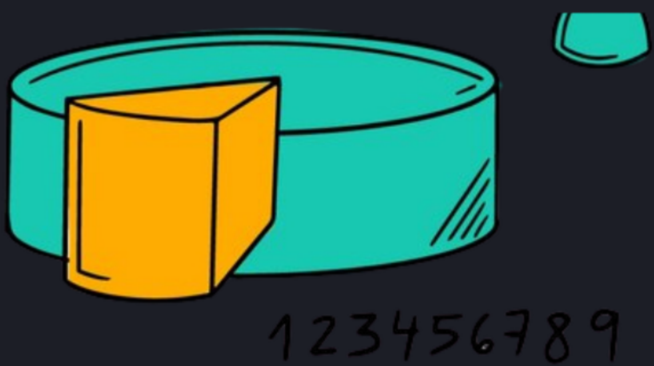
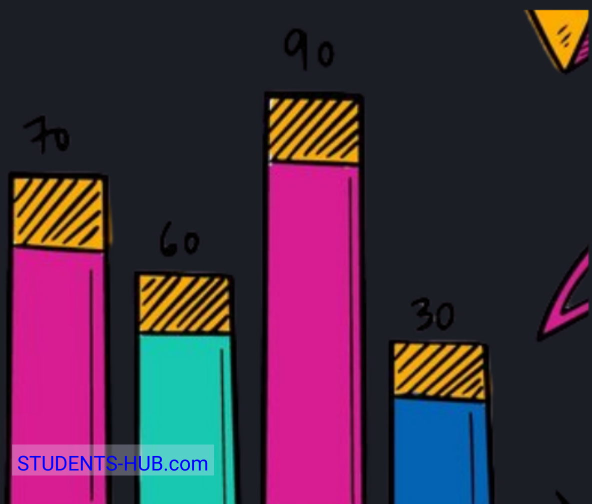
STATISTICS

By Rawan Alfares





1. Intersection $A \cap B$ means And $(\times)$
2. union $A \cup B$ means or $(+)$
3. Mutually Exclusive or disjoint $(A \cap B = \phi)$. $\longrightarrow$ $p(A \cup B) = p(A) + p(B)$.
4. Complement A^c or $\bar{A}$

5. Demorgan's law : $\overline{(A \cap B)} = \bar{A} \cup \bar{B}$
$\overline{(A \cup B)} = \bar{A} \cap \bar{B}$

6. $p(A) = \dfrac{\text{number of outcomes in event } A}{\text{number of total outcomes in } S}$

7. $1 \geq p(A) \geq 0$

⇒ **Conditional Probability and statistical independence.**

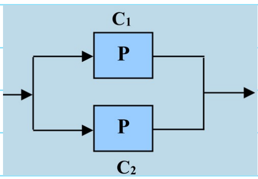
1. $p(A/B) = \dfrac{p(A \cap B)}{p(B)}$

2. $p(A \cap B) = p(A/B) \cdot p(B)$

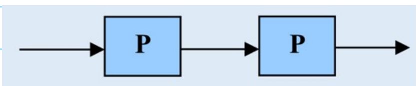
3. $p(A \cap B) = p(A) \cdot p(B)$ $\longrightarrow$ A and B are statistically indep.

* if A and B are S. Indep, then:
1. $p(A/B) = p(A)$
2. $p(B/A) = p(B)$

to prove that A, B and C are S. Indep :-
1. $p(A \cap B) = p(A) \cdot p(B)$
2. $p(A \cap C) = p(A) \cdot p(C)$
3. $p(B \cap C) = p(B) \cdot p(C)$
4. $p(A \cap B \cap C) = p(A) \cdot p(B) \cdot p(C)$.



$p(\text{system works}) = p(C_1 \cup C_2) = p(C_1) + p(C_2) - p(C_1 \cap C_2)$
$= P + P - (p(C_1) \cdot p(C_2))$ called Realiability of the system.
$= 2P - P^2$



$p(\text{system works}) = p(C_1 \cap C_2) = p(C_1) \cap p(C_2)$
$= P \cap P = P \cdot P = P^2$

Counting techniques :-

1. multiplication Rule :- operation A can be performed in n_1 different ways. and operation B can be performed in n_2 different ways. then the sequence (A, B) can be performed in $(n_1 \times n_2)$ different ways.

الترتيب مهم / التكرار غير مسموح

2. Permutation :- A- Sampling without replacement :- $N = \dfrac{n!}{(n-R)!}$ — احتمالية أمر معين بدون ترتيب

عدد مرات التجربة

B- Sampling with Replacement :- $N = n^R$ — الترتيب مهم التكرار مسموح

عدد الاحتمالات

3. Combination :- الترتيب غير مهم , التكرار غير مسموح .

$\binom{n}{k} = \dfrac{n!}{k!\,(n-k)!}$, n choose k = $C(k, n)$

Random Variable

Continuoes

PDF

$f_x(x)$

$\ast$ $\int_{-\infty}^{\infty} f_x(x)\,dx = 1$

$\ast$ $F_x(x) = \int_{-\infty}^{\infty} f_x(x)\,dx$

$\ast$ $f_x(x) = \dfrac{dF_x(x)}{dx}$

$\ast$ $P\{x_1 \leq X \leq x_2\} = \int_{x_1}^{x_2} f_x(x)\,dx$ = Area under the Curve.

