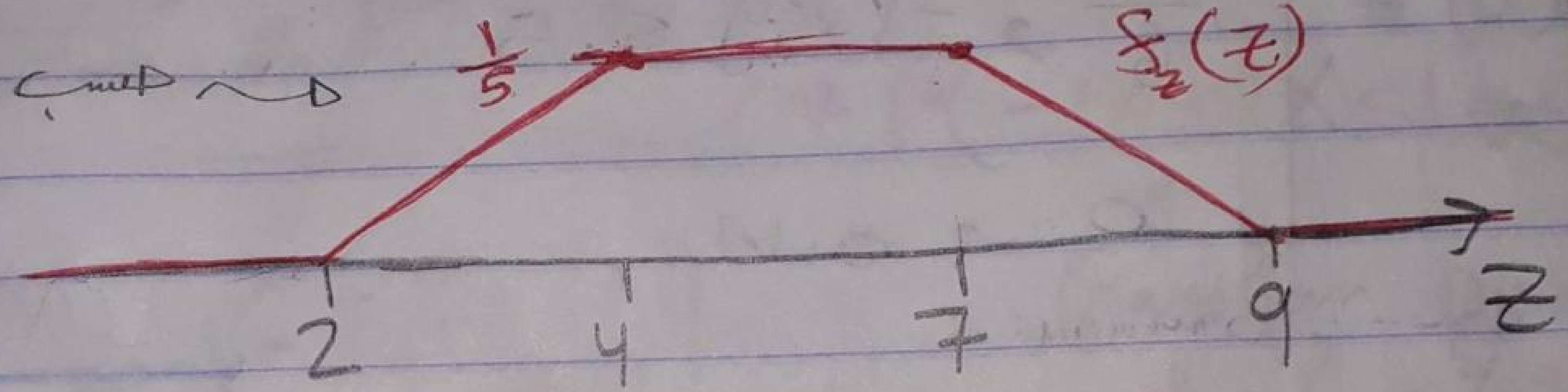
$$Z-4>5$$

$$Z>9$$

$$f_Z(z)=0$$

$$\therefore f_Z(z)=\begin{cases} 0 & , z<2 \\ (1/10)(z-2) & , 2\le z<4 \\ 1/5 & , 4\le z<7 \\ (1/10)(9-z) & 7\le z<9 \\ 0 & 9\le z \end{cases}$$

حسب ال pdf



طريقة سهلة عشان أتأكد من الفترات اللي أنا عاملاهم هنا :

الفترة بتاعة ال $f_x(x)$

$f_x(x) \longrightarrow$ 0 5

$f_y(y) \longrightarrow$ 2 4

2 7 4 9

نغير ترتيبها

$\longleftarrow$ 2 $\longleftrightarrow$ 4 $\longleftrightarrow$ 7 $\longleftrightarrow$ 9 $\longrightarrow$

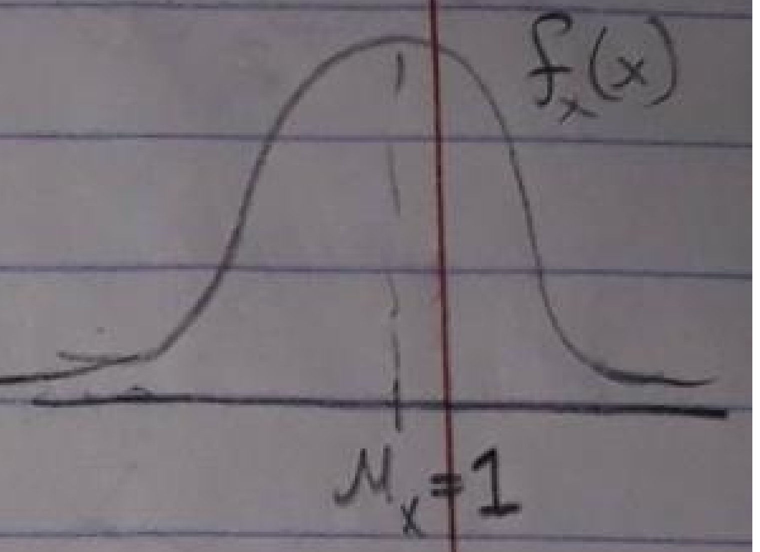
(I) (II) (III) (IV) (V)

$\boxed{\text{Example}}$ let x be a Normal R.V with $\mu_x = 1$, and $\sigma_x^2 = 9$. y is another R.V independent of x with a uniform distribution over the internal $[-1, 5]$. $Z = x + y$. Determine the pdf of Z at $\boxed{Z=0}$.

$$f_Z(Z) = \int_{-\infty}^{\infty} f_x(x)\, f_y(Z - x)\, dx$$

$$f_x(x) = \frac{1}{\sqrt{2\pi * 9}}\, e^{\frac{-(x-1)^2}{2*9}} \quad , \quad -\infty < x < \infty$$

لأنو Normal ← Gaussian



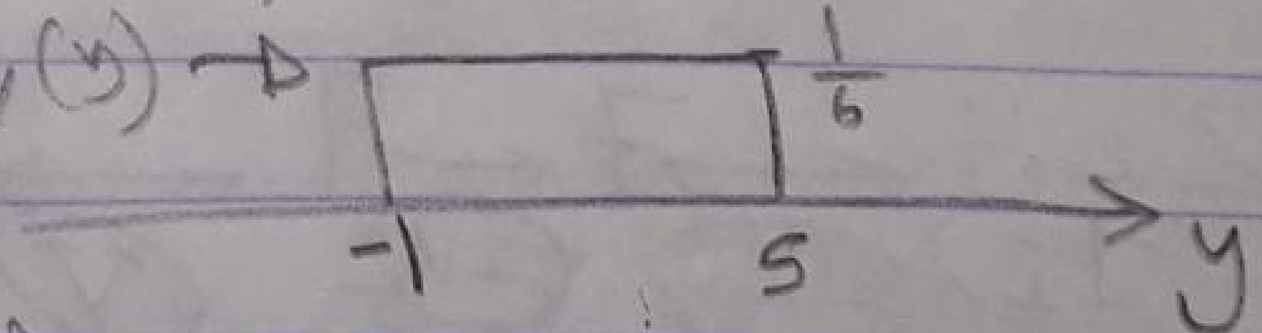
$$f_y(y) = \begin{cases} \frac{1}{6} & , -1 \le y \le 5 \\ 0 & , \text{o.w} \end{cases}$$

$$\therefore\ f_y(Z - x) = \begin{cases} \frac{1}{6} & , -1 \le Z - x \le 5 \\ 0 & , \text{o.w} \end{cases}$$

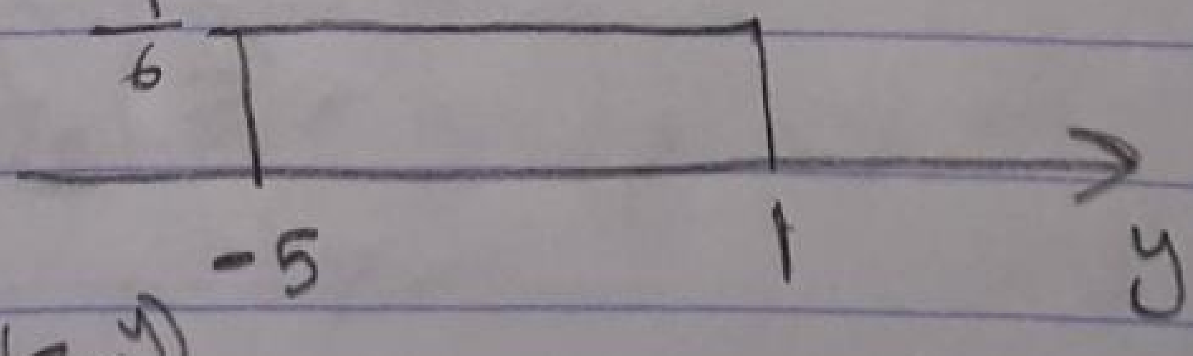
$$= \begin{cases} \frac{1}{6} & , Z-5 \le x \le Z+1 \\ 0 & , \text{o.w} \end{cases}$$

طريقة سهلة لرسم $f_y(Z-x)$ ، بدل ما أحول y بدلالة Z و x أول بحرف أرسم:

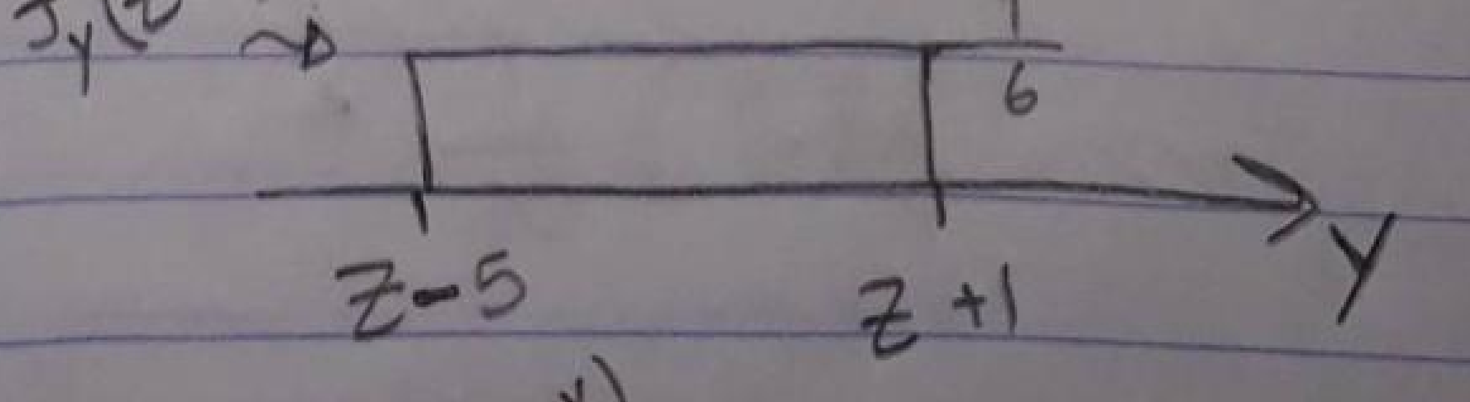
$\boxed{1}$ بورسم بدلالة y عادي نفس هاد



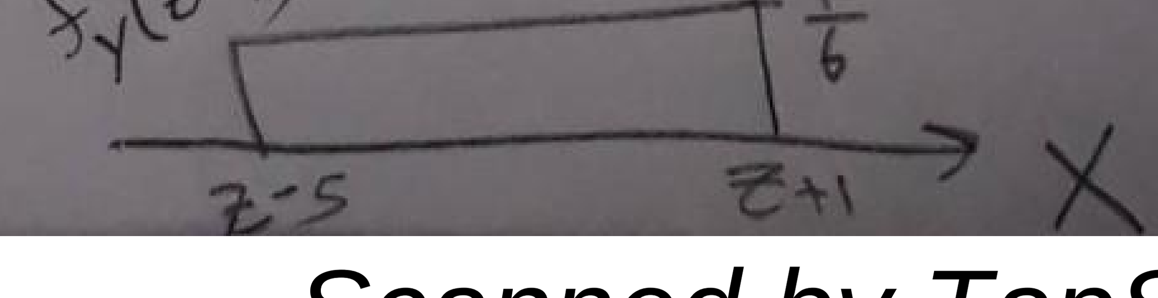
$\boxed{2}$ بورسم $f_y(-y)$ ، يعني بضرب حدود y بـ سالب ، بالتالي رح أعكس أماكنهم لأنو الكبير بصير ، والصغير بيكبر بعد السالب.



$\boxed{3}$ بضيف Z للحدود ولداخل الـ $f_y(-y)$ فتصبح $f_y(Z-y)$

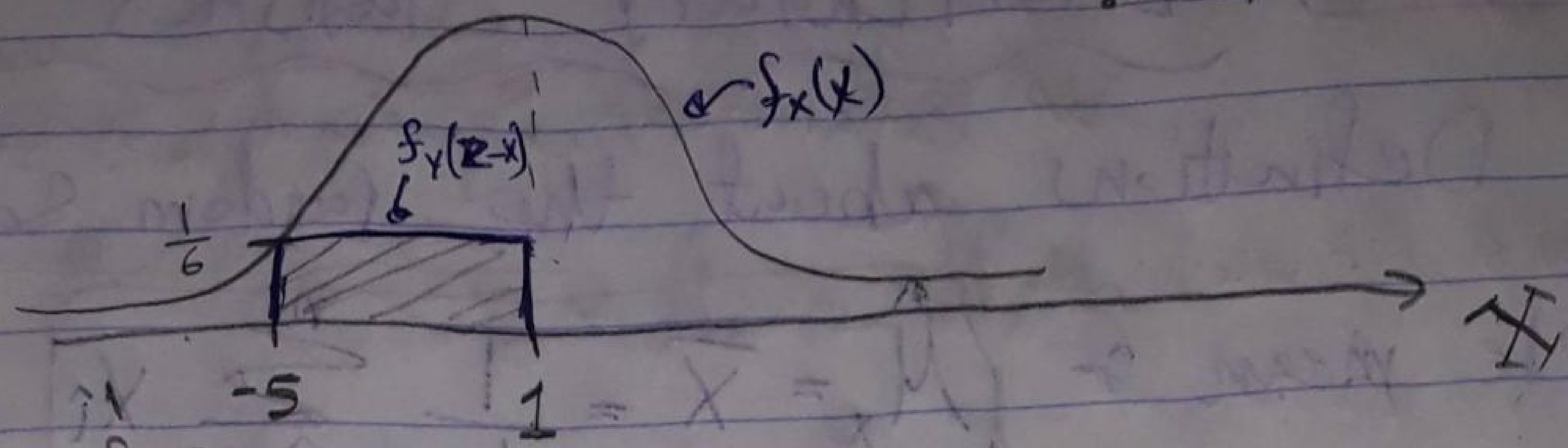


$\boxed{4}$ بدل الرمز y ، بحط الرمز X



وهيك بكون حصلت على $f_Y(z-x)$ بطريقة سريعة، وبقدر أقارنها مع $f_X(x)$ على خط X .

at $\boxed{z=0}$:-



$$\therefore f_Z(z=0) = \int_{-5}^{1} f_X(x)\, f_Y(z-x)\, dx$$

$$= \int_{-5}^{1} \left(\frac{1}{6}\right)\left(\frac{1}{\sqrt{2\pi * 9}}\right) e^{-\frac{(x-1)^2}{2*9}}\, dx$$

$$= \frac{1}{6} \int_{-5}^{1} \frac{1}{\sqrt{2\pi * 9}}\, e^{-\frac{(x-1)^2}{2*9}}\, dx \rightarrow P(-5 \le X \le 1)$$

Gaussian Distribution $\mu_x = 1,\ \sigma_x^2 = 9$

$$\therefore = \frac{1}{6}\left[\phi\left(\frac{1-1}{\sqrt{9}}\right) - \phi\left(\frac{-5-1}{\sqrt{9}}\right)\right]$$

$$= \frac{1}{6}\left[\phi(0) - \phi(-2)\right]$$

$$= \frac{1}{6}\left[\phi(0) - [1 - \phi(2)]\right]$$

$$= \frac{1}{6}\left[0.5 - 1 + \phi(2)\right]$$

بطلع هاي من الجدول .

Scanned by TapScanner

Chapter #4 Elementary Statistics

Basic Definitions about the random Sample :-

هو عينة أجمعها n مأخوذة من population كبيرة أصلية

ما نميز الـ Sample عن الأصلية هو الـ $(\hat{\ })$ hat.

[1] Sample mean :-

$$\hat{\mu}_x = \bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$$

[2] Sample Variance :- S_x or $\hat{\sigma}_x^2$

$$S_x = \hat{\sigma}_x^2 = \frac{1}{n} \sum_{i=1}^{n} (X_i - \mu_x)^2$$

μ_x is known (true mean is known).

OR

$$S_x = \hat{\sigma}_x^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \hat{\mu}_x)^2$$

True Mean is Unknown so we use $\hat{\mu}_x$

Population: $\mu_x, \sigma_x^2, \sigma_x, b, \lambda$ (exp-nor, binomial distribution)

Random Sample: $X_1, X_2, \ldots, X_n$

size of Sample is (n)

[3] Sample Standard deviation $S_x, \hat{\sigma}_x$

$$S_x = \hat{\sigma}_x = \sqrt{S_x^2} = \sqrt{\hat{\sigma}_x^2}$$

زي ما بنعرف

-: Some Relation بين 2 Random Variables إذا كان

Sample Covariance $= \hat{\mu}_{x,y} = C_{x,y}$:-

$$C_{x,y} = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \hat{\mu}_x)(Y_i - \hat{\mu}_y)$$

Sample Correlation Coefficient $= r_{xy}$

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

برضو زي ما أخذنا تعاونا الـ correlation coefficient ρ_{xy} بتشابه.

بس هون اختلفت الرموز.

من معادلة الـ variance $(\hat{S}_x^2)$ بنا نشتق معادلة جديدة (التالية):

$$S_x^2 = \hat{\sigma}_x^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \hat{\mu}_x)^2$$

$$= \frac{1}{n-1}\sum_{i=1}^{n}\left[x_i^2 - 2x_i\hat{\mu}_x + \hat{\mu}_x^2\right] \leftarrow \text{فكينا التربيع}$$

$$= \frac{1}{n-1}\left[\sum_{i=1}^{n}x_i^2 - 2\underbrace{\hat{\mu}_x}_{\text{ثابت}}\sum_{i=1}^{n}x_i + \underbrace{\hat{\mu}_x^2}_{\text{ثابت}}\sum_{i=1}^{n}1\right]$$

$(1\times n)$

$$= \frac{1}{n-1}\left[\sum_{i=1}^{n}x_i^2 - 2\hat{\mu}_x\, n\hat{\mu}_x + \hat{\mu}_x^2\, n\right]$$

$$= \frac{1}{n-1}\left[\sum_{i=1}^{n}x_i^2 - \cancel{2}n\hat{\mu}_x^2 + \cancel{n\hat{\mu}_x^2}\right]$$

$$= \frac{1}{n-1}\left[\sum_{i=1}^{n}x_i^2 - n\hat{\mu}_x^2\right]$$

$$= \frac{1}{n-1}\left[\sum_{i=1}^{n}x_i^2 - n\left(\frac{1}{n^2}\right)\left(\sum_{i=1}^{n}x_i\right)^2\right]$$

نضرب $\sum_{i}^{n}x_i^2$ بـ $\frac{n}{n}$
ثم نخرج $\frac{1}{n}$ عامل مشترك من المعادلة

$$\therefore\ S_x^2 = \frac{1}{n(n-1)}\left[n\sum_{i=1}^{n}x_i^2 - \left(\sum_{i=1}^{n}x_i\right)^2\right]$$

بستخدمها إذا معيش لا μ_x ولا $\hat{\mu}_x$

أمّا هلأ فمن معادلة الـ covariance نبدأ نشتق معادلة جديدة:

$$C_{xy} = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \hat{\mu}_x)(y_i - \hat{\mu}_y)$$

$$= \frac{1}{n-1}\sum_{i=1}^{n}\left[x_iy_i - x_i\hat{\mu}_y - y_i\hat{\mu}_x + \hat{\mu}_x\hat{\mu}_y\right]$$

$$= \frac{1}{n-1}\left[\sum_{i=1}^{n}x_iy_i - \hat{\mu}_y\sum_{i=1}^{n}x_i - \hat{\mu}_x\sum_{i=1}^{n}y_i + \hat{\mu}_x\hat{\mu}_y\sum_{i=1}^{n}1\right]$$

$$= \frac{1}{n-1}\left[\sum_{i=1}^{n}x_iy_i - \hat{\mu}_y\, n\hat{\mu}_x - \cancel{n\hat{\mu}_y\hat{\mu}_x} + \cancel{n\hat{\mu}_x\hat{\mu}_y}\right]$$

$$= \frac{1}{n-1}\left[\sum_{i=1}^{n} x_i y_i - \cancel{n}\left(\frac{1}{\cancel{n}}\sum^{n} x_i\right)\left(\frac{1}{n}\sum_{i=1}^{n} y_i\right)\right]$$

$$\therefore\ C_{xy} = \frac{1}{n(n-1)}\left[n\sum_{i=1}^{n} x_i y_i - \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right)\right]$$

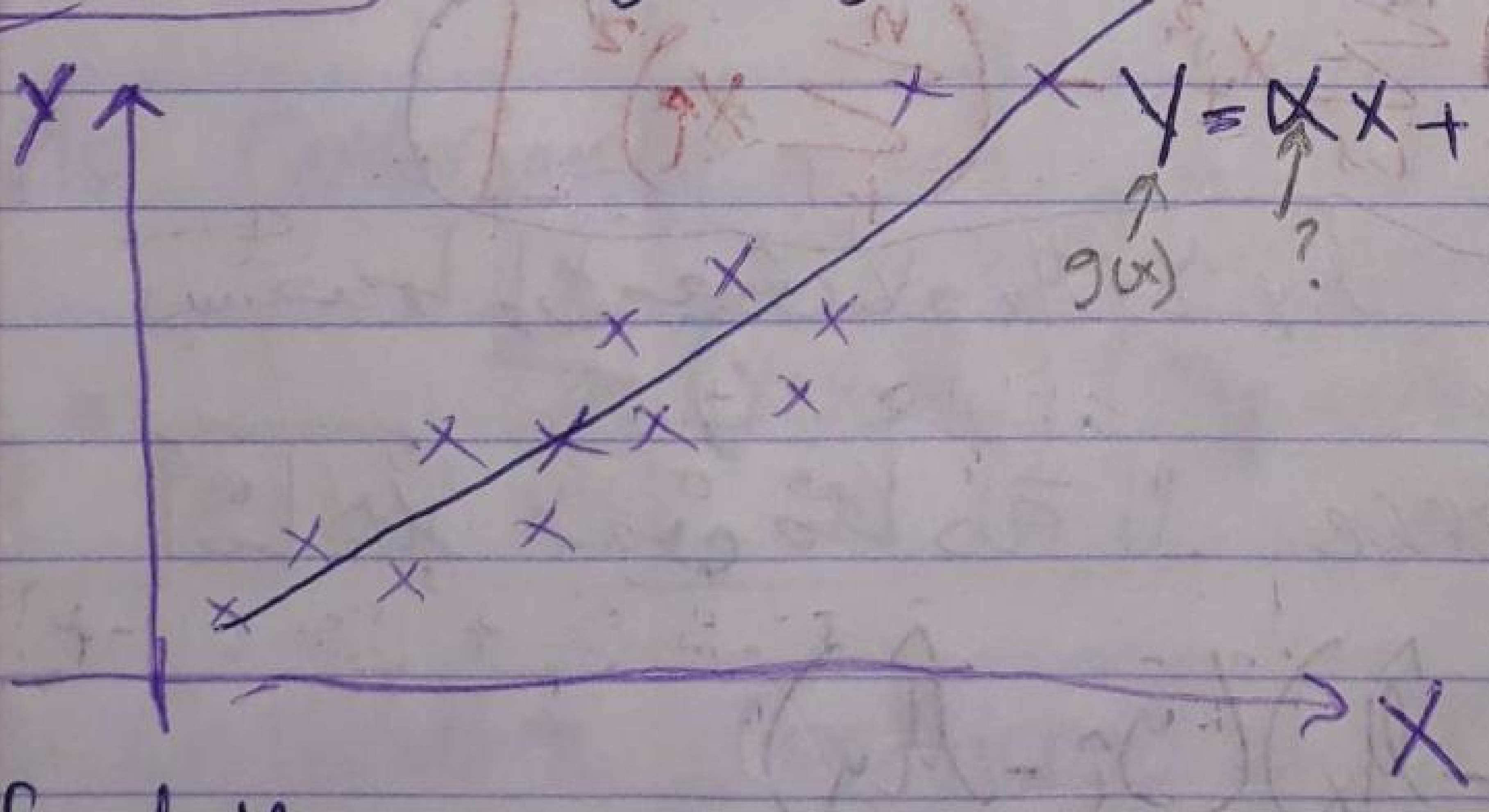
نلاحظ إنو إذا عوّضت مكان y بـ x ، فبطلع نفس قانون الـ S_x^2
اللي اشتقيناه هلا .

Regression Techniques :- إيجاد علاقة بين two Random Variables

X_i	x_1	x_2	x_3				X_n
y_i	y_1	y_2	y_3				y_n ← Practical measurment
$Y = g(x)$	$g(x_1)$	$g(x_2)$	$g(x_3)$				$g(X_n)$ ← theoritical

Random Sample
$\langle x_1, y_1\rangle, \langle x_2, y_2\rangle, \ldots, \langle x_n, y_n\rangle$

So draw the plot out of these points, we try to find the best $\boxed{\text{fit line}}$ passing through the points, and $\boxed{\text{Equation}}$.



$$\underset{g(x)}{Y} = \underset{?}{\alpha} X + \underset{?}{\beta}$$

to find the equation that describes the relationship between X and Y; a $\boxed{g(x) \text{ function } = Y}$ needs to be defined here.

وبدنا نحاول إنو تكون القيمة العملية أقرب ما يكون للنظرية، وبالطبع رح يكون في نسبة خطأ ناتجة سواء من ال measurments أو أي شيء آخر، و إحنا منهدف لتكون أصغر ما يمكن، فبنعرف ال Error Function :
هو الفرق بين النظري والعملي وبكدمش عند نقطة واحدة، إنما عند كل النقاط، عشان هيك حطينا $\sum$ ، وبعدها بناخد ال average وذلك بالقسمة على n ، ولتلاشي السوالب حطينا تربيع، كالتالي:

$$E = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - g(x_i)\right)^2$$ ← Mean Square error

From Equation $g(x_i) = \alpha x_i + \beta$:-

$$E = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \underset{?}{\alpha} x_i - \underset{?}{\beta}\right)^2$$

To find $\underbrace{\alpha \text{ and } \beta}$, we derive the equation twice:

لأنهم عشان يطلع الأكبر قليل ال minimum values

1 In respect to β :- $\frac{dE}{d\beta} = 0 \Rightarrow \frac{\cancel{1}}{\cancel{n}}\sum_{i=1}^{n}\cancel{2}\left(y_i - \alpha x_i - \beta\right)\underset{\text{معامل } \beta}{(-1)} = 0$

$$\therefore -\sum y_i + \alpha\sum_{i=1}^{n} x_i + \beta n = 0$$

$$\therefore \boxed{\beta n + \alpha\sum_{i=1}^{n} x_i = \sum y_i} \quad \text{--- (1)}$$

بجرب جدول القيم فبس جمع جدول القيم

نفس الإشي هون، فضفية بس بوفش α و β .

رامواشان الصفر بايسير عنصبودهم وعدمه واحد.

2 In respect to α :- $\frac{dE}{d\alpha} = 0 \Rightarrow \frac{\cancel{1}}{\cancel{n}}\sum_{i=1}^{n}\cancel{2}\left(y_i - \alpha x_i - \beta\right)\underset{\text{معامل } \alpha}{(-x_i)} = 0$

$$-\sum_{i=1}^{n} y_i x_i + \alpha\sum_{i=1}^{n} x_i^2 + \beta\sum_{i=1}^{n} x_i = 0$$

$$\therefore \boxed{\beta\sum_{i=1}^{n} x_i + \alpha\sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i y_i} \quad \text{--- (2)}$$

بنربط أحل هدول المعادلتين باستخدام ال linear Algebra ← المصفوفات.

تكملة →

المعادلة الأصلية :

$$Y = \alpha x^1 + \beta x^0$$

① :-

$$\beta n + \alpha \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i$$

② :-

$$\beta \sum_{i=1}^{n} x_i + \alpha \sum x_i^2 = \sum_{i=1}^{n} x_i y_i$$

الأشي اللي مضروبة بـ β بالمعادلة الأولى

اللي فوق معاملة أقل خطه أول

الأشي اللي مضروبة بـ α بالمعادلة ①

الأشي اللي مضروبة بـ β بمعادلة ②

الأشي اللي مضروبة بـ α بمعادلة ②

ناتج معادلة ①

ناتج معادلة ②

$$\begin{bmatrix} n & \sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} x_i & \sum_{i=1}^{n} x_i^2 \end{bmatrix} \begin{bmatrix} \beta \\ \alpha \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} x_i y_i \end{bmatrix}$$

بدلنا موقع β بالناتج

$$\beta = \frac{\begin{vmatrix} \sum y_i & \sum x_i \\ \sum x_i y_i & \sum x_i^2 \end{vmatrix}}{\begin{vmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^2 \end{vmatrix}} = \frac{\sum y_i \sum x_i^2 - \sum x_i \sum x_i y_i}{n \sum x_i^2 - (\sum x_i)^2}$$

تكملة

$$\alpha = \frac{\begin{vmatrix} n & \sum y_i \\ \sum x_i & \sum x_i y_i \end{vmatrix}}{\begin{vmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^2 \end{vmatrix}} = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{n \sum x_i^2 - (\sum x_i)^2}$$

نلاحظ أنه: أرجع لمعادلة C_{xy} و S_x^2

$$= \frac{\cancel{n(n-1)}\, C_{xy}}{\cancel{n(n-1)}\, S_x^2}$$

$$\therefore \boxed{\alpha = \frac{C_{xy}}{S_x^2}}$$

بوزبط أعطي كمان

$$Y = \alpha X + \beta$$

$$\hat{\mu}_y = \alpha \hat{\mu}_x + \beta$$

$$\therefore \boxed{\beta = \hat{\mu}_y - \alpha \hat{\mu}_x}$$

So we can rewrite the equation :- $y = \alpha x + \beta$ as follows :-

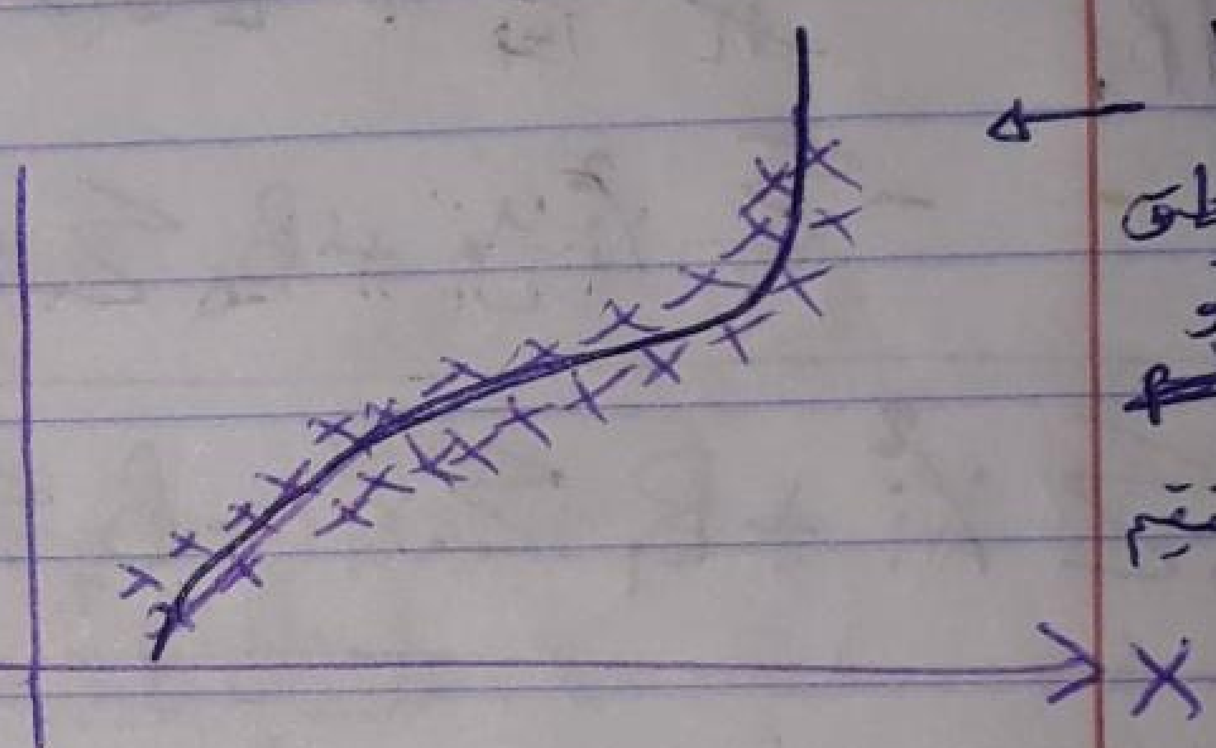
$$y = \frac{C_{xy}}{S_x^2} x + \hat{\mu}_y - \frac{C_{xy}}{S_x^2} \hat{\mu}_x$$

$$y - \hat{\mu}_y = \frac{C_{xy}}{S_x^2} (x - \hat{\mu}_x) \xrightarrow[\text{المعادلة على } S_y]{\text{نقسم طرفي}} \frac{y - \hat{\mu}_y}{S_y} = \frac{C_{xy}}{S_x S_y} \left[\frac{x - \hat{\mu}_x}{S_x} \right]$$

$$\therefore \boxed{\frac{y - \hat{\mu}_y}{S_y} = r_{xy} \left(\frac{x - \hat{\mu}_x}{S_x} \right)}$$

Polynomial Regression :-

$$y = \underline{\underline{B_0}} x^0 + \underline{\underline{B_1}} x^1 + \underline{\underline{B_2}} x^2$$



إذا في هنا زي هيك نقاط فالمنطق بيحكي إنو لو نرسم polynomial مش خط مستقيم

x_i	x_1	x_2	------	x_n
y_i	y_1	y_2		y_n
$g(x)$	$g(x_1)$	$g(x_2)$		$g(x_n)$

$$E = \frac{1}{n} \sum_{i=1}^{n} (y_i - g(x))^2$$

$$E = \frac{1}{n} \sum_{i=1}^{n} \left(y_i - [B_0 + B_1 x_i + B_2 x_i^2] \right)^2$$

$$\boxed{E = \frac{1}{n} \sum_{i=1}^{n} \left[y_i - B_0 - B_1 x_i - B_2 x_i^2 \right]^2}$$

To find B_0, B_1 & B_2 $\longrightarrow$ DERIVE !

$$\frac{dE}{dB_0} = \frac{1}{\cancel{n}} \sum_{i=1}^{n} \cancel{2} [y_i - B_0 - B_1 x_i - B_2 x_i^2) (-1) = 0$$

$$\Rightarrow -\sum_{i=1}^{n} y_i + n B_0 + B_1 \sum_{i=1}^{n} x_i + B_2 \sum_{i=1}^{n} x_i^2 = 0$$

$$\boxed{nB_0 + B_1 \sum x_i + B_2 \sum x_i^2 = \sum y_i} \quad \text{— (1)}$$

Derive in Respect to B_1 :-

$$\frac{dE}{dB_1} = \cancel{\tfrac{n}{2}} \times \frac{1}{\cancel{n}} \sum_{i=1}^{n} \cancel{2} [y_i - B_0 - B_1 x_i - B_2 x_i^2](-x_i) = 0 \times \frac{n}{2}$$

$$\therefore \; -\sum x_i y_i + B_0 \sum x_i + B_1 \sum x_i^2 + B_2 \sum x_i^3 = 0$$

$$\Rightarrow \boxed{B_0 \sum x_i + B_1 \sum x_i^2 + B_2 \sum x_i^3 = \sum x_i y_i} \quad \text{— (2)}$$

Derive in respect to B_2 :-

$$\frac{dE}{dB_2} = \frac{1}{\cancel{n}} \sum_{i=1}^{n} 2 [y_i - B_0 - B_1 x_i - B_2 x_i^2](-x_i^2) = 0$$

$$-\sum x_i^2 y_i + B_0 \sum x_i^2 + B_1 \sum x_i^3 + B_2 \sum x_i^4 = 0$$

$$\boxed{B_0 \sum x_i^2 + B_1 \sum x_i^3 + B_2 \sum x_i^4 = \sum x_i^2 y_i} \quad \text{— (3)}$$

Using linear algebra :-

$$\begin{bmatrix} n & \sum x_i & \sum x_i^2 \\ \sum x_i & \sum x_i^2 & \sum x_i^3 \\ \sum x_i^2 & \sum x_i^3 & \sum x_i^4 \end{bmatrix} \begin{bmatrix} B_0 \\ B_1 \\ B_2 \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum x_i y_i \\ \sum x_i^2 y_i \end{bmatrix}$$

ونحل الحل باستخدام المحددات زي ما حالة α و β بالسؤال القبل .

Fitting an Exponential by the Method of Least Squares

If we have the equation :- $Y = a e^{bx}$, to solve it easily :-

We use the Linearization Method :-

يعني نحوله لـ linear function :-

Take the (ln) for both sides :-

$$\ln[Y] = \ln[a e^{bx}]$$

$$= \ln[a] + \underbrace{\ln[e^{bx}]}_{bx}$$

$$\ln[Y] = \ln[a] + bx$$

$$Y_{new} = B + \alpha x$$ ← نحرّا السابق شبايه بصير خطي

$\boxed{B = \ln a}$ $\boxed{\alpha = b}$ $\boxed{Y_{new} = \ln[Y]}$

وبحل هاي المعادلة زي ما حلينا بالدرس السابق .

$\boxed{Ex}$ If $y = \dfrac{L}{1+e^{a+bx}}$

$$y(1+e^{a+bx}) = L \Rightarrow y + y e^{a+bx} = L$$

$$e^{a+bx} = \frac{L-y}{y} \xrightarrow[\text{for both sides}]{\text{take (ln)}} a + bx = \underbrace{\ln\left(\frac{L-y}{y}\right)}_{y_{new}}$$

$\boxed{y_{new} = \alpha + \beta x}$ → linear , $y_{new} = \ln\left(\frac{L-y}{y}\right)$, $\alpha = a$, $\beta = B$

Scanned by TapScanner

Central Limit Theorem

بتحكيلي كيف بقدر أحسب Probability معينة على sample مأخوذة من الـ population

$P(\hat{\mu}_x \leq 11) = ??$

<u>Note:-</u> $Y = c_1 x_1 + c_2 x_2 + c_3 x_3$

$$E\{Y\} = c_1 \mu_{x_1} + c_2 \mu_{x_2} + c_3 \mu_{x_3}$$

$$Var\{Y\} = c_1^2 \sigma_{x_1}^2 + c_2^2 \sigma_{x_2}^2 + c_3^2 \sigma_{x_3}^2$$

$$+ 2 c_1 c_2 \sigma_{x_1} \sigma_{x_2} \rho_{x_1,x_2} \leftarrow x_1 \text{ and } x_2$$

$$+ 2 c_1 c_3 \sigma_{x_1} \sigma_{x_3} \rho_{x_1,x_3} \leftarrow x_1 \text{ and } x_3$$

$\rho_{x_1,x_3} \rightarrow$ correlation coefficient

$$+ 2 c_2 c_3 \sigma_{x_2} \sigma_{x_3} \rho_{x_2,x_3} \leftarrow x_2 \text{ and } x_3$$

* if x_1, ~~and~~, x_2, and x_3 are S.I, then :-

$\rho_{x_1,x_2} = \rho_{x_1,x_3} = \rho_{x_2,x_3} = 0$ ← أخذناها آخر ch3

Ex Let x_1 and x_2 be two Gaussian random variables such that :- $\mu_1 = 0$, $\sigma_1^2 = 4$, $\mu_2 = 10$, $\sigma_2^2 = 9$, $\rho_{1,2} = 0.25$. Define $Y = 2x_1 + 3x_2$.

Q Find the mean and variance of Y.

$$\mu_Y = 2\mu_{x_1} + 3\mu_{x_2} = 2(0) + 3(10) = 30$$

$$\sigma_Y^2 = (2)^2 \sigma_{x_1}^2 + (3)^2 \sigma_{x_2}^2 + 2\sigma_{x_1}\sigma_{x_2}\rho_{x_1,x_2}$$

$$= 4(\overset{16}{4}) + 9(9) + 2\sqrt{4}\sqrt{9}(0.25)$$

$$= 115$$

$\boxed{b}$ Find $P(Y \leq 35)$. Y is Gaussian ✓

$\mu_y = 30$, $\sigma_y^2 = 115$

$\therefore \phi\left(\frac{35-30}{\sqrt{115}}\right) = \phi(0.466) = 0.6794$

$\boxed{Ex}$ Let x_1 and x_2 be ~~two~~ independent Gaussian Random Variables such that: $\mu_1 = 0$, $\sigma_1^2 = 4$, $\mu_2 = 10$, $\sigma_2^2 = 9$. Define $Y = 2x_1 + 3x_2$.

$\boxed{\rho_{x_1, x_2} = 0}$

$\boxed{a}$ Find the mean and variance of Y.

$\mu_Y = 2\mu_{x_1} + 3\mu_2$

$= 2(0) + 3(10)$

$= 30$

$\sigma_Y^2 = 4\sigma_{x_1}^2 + 9\sigma_{x_2}^2 + \cancel{2\sigma_{x_1}\sigma_{x_2}\rho_{x_1,x_2}}$

$= 4(4) + 9(9)$ ($16 + 81$)

$= 97$

$\boxed{b}$ $P(Y \leq 35)$ → Gaussian ✓ , $\mu_Y = 30$, $\sigma_x^2 = 97$

$P(Y \leq 35) = \phi\left(\frac{35-30}{(\sqrt{97})}\right) = \phi(0.5077) = 0.6942$

$\boxed{Ex}$ Soft drink cans are filled by an automated filling machine. The mean fill volume is 330 ml, and the standard deviation is $\boxed{1.5 \text{ ml}}$. Assume that the fill volume of the cans are $\boxed{\text{independent}}$ Gaussian Random Variables. What is the probability that the average volume of 10 cans selected at random from this process is less than 328 ml.

$\rho_{x,x} =$

$P(\hat{\mu}_x \leq 328) = ??$

→

$\mu_x = 330$ ml
$\sigma_x = 1.5$
Population cans

$x_1, x_2, x_3, \dots, x_{10}$ $\quad n=10$

As $P(\hat{\mu}_x \le 328) = ??$
we deal with $\hat{\mu}_x$ as if it is Y

$$\therefore \hat{\mu}_x = Y = \frac{1}{n}\sum x_i$$

$$= \frac{1}{n}x_1 + \frac{1}{n}x_2 + \frac{1}{n}x_3 + \dots + \frac{1}{n}x_{10}$$

$$E\{Y\} = E\{\hat{\mu}_x\} = E\{\quad\}$$

بنحطها ↓ ..

E بتوزيع

$$\therefore E\{\hat{\mu}_x\} = \frac{1}{n}E\{x_1\} + \frac{1}{n}E\{x_2\} + \frac{1}{n}E\{x_3\} + \dots \frac{1}{n}E\{x_n\}$$

$$= \frac{1}{n}\mu_x + \frac{1}{n}\mu_x + \frac{1}{n}\mu_x + \dots + \frac{1}{n}\mu_x$$

$$= \frac{1}{\cancel{n}}(\cancel{n})(\mu_x)$$

عشان الـ X هي مأخوذة من الـ Population، فالـ E بتعتها هي نفسها الـ E الموجودة بالـ Population.

$$= \mu_x$$
$$= 330$$

* بتقول الـ Central Limit Theorem :- إذا باخد Sample من الـ Population الـ expected value للـ Sample Mean ، $E\{\hat{\mu}_x\}$ يساوي الـ mean الحقيقي.

$$\boxed{E\{\hat{\mu}_x\} = \mu_x}$$

$$Var\{\hat{\mu}_x\} = \left(\frac{1}{n}\right)^2 \sigma_{x_1}^2 + \left(\frac{1}{n}\right)^2 \sigma_{x_2}^2 + \dots + \left(\frac{1}{n}\right)^2 \sigma_{x_{10}}^2$$

نفس الإشي بما أنا نعتبر كل X مأخوذة من الـ Population فهي لها نفس σ_x^2

$$= \frac{1}{n^2}(n)\sigma_x^2 = \frac{\sigma_x^2}{n} = \frac{(1.5)^2}{10} = 0.225$$

$\therefore$ Gaussian $\hat{\mu}_x = 330$, $\hat{\sigma}_x^2 = 0.225$

$$P(\hat{\mu}_x < 328) = \phi\left(\frac{328-330}{\sqrt{0.225}}\right) = \phi\left(\frac{-2}{0.4743}\right)$$

$$= \phi(-4.2163)$$

وبطلع الجواب من الجدول .

Scanned by TapScanner

$\boxed{\text{Ex 5-10}}$ An Electric company manufactures resistors that have a mean resistence of 100 Ω and a standard deviation of 10 Ω. Find the probability that a random sample of $\boxed{n=25}$ resistors will have an average resistance less than 95 Ω.

Random Sample → إزاي هعرف Random Sample طول لازم أعرف إنهم independent.

$\mu_x = 100\,\Omega \quad , \quad \sigma_x = 10 \quad , \quad n = 25$

μ_x vs

$P(\hat{\mu}_x < 95) = ??$ → هنا ح نطلع $\hat{\mu}_x$ كـ متغير عادي x, y وهكذا لازم mean و Var.

Using the centeral theorem :-

$$\boxed{\underset{\text{العينة}}{\hat{\mu}_x} = \underset{\text{pop. الـ}}{\mu_x} = 100}$$

$$\hat{\sigma}_x^2 = \frac{\sigma_x^2}{n} = \frac{(10)^2}{25} = 4$$

$$\therefore P(\hat{\mu}_x < 95) = \phi\left(\frac{95-100}{\sqrt{4}}\right) = \phi\left(\frac{-5}{2}\right) = 0.00621$$

↳ suppose it is a Gaussian ✓

$\hat{\mu}_x = 100$ ✓

$\hat{\sigma}_x^2 = 4$ ✓

$\boxed{\text{Ex}}$ The lifetime of a special type of battery is a R.V with mean 40 hours and standard deviation 20 hours. A battery is used until it fails, then it is immediately replaced by a new one. Assume we have 25 such batteries, the lifetime of which are independent, approximate the probability that at least $\boxed{1100 \text{ hours}}$ of use can be obtained.

$\boxed{\mu_x = 40}$ $\boxed{\sigma_x = 20}$ $\boxed{n = 25}$

Let $X_1, X_2, X_3, \ldots X_{25}$ be the lifetimes of batteries.

Let $Y = X_1 + X_2 + \ldots + X_{25}$ be the overall lifetime of the system.

since x_i are independent, Using Gaussian :-

$$\mu_Y = \mu_1 + \mu_2 + \ldots \mu_{25} = 25\mu_x = 25 \times 40 = 1000$$

$$\sigma_Y^2 = \sigma_{x_1}^2 + \sigma_{x_2}^2 + \ldots + \sigma_{x_{25}}^2 = 25(20)^2 = 10000$$

$$P(Y \geq 1100) = 1 - P(Y < 1100) = 1 - \phi\left(\frac{1100-1000}{\sqrt{10000}}\right) = 1 - \phi\left(\frac{100}{100}\right)$$

$$= 1 - \phi(1) = 0.158655$$

Estimation of Parameters :-

ع جدول لصفحة 77 موجودة في القوانين اللي أخذنا هو حيل .

Estimator :- is a function of the observable Sample data that is used to estimate an unknown Population parameter (μ_x, σ_x^2, ...).

We consider two types of estimators :
1. Point Estimator
2. Interval Estimator.

Point Estimator

Properties : [1] An estimator should be close to the true value of the unknown Parameters.

Def : A Point estimator $(\hat{\theta})$ is unbiased estimator of (θ) if $\boxed{E(\hat{\theta}) = \theta}$.

- If the estimator is biased, then $\boxed{E(\hat{\theta}) - \theta = B}$ is called the bias of the estimator $(\hat{\theta})$

- Let $\hat{\theta}_1, \hat{\theta}_2$ be unbiased estimators of (θ), the one with the Smallest Variance is called the minimum Variance unbiased estimator (MVUE).

In other words

$E\{\hat{\mu}_x\} = \mu_x \leftarrow \hat{\mu}_x$ is an unbiased estimator for the mean μ_x.

$E\{\hat{P}\} = P \leftarrow \hat{P}$ is an unbiased estimator for the Probability of success P.

Scanned by TapScanner

بنحسب $E\{\hat{\mu}_x\}$ حسبها عادي زي ما أخذنا بـ chapter2 :-

$$E\{X\} \triangleq \sum_{x=-\infty}^{\infty} x\, P(X=x) \longrightarrow \text{PMF}$$

$$\triangleq \int_{-\infty}^{\infty} x\, f_x(x)\, dx \longrightarrow \text{pdf}$$

$\boxed{Ex}$ let X_1 and X_2 be a random sample of size two from a Population with mean μ_x and variance σ_x^2. Two estimators for μ_x are proposed:-

$\mu_1 = \frac{X_1+X_2}{2}$ and $\mu_{x_2} = \frac{X_1+2X_2}{3}$. which estimator is better and in what sense?

هي الـ estimator الأول، مش لحاله واحد

$$E\{\mu_{x_1}\} = E\left\{\frac{X_1+X_2}{2}\right\} = E\left\{\frac{X_1}{2}\right\} + E\left\{\frac{X_2}{2}\right\} = \frac{1}{2}\mu_x + \frac{1}{2}\mu_x$$

$= \mu_x \Rightarrow$ as $E\{\hat{\mu}_{x_1}\} = \mu_1 \Rightarrow \therefore \hat{\mu}_{x_1} = \frac{X_1+X_2}{2}$ is unbiased.

For $\hat{\mu}_{x_2} \Rightarrow E\{\hat{\mu}_{x_2}\} = E\left\{\frac{X_1+2X_2}{3}\right\} = E\left\{\frac{X_1}{3}\right\} + \frac{2}{3}E\{X_2\}$

$= \frac{1}{3}\mu_x + \frac{2}{3}\mu_x = \mu_x \rightarrow \hat{\mu}_{x_2} = \frac{X_1+2X_2}{3}$ is unbiased

$\rightarrow$ Estimation إضافي مو كتير :- $\mu_{x_3} = \frac{X_1+X_2+1}{3}$

$$\therefore E\{\mu_{x_3}\} = E\left\{\frac{1}{3}X_1 + \frac{1}{3}X_2 + \frac{1}{3}\right\} = \frac{1}{3}E\{X_1\} + \frac{1}{3}E\{X_2\} + \frac{1}{3}$$

$= \frac{2}{3}\mu_x + \frac{1}{3} \rightarrow$ is a biased estimator

مقدار الـ biase هيك بنحسبه :-

$$\boxed{B = E\{\hat{\theta}\} - \theta}$$

$$= E\{\hat{\mu}_x\} - \mu_x$$

$$= \frac{2}{3}\mu_x + \frac{1}{3} - \mu_x = \frac{1}{3} - \frac{1}{3}\mu_x$$

$$\therefore \boxed{B = \frac{1-\mu_x}{3}}$$

بما إنو طلع عندي 2 biased و 1 unbiased ، فأكيد رح أستثني الـ biaseds ، باخذ الـ unbiased ، وهنا ~ أشوف أنو أفضل من بين الـ estimator اللي الـ variance تبعو أقل .

$$Var\{\hat{\mu}_{x,2}\} = Var\left\{\frac{1}{3}x_1 + \frac{2}{3}x_2\right\}$$

$$= \frac{1}{9}\sigma_x^2 + \frac{4}{9}\sigma_x^2$$

$$\boxed{\sigma^2_{\hat{\mu}_{x,2}} = \frac{5}{9}\sigma_x^2}$$

حسب ما أخذنا بالـ central Theo. إذا كان x_1, x_2 SI و

$$y = c_1x_1 + c_2x_2$$

$$\therefore \sigma_y^2 = c_1^2\sigma_x^2 + c_2^2\sigma_{x_2}^2$$

$$Var\{\mu_{x,1}\} = Var\left\{\frac{x_1 + x_2}{2}\right\}$$

$$= \frac{1}{4}\sigma_x^2 + \frac{1}{4}\sigma_x^2 = \frac{1}{2}\sigma_x^2$$

As $Var\{\hat{\mu}_{x,1}\} < Var\{\hat{\mu}_{x,2}\}$

$\therefore$ $\hat{\mu}_{x,1}$ is the best estimator for μ_x .

Example Check whether the following estimator is biased or unbiased. Then try to modify the estimator to be unbiased if it is found to be biased.

$$\hat{\sigma}_x^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \hat{\mu}_x)^2$$

$$E\{\hat{\sigma}_x^2\} = E\left\{\frac{1}{n}\sum_{i=1}^{n}(x_i - \hat{\mu}_x)^2\right\} = E\left\{\frac{1}{n}\sum\left(x_i^2 - 2x_i\hat{\mu}_x + \hat{\mu}_x^2\right)\right\}$$

إزا

بتوزيع المجموع

$$= \frac{1}{n}E\left\{\sum_{x=1}^{n}x_i^2 - 2\hat{\mu}_x\sum_{x=1}^{n}x_i + n\hat{\mu}_x^2\right\}$$

$$= \frac{1}{n}E\left\{\sum_{x=1}^{n}x_i^2 - 2\hat{\mu}_x\, n\hat{\mu}_x + \cancel{n\hat{\mu}_x^2}\right.$$

حسب ما أخذنا

$$\frac{\sum x_i}{n} = \hat{\mu}_x$$ بتعريف

$$= \frac{1}{n} E\left\{\sum_{i=1}^{n} x_i^2 - n\hat{\mu}_x^2\right\}$$

$$= \frac{1}{n}\left[\sum E\{x_i^2\} - n E\{\hat{\mu}_x^2\}\right]$$

تكملة

* Note :- (1) $Var\{?\} = E\{?^2\} - (E\{?\})^2$

أضفناها بتشابتر 2 أظن ما طلعة عالصفحة بجمة مس تح عنده نوخها

$\therefore Var\{x\} = E\{x^2\} - (E\{x\})^2$

$\therefore \sigma_x^2 = E\{x^2\} - \mu_x^2$

$\therefore \boxed{E\{x^2\} = \sigma_x^2 + \mu_x^2}$ —— (1)

(2) $\therefore$ According to the Previous Note :-

$$Var\{\hat{\mu}_x\} = E\{\hat{\mu}_x^2\} - (E\{\hat{\mu}_x\})^2$$

According Central Limit theorem

$$\frac{\sigma_x^2}{n} = E\{\hat{\mu}_x^2\} - \mu_x^2$$

$\therefore \boxed{E\{\hat{\mu}_x^2\} = \frac{\sigma_x^2}{n} + \mu_x^2}$ —— (2)

$$\therefore = \frac{1}{n}\left(\sum \sigma_{x_i}^2 + \mu_x^2\right) - n\left(\frac{\sigma_x^2}{n} + \mu_x^2\right)$$

$$= \frac{1}{n}\left(n\sigma_x^2 + n\mu_x^2 - \sigma_x^2 - n\mu_x^2\right)$$

$$\therefore E\{\hat{\sigma}_x^2\} = \frac{1}{n}\left(n\sigma_x^2 - \sigma_x^2\right) = \frac{n-1}{n}\sigma_x^2 \rightarrow \text{biased}$$

لأنو ما طلع = σ_x^2 لحالها حسب تعريف الـ unbiased.

فنجرب نعكوس المقابل تبع σ_x^2 عشان يطلع unbiased هيك

$$\therefore E\left\{\frac{n}{n-1}\hat{\sigma}_x^2\right\} = E\left\{\frac{\cancel{n}}{n-1} * \frac{1}{\cancel{n}} \sum (x_i - \hat{\mu}_x)^2\right\} = \frac{\cancel{n}}{\cancel{n-1}} * \frac{\cancel{n-1}}{\cancel{n}}\sigma_x^2$$

$$\therefore \hat{\sigma}_{x_{new}}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \hat{\mu}_x)^2 \Rightarrow E\{\hat{\sigma}_{x_{new}}^2\} = \sigma_x^2$$

إذا طلع العامل 20 بس نجرب الـ 19 بالمجموع وخلص فش داعي للـ 20 بالبسط عشان نحكيها $\hat{\sigma}_x^2$

The Maximum Likelihood (ML) Estimator

Ex The probability $P = P(H)$ of a coin may be 0.1 or it may be 0.9. To resolve the uncertainty, the coin was tossed 10 times and 3 heads were observed. Find a maximum likelihood estimate for P.

$\rightarrow$ Probability of success

$P = 0.1$ or $P = 0.9$ + تلاقي إنو binomial

X: number of heads observed is a binomial R.V.

$$P(X=x) = \binom{n}{x} P^x [1-P]^{n-x}, \quad n=10$$

$$\therefore P(X=x) = \binom{10}{x} P^x [1-P]^{10-x}$$

$$P(X=3 \overset{as}{;} P=0.1) = \binom{10}{3}(0.1)^3(0.9)^7 = 0.0574$$

But, when:-

$$P(X=3; P=0.9) = \binom{10}{3}(0.9)^3(0.1)^7 = 8.748 \times 10^{-6}.$$

As when $\boxed{P=0.1}$ has the higher probability, thus we choose it as our estimator. $\therefore \boxed{\hat{P} = 0.1}$

B if $\boxed{X=8}$; $P(X=8; P=0.1) = \binom{10}{8}(0.1)^8(0.9)^2 = 3.645 \times 10^{-\ldots}$

$$P(X=8; P=0.9) = \binom{10}{8}(0.9)^8(0.1)^2 = 0.1937.$$

In this case, we choose $\boxed{\hat{P} = 0.9}$ as the estimator.

Example Let P be the probability of success in a binomial distribution. This probability is unknown. To estimate P, the experiment is performed 10 times and 3 successes were observed. Find a maximum likelihood estimate for p.

binomial

$n = 10$, $X = 3$

$P = ??$ → we find it

$$f(p) = P(X=3\,;\,p) = \binom{10}{3} p^3 (1-p)^7 = \frac{10!}{3!\,(10-7)!} p^3 (1-p)^7$$

عشان نوجد أعلى قيمة للـ success ، بنشتق المعادلة

بالنسبة لـ P ونساويها بالصفر.

$$\therefore \frac{df}{dp} = \binom{10}{3}\left[-7p^3(1-p)^6 + 3p^2(1-p)^7\right] = 0$$

وبيطلع معي إشي 0.25 , 0.35 و $P = 0.3$

ونعوض بهاي المعادلة وبشوف أي نتيجة كانها البروبابلتي أعلى

والقيمة هاي هي بتكون $\hat{P}$ ، وبالآخر طلع معنا

$$\boxed{\hat{P} = \frac{3}{10} = 0.3}$$

For the continuous case :- بكتب ال cdf الخاص بالفنكشن
زي ما كتبنا ال pdf بال Discrete .

$\boxed{\text{Example (6-3)}}$ Given a random sample of size (n) taken from a Gaussian population with parameters μ_x and σ_x^2. Use the ML technique to find estimators for the cases :-

$\boxed{a}$ The mean μ_x when the variance σ_x^2 is assumed known

Gaussian μ_x , σ_x^2 → random sample size n → $x_1, x_2, \dots, x_n$

$$f_{x_1}(x_1) = \frac{1}{\sqrt{2\pi\sigma_{x_1}^2}} e^{\frac{-(x_1-\mu_{x_1})^2}{2\sigma_{x_1}^2}}$$

$\mu_{x_1} = \mu_x$ لأنها من و $\sigma_{x_1}^2 = \sigma_x^2$

$$= \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(x_1-\mu_x)^2}{2\sigma_x^2}}$$

بس $x_1 \neq x$ بس طبعاً إذا كان $z = x_1$ أمشي على ال x لـ z .

$\boxed{\text{And}}$

$$f_{x_2}(x_2) = \frac{1}{\sqrt{2\pi\sigma_x^2}} e^{\frac{-(x_2-\mu_x)^2}{2\sigma_x^2}}$$

$\vdots$

$$\therefore f_{x_n}(x_n) = \frac{1}{\sqrt{2\pi\sigma_x^2}} e^{\frac{-(x_n-\mu_x)^2}{2\sigma_x^2}}$$

بما أنه إحنا ماخدين random sample فالعينة independent يعني كل واحد لا يعتمد على الأخرى

$$\therefore f_{x_1, x_2, \dots, x_n}(x_1, x_2, \dots, x_n) = f_{x_1}(x_1) f_{x_2}(x_2) \dots f_{x_n}(x_n)$$

$$= \left[\frac{1}{\sqrt{2\pi\sigma_x^2}}\right]^n e^{\frac{-1}{2\sigma_x^2}\sum_{i=1}^{n}(x_i-\mu_x)^2}$$

حطينا مجموع علشان لما نضرب e نجمع الأسس .

$$(2\pi\sigma_x^2)^{\frac{-n}{2}} e^{\frac{-1}{2\sigma_x^2}\sum_{i=1}^{n}(x_i-\mu_x)^2}$$

$\ln[f_{x_1,\dots,x_n}(x_1,\dots,x_n)]$

$$L(\mu_x, \sigma_x^2) = \frac{-n}{2}\ln(2\pi\sigma_x^2) - \frac{1}{2\sigma_x^2}\sum_{i=1}^{n}(x_i-\mu_x)^2$$

صار اسمه ML function
Maximum Likelihood

Scanned by TapScanner

هلأ بلّش حل $\boxed{a}$ حيث إنو رح أستخدم المعادلة السابقة في إيجاد الـ $\hat{\mu}_x$ المقدّر، على فرض إني بعرف σ_x^2.

$\boxed{a}$ Find $\hat{\mu}_x$, provided that σ_x^2 is known:-

$$L(\hat{\mu}_x) = \frac{-n}{2}\ln(2\pi\sigma_x^2) - \frac{1}{2\sigma_x^2}\sum_{i=1}^{n}(x_i-\hat{\mu}_x)^2$$

هو بركّز الاشي اللي بدي أوجد الـ estimation الو، وبشتق المعادلة بالنسبة له.

$$\frac{dL(\hat{\mu}_x)}{d\hat{\mu}_x} = \underbrace{0}_{\text{مشتقتها بالنسبة لـ } \hat{\mu}_x} - \frac{1}{2\sigma_x^2}\sum_{i=1}^{n}2(x_i-\hat{\mu}_x)(-1) = 0$$

$$\therefore \frac{-1}{2\sigma_x^2}\sum_{i=1}^{n}2(x_i-\hat{\mu}_x)(-1) = 0$$

$$\sum_{i=1}^{n}x_i - \sum_{i=1}^{n}\hat{\mu}_x = 0$$

$$\therefore \sum_{i=1}^{n}x_i = n\hat{\mu}_x \Rightarrow \boxed{\hat{\mu}_x = \frac{\sum x_i}{n}}$$ ← ML estimator for the mean when the variance is known.

Now we check for biasing:-

$$E\{\hat{\mu}_x\} = E\left\{\frac{1}{n}\sum x_i\right\} = \frac{1}{n}\sum_{i=1}^{n}E\{x_i\} = \frac{1}{n}\sum_{i=1}^{n}\mu_x$$

$$= \frac{1}{n}(n)\mu_x = \mu_x$$

$$\boxed{E\{\hat{\mu}_x\} = \mu_x}$$ so it is unbiased.

$\boxed{a_2}$ $\hat{\mu}_x$, Variance is unknown.

$$L(\hat{\mu}_x, \hat{\sigma}_x^2) = -\frac{n}{2}\ln(2\pi\hat{\sigma}_x^2) - \frac{1}{2\hat{\sigma}_x^2}\sum_{i=1}^{n}(x_i-\hat{\mu}_x)^2$$

$$\frac{dL(\hat{\mu}_x, \hat{\sigma}_x^2)}{d\hat{\mu}_x} = 0 + \frac{1}{\cancel{2\hat{\sigma}_x^2}}\sum_{i=1}^{n}\cancel{2}(x_i-\hat{\mu}_x)\cancel{(-1)} = 0$$

$$\therefore \sum_{i=1}^{n}x_i - \sum_{i=1}^{n}\hat{\mu}_x = 0 \Rightarrow \boxed{\hat{\mu}_x = \frac{1}{n}\sum_{i=1}^{n}x_i}$$ ← ML estimator for the mean, when the variance is unknown;

تكملة

Check for biasing: $E\{\hat{\mu}_x\} = \ldots = \mu_x$

So it is biased

زي الفرع اللي قبله يعني.

[b] Estimator for σ_x^2, μ_x is known.

$$L(\hat{\sigma}_x^2) = \frac{-n}{2}\ln(2\pi\hat{\sigma}_x^2) - \frac{1}{2\hat{\sigma}_x^2}\sum_{i=1}^{n}(x_i-\mu_x)^2$$

$$\frac{dL(\hat{\sigma}_x^2)}{d\hat{\sigma}_x^2} = \left(-\frac{n}{2} * \frac{1}{2\pi\hat{\sigma}_x^2} * 2\pi\right) + \frac{1}{2}\left(\frac{+1}{(\hat{\sigma}_x^2)^2}\right)\sum_{i=1}^{n}(x_i-\mu_x)^2 = 0$$

$$\frac{+n}{\cancel{2}\hat{\sigma}_x^2} = \frac{1}{\cancel{2}(\hat{\sigma}_x^2)^2}\sum_{i=1}^{n}(x_i-\mu_x)^2$$

$$\hat{\sigma}_x^2 = \frac{\sum(x_i-\mu_x)^2}{n}$$ $\rightarrow$ ML estimator for the variance when the mean is known.

$\Rightarrow$ check for biasing:

$$E\{\hat{\sigma}_x^2\} = E\left\{\frac{1}{n}\sum_{i=1}^{n}(x_i-\mu_x)^2\right\}$$

$$= \frac{1}{n}\sum_{i=1}^{n}\underbrace{E\{(x_i-\mu_x)^2\}}$$

$$= \frac{1}{n}\sum_{i=1}^{n}\sigma_x^2 = \frac{1}{\cancel{n}} * \cancel{n}\,\sigma_x^2 = \sigma_x^2$$

So it is an unbiased estimator for the variance when the mean is known.

[c] $\hat{\sigma}_x^2$ when μ_x is unknown.

$$L(\hat{\mu}_x, \hat{\sigma}_x^2) = \frac{-n}{2}\ln(2\pi\hat{\sigma}_x^2) - \frac{1}{2\hat{\sigma}_x^2}\sum_{i=1}^{n}(x_i-\hat{\mu}_x)^2$$

هون بنحط الاثنين unknown

$$\frac{dL(\hat{\mu}_x, \hat{\sigma}_x^2)}{d\sigma_x^2} = \frac{-n}{2} * \frac{1}{\cancel{2\pi}\hat{\sigma}_x^2} * \cancel{2\pi} - \frac{1}{2} * \frac{-1}{(\hat{\sigma}_x^2)^2}\sum_{i=1}^{n}(x_i-\hat{\mu}_x)^2 = 0$$

هون ما بنشتق بدون estimator حسب ليش بالسؤال

$$\frac{n}{2} * \frac{1}{\hat{\sigma}_x^2} = \frac{1}{2} * \frac{1}{(\hat{\sigma}_x^2)^2} \sum_{i=1}^{n} (x_i - \hat{\mu}_x)^2$$

$$\therefore \boxed{\hat{\sigma}_x^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \hat{\mu}_x)^2} \rightarrow$$ ML estimator for the variance when the mean is unknown.

→ Check for biasing :-

$$E\{\hat{\sigma}_x^2\} = E\left\{\frac{1}{n} \sum_{i=1}^{n} (x_i - \hat{\mu}_x)^2\right\}$$

$$= E\left\{\frac{1}{n} \sum_{i=1}^{n} (x_i^2 - 2x_i\hat{\mu}_x + \hat{\mu}_x^2)\right\}$$

$$= \frac{1}{n} E\left\{\sum x_i^2 - 2\hat{\mu}_x \sum_{i=1}^{n} x_i + \sum \hat{\mu}_x^2\right\}$$

$$= \frac{1}{n} E\left\{\sum x_i^2 - \cancel{2}\hat{\mu}_x\, n\hat{\mu}_x + \cancel{n}\hat{\mu}_x^2\right\}$$

$$* \quad \sigma_x^2 = E\{x^2\} - \mu_x^2$$

$$\therefore E\{x^2\} = \sigma_x^2 + \mu_x^2$$

$$* \operatorname{Var}\{\hat{\mu}_x\} = \frac{\sigma_x^2}{n}$$

$$\therefore E\{\hat{\mu}_x^2\} - (E\{\hat{\mu}_x\})^2 = \frac{\sigma_x^2}{n}$$

$$\therefore E\{\hat{\mu}_x^2\} = \frac{\sigma_x^2}{n} + \mu_x^2$$

$$= \frac{1}{n} E\left\{\sum x_i^2 - n\hat{\mu}_x^2\right\}$$

$$= \frac{1}{n}\left[\sum E\{x_i^2\} - n E\{\hat{\mu}_x^2\}\right]$$

$$= \frac{1}{n}\left[\sum_{i=1}^{n} [\sigma_x^2 + \mu_x^2] - n\left[\frac{\sigma_x^2}{n} + \mu_x^2\right]\right]$$

$$= \frac{1}{n}\left[n\sigma_x^2 + \cancel{n\mu_x^2} - \sigma_x^2 - \cancel{n\mu_x^2}\right]$$

$$= \frac{n-1}{n}\sigma_x^2 .$$

* Check for biasing :-

$$E\{\hat{\sigma}_x^2\} = E\left\{\frac{1}{n} \sum_{i=1}^{n} (x_i - \hat{\mu}_x)^2\right\} = \frac{n-1}{n}\sigma_x^2$$

علشان نخليها unbiased ← بنضرب بـ $\frac{n}{n-1}$ علشان نخلي σ_x^2 لحالها ويصير unbiased

$$\therefore E\{\hat{\sigma}_{x_{new}}^2\} = E\left\{\frac{n}{n-1} * \frac{1}{n} \sum_{i=1}^{n} (x_i - \hat{\mu}_x)^2\right\} = \frac{\cancel{n}}{\cancel{n-1}} * \frac{\cancel{n-1}}{\cancel{n}} \sigma_x^2$$

$$\therefore \boxed{\hat{\sigma}_{x_{new}}^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \hat{\mu}_x)^2} \rightarrow$$ An unbiased ML estimator of the variance, when the mean IS unknown.

Ex 6-4 Given a random sample of size (n) taken from a distribution X with pdf :-

$$f(x) = (\alpha+1)\,x^{\alpha} \quad , 0 < x < 1 .$$

Use ML technique to find an estimator for α.

The likelihood function is:-

$$L(\alpha) = f(x_1)\, f(x_2)\, f(x_3) \ldots f(x_n)$$

وأفضلنا كامل مشترك

$$L(\alpha) = (\alpha+1)\, x_1^{\alpha} \ldots (\alpha+1)\, x_n^{\alpha} = (\alpha+1)^n\, x_1^{\alpha} \ldots x_n^{\alpha}$$

Enter Ln to both sides:-

$$\ln[L(\alpha)] = n \ln(\alpha+1) + \alpha \ln x_1 + \alpha \ln x_2 + \ldots \alpha \ln x_n$$

$$\ln[L(\alpha)] = n \ln(\alpha+1) + \alpha \sum_{i=1}^{n} \ln(x_i)$$

$$\therefore \frac{d \ln[L(\alpha)]}{d\alpha} = n \frac{1}{1+\alpha} + \sum_{i=1}^{n} \ln(x_i) = 0$$

المعكوس

$$\frac{-n}{1+\hat{\alpha}} = \sum_{i=1}^{n} \ln(x_i) \Rightarrow \frac{1+\hat{\alpha}}{-n} = \frac{1}{\sum_{i=1}^{n} \ln(x_i)}$$

$$1+\hat{\alpha} = \frac{-n}{\sum_{i=1}^{n} \ln(x_i)} \Rightarrow \boxed{\hat{\alpha} = \frac{-n}{\sum_{i=1}^{n} \ln(x_i)} - 1}$$

هاد ال estimator المطلوب بالسؤال

Scanned by TapScanner

Finding Interval Estimators for the mean and Variance

$$P(\Theta_1 \leq \Theta \leq \Theta_2) = 1-\alpha \quad ; \quad 0 < \alpha < 1$$

Where:- (Θ) is the unknown parameter.

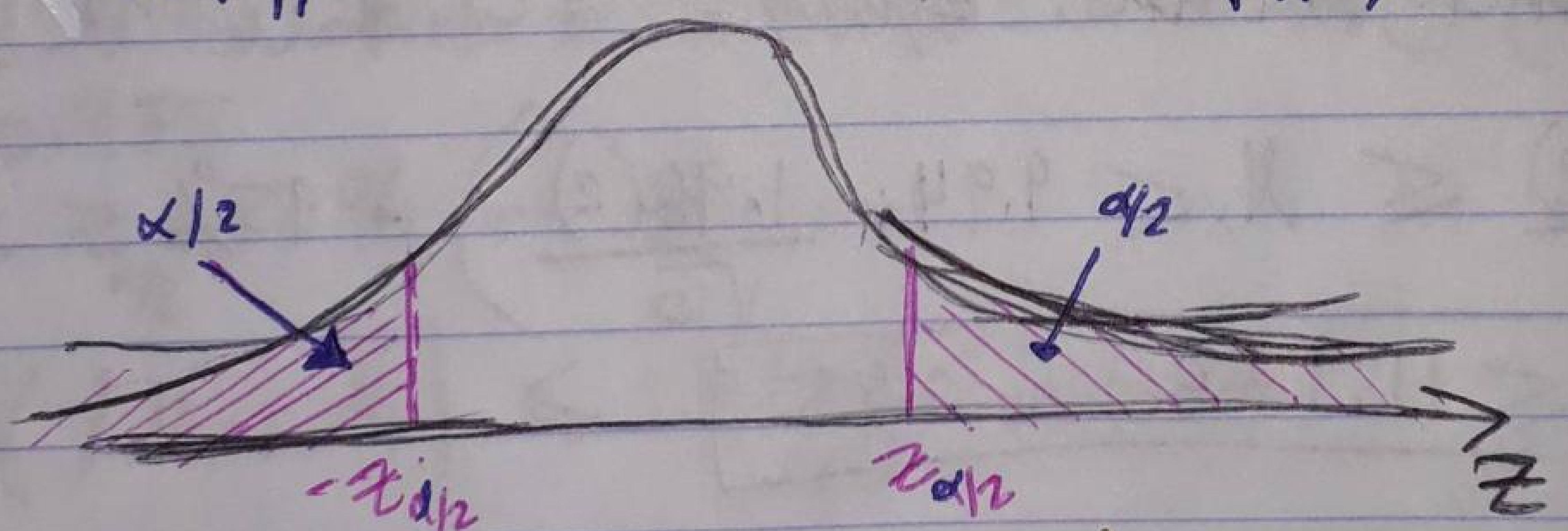
$(1-\alpha)$ is the confidence coefficient

(α) is the confidence level.

(Θ_1) and (Θ_2) are the lower and the upper confidence limits.

$\boxed{I}$ Confidence Interval on the Mean: (Known Variance).

$$P\left(\hat{\mu}_x - Z_{\alpha/2}\sqrt{\frac{\sigma_x^2}{n}} \leq \mu_x \leq \hat{\mu}_x + Z_{\alpha/2}\sqrt{\frac{\sigma_x^2}{n}}\right) = 1-\alpha$$



$\boxed{Ex}$ The following samples are drawn from a population that is known to be Gaussian.

7.31	10.80	11.27	11.91	5.51	8.00	9.03	14.42	10.24	10.91

Find the confidence limits for 95% confidence level if the variance of population is 4.

Gaussian ✓

Variance known ✓

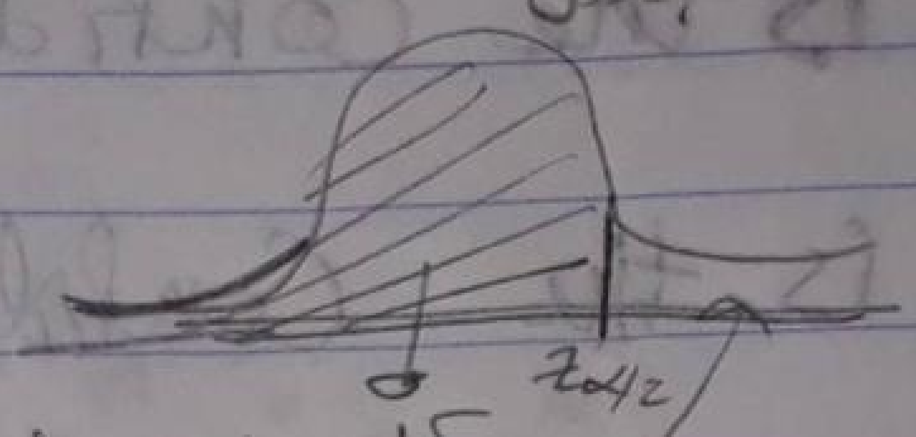
$\therefore$

$$P\left(\hat{\mu}_x - Z_{\alpha/2}\sqrt{\frac{\sigma_x^2}{n}} \le \mu_x \le \hat{\mu}_x + Z_{\alpha/2}\sqrt{\frac{\sigma_x^2}{n}}\right) = 1-\alpha$$

$\boxed{n = 10}$ $\boxed{1-\alpha = 0.95}$ $\boxed{\sigma_x^2 = 4}$ $\boxed{\hat{\mu}_x = \frac{1}{n}\sum_{i=1}^{n} x_i = 9.94}$

مجموع ال xات اللي بالجدول

$1-\alpha = 0.95 \Rightarrow \alpha = 0.05 \Rightarrow \alpha/2 = 0.025$



$\therefore \Phi(Z_{\alpha/2}) = 1 - 0.025$
$= 0.975$

وبروح عالجدول بشوف هاي القيمة نتيجة إيش z ؟ بتطلع z أكبر

كل هاد بدنا إياه ناقص هاي اللي هي $z_{\alpha/2}$

$\therefore \boxed{Z_{\alpha/2} = 1.96}$

$\therefore$ Every thing in the equation is with us ✓ !

$$P\left(9.94 - \frac{1.96(2)}{\sqrt{10}} \le \mu_x \le 9.94 + \frac{1.96(2)}{\sqrt{10}}\right) = 95\%$$

$$\boxed{P(8.70 \le \mu_x \le 11.1796) = 0.95}$$

* Note ①: ممكن يسألني عن ال Confidence Interval الخاصة بال Error ~~كالتالي~~ وحين ننقل $\hat{\mu}_x$ عند ال μ_x بالنص وخلص القصة ، كالتالي :

$$P\left(-Z_{\alpha/2}\sqrt{\frac{\sigma_x^2}{n}} \le \mu_x - \hat{\mu}_x \le Z_{\alpha/2}\sqrt{\frac{\sigma_x^2}{n}}\right) = 1-\alpha$$

وهاذا بدي أجاوب بمعطيات a بإطلاع مع الحل ، يعني بيطلع معطاة أشوي ع $\hat{\mu}_x$

$$P\left(-1.96 * \sqrt{\frac{4}{10}} \le \mu_x - \hat{\mu}_x \le 1.96 * \sqrt{\frac{4}{10}}\right) = 0.95$$

Note ② بلاحظ بـ note [1] إنو الـ LHS و الـ RHS هم نفس الاشي ولكن تتغير إشارتهما ، وهذا يدل على القيمة المطلقة $\hat{\mu}$ ، إذا "برضو" ممكن يسألني عن ~~القيمة~~ الـ (Absolute value for the error) ، بمعنى :-

$$P\left(|\mu_x - \hat{\mu}_x| \leq Z_{\alpha/2}\sqrt{\frac{\sigma_x^2}{n}}\right) = 1-\alpha$$

For ⓐ → $P\left(|\mu_x - \hat{\mu}_x| \leq 1.96 * \sqrt{\frac{4}{10}}\right) = 0.95$

Note ③ Find the value of n such that the absolute value of the error in the mean [between the true mean and the sample mean] is less than 0.1, with 95% confidence coefficient where $\sigma_x^2 = 4$.

$$P\left(|\mu_x - \hat{\mu}_x| \leq Z_{\alpha/2}\sqrt{\frac{\sigma_x^2}{n}}\right) = 0.95$$

$1-\alpha = 0.95 \Rightarrow \alpha = 0.05 \Rightarrow \alpha/2 = 0.025 \Rightarrow \Phi(Z_{\alpha/2}) = 1 - 0.025$
$= 0.975$
$\therefore Z_{\alpha/2} = 1.96$ ← من الجدول نحوّل

$\therefore P\left(|\mu_x - \hat{\mu}_x| \leq \frac{1.96 * 2}{\sqrt{n}}\right) = 0.95$

$\Rightarrow 1.96 * \frac{2}{\sqrt{n}} < \frac{0.1}{1} \Rightarrow \left(\frac{1.96 * 2}{0.1}\right)^2 < (\sqrt{n})^2$

$\therefore \boxed{n = 1537}$

ملاحظة: كلما زادت نسبة التوقع اللي هي $\boxed{1-\alpha}$ ، كلما

زادت الـ interval بتعتنا، زي لما الأستاذ يحكي كم بتتوقعوا بنسبة 90% العلامات بتعت الطلاب بآخر امتحان ، فاحكي رح يحكي 10-15 ، 15-20 ، 15-25 . أما إذا قال شو بنسبة 99% تكون علامات الطلاب ، بيحكي لازم أشمل كل الطلاب ونسبة الخطأ تكون بس 1% ، فرضيًا المعظم رح يقول انحصر من 5-25 هاي علامات الطلاب رح تكون بنسبة 99% ، نسبة قليلة تجيب علامة أقل من 5 ،

2 Confidence Interval on the Mean: (Variance Unknown)

$$P\left\{\hat{\mu}_x - t_{\alpha/2}\sqrt{\frac{\hat{\sigma}_x^2}{n}} \leq \mu_x < \hat{\mu}_x + t_{\alpha/2}\sqrt{\frac{\hat{\sigma}_x^2}{n}}\right\} = 1-\alpha$$

على نفس سؤال قبل 6 لبس على اعتبار بعرفش ال Varince

$$\boxed{\hat{\mu}_x = \frac{1}{n}\sum x_i = 9.94} \qquad \hat{\sigma}_x^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \mu_x)^2 = 6.51$$

Number of degrees of freedom $= n-1 = 10-1 = 9$.

$$1-\alpha = 0.95 \Rightarrow \boxed{\alpha = 0.05} \Rightarrow \therefore \boxed{\alpha/2 = 0.025}$$

$\therefore$ From the table of t-distribution:-

$$t_{(\alpha/2), dof} = t_{0.025, 9} = 2.263$$

$$\therefore P\left\{\hat{\mu}_x - t_{\alpha/2, n-1}\frac{\hat{\sigma}_x}{\sqrt{n}} \leq \mu_x \leq \hat{\mu}_x + t_{\alpha/2, n-1}\frac{\hat{\sigma}_x}{\sqrt{n}}\right\} = 1-\alpha$$

$$\therefore P\left\{9.94 - 2.263\frac{\sqrt{6.51}}{\sqrt{10}} \leq \mu_x \leq 9.94 + 2.263\frac{\sqrt{6.51}}{\sqrt{10}}\right\} = 0.95$$

$$\therefore P\{8.11 \leq \mu_x \leq 11.77\} = 0.95$$

تم بحمد الله