CDF

$F_x(x)$

$P(a \leq x \leq b) = F_x(b) - F_x(a)$

discreate

PMF

$P(X = x_i)$

"عند نقطة معينة"

CDF

$P(X \leq x) = F_x(x)$

" النسبة التراكمية "

---

* إذا كان عندي CDF وفيه مجاهيل فبستخدم هدول ⟸ $\begin{cases} F_x(-\infty) = 0 \\ F_x(\infty) = 1 \end{cases}$

+ عشان أقدر أحوله لـ pdf وأقدر أوجد احتمالية كل نقطة بستخدم قانون

$P(X = x) = F(x) - F(x^-)$
$= P(X \leq x) - F(X < x)$

---

mean, expected value, average: μ_x

discreate

$E[g(x)] = \sum_{x=-\infty}^{\infty} g(x)\,P(X = x_i)$

Continuoes

$E[g(x)] = \int_{-\infty}^{\infty} g(x)\,f_x(x)\,dx$

Variance of x: → يوجد قانونين

① $Var[X] = \sigma_x^2$
$= E[(X - \mu_x)^2]$

② $\sigma_x^2 = E[x^2] - \mu_x^2$

Standard deviation of x :

$\sigma_x = \sqrt{\sigma_x^2}$

The median :

$\int_{-\infty}^{0} f_x(x)\,dx = 1/2$, النقطة التي يكون الاحتمال قبلها = بعدها

The mod :

$\dfrac{d f_x(x)}{dx} = 0$ قيمة x التي يكون عندها الاقتران أعلى شي ,

نتنبه إنو قيمة x وليس أعلى قيمة للاقتران

Linear transformation:

$y = aX + b$

$\mu_y = a\mu_x$

$\sigma_y^2 = a^2\sigma_x^2$

ربنا تقبل منا إنك
أنت السميع العليم ..

روان الفار

Ch.2

1. Binomial distribution :-

$$p(X=x) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x} & , x = 0,1,2,\dots \\ 0 & , o.w \end{cases}$$

, n : number of trials.

x : number which success will occure

mean value :- $\mu_x = E(x) = np$

Variance $= \sigma_x^2 = var(x) = np(1-p)$

2. Geometric Distribution :-

$$p(X=x) = \underbrace{(1-p)^{x-1}}_{\text{probability of fail}} \underbrace{p}_{\text{probability of success}} \quad , X=1,2,3,\dots$$

mean value $= \mu_x = E[x] = \frac{1}{p}$

Variance $= \sigma_x^2 = Var(x) = \frac{1-p}{p^2}$

3. Hype geometric distribution :-

- $p(X=x) = \frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}}$
- mean value $= \frac{nK}{N} = np$
- $\sigma_x^2 = np(1-p)\left[\frac{N-n}{N-1}\right]$

4. Poisson distribution

$$p(X=x) = e^{-b}\frac{b^x}{x!} \quad , X=0,1,2,\dots$$

$\mu_x = E[x] = b$, $b = \underbrace{\lambda}_{\text{average of occurences}} \underbrace{T}_{\text{unite}}$

common continuous random variable

1. uniform distribution.

- $f_X(x) = \begin{cases} \frac{1}{b-a} &, a \leq x \leq b \\ 0 &, o.w \end{cases}$, • $\mu_X = \frac{a+b}{2}$, $\sigma_X^2 = \frac{(b-a)^2}{12}$

2. exponential Distribution :-

$f_X(x) = \begin{cases} \lambda e^{-\lambda x} &, x \geq 0 \\ 0 &, o.w \end{cases}$, $\mu_X = \frac{1}{\lambda}$, $\sigma_X^2 = \frac{1}{\lambda^2}$

3. Rayleigh Distribution.

$f_X(x) = \frac{2}{b} x e^{\frac{-x^2}{b}}$, $F_X(x) = 1 - e^{\frac{-x^2}{b}}$

$\mu_X = E[X] = \sqrt{\frac{\pi b}{4}}$

$\sigma_X^2 = \frac{b(4-\pi)}{4}$

4. Cauchy Random Variable

$f_X(x) = \frac{\frac{\alpha}{\pi}}{x^2 + \alpha^2}$, $F_X(x) = \frac{1}{2} + \frac{1}{\pi} \tan^{-1}\left(\frac{x}{a}\right)$

Gaussian normal distribution

$f_X(x) = \frac{1}{\sqrt{2\pi\sigma_X^2}} e^{\frac{-(x-\mu_X)^2}{2\sigma_X^2}}$ $\longrightarrow$ mean = zero , variance = 1

- $P(X \leq 0) = \int_{-\infty}^{0} f_X(x)\, dx = \Phi(0) =$ الجواب من الجدول

- $P(X \leq 1.12) = \int_{-\infty}^{1.12} f_X(x) = \Phi(1.12) =$ من الجدول

- $\Phi(-a) = 1 - \Phi(a)$
- $P(X \geq 3.12) = 1 - P(X \leq 3.12)$
  $= 1 - \Phi(3.12)$
- $P(0.5 \leq X \leq 1.7) = \Phi(1.7) - \Phi(0.5)$

- if $\mu_X \neq 0$ and $\sigma_X \neq 1$ $\rightarrow$ "not standard"

$\Phi\left(\frac{X-\mu_X}{\sigma_X}\right)$

normal approximation for binomial and poisson distribution

- evalute μ_x and σ_x then use gaussian distribution to find the propability.

$$\Phi\left(\frac{x-\mu_x}{\sigma_x}\right)$$

Transformation of Random Variables

discreate

they are equal

X	$P(X=x)$	$Y=g(x)$	$P(Y=y)$

Continueos :

$$f_y(y)=\frac{f_x(x)}{\left|\frac{dy}{dx}\right|}$$

* ① $\frac{dy}{dx}$, ② X = in terms of Y , ③ Substitute in $f_y(y)$ with Cases

---

Note :-

Y is gaussian with mean $\mu_y = a\mu_x + b$ and variance $\sigma_y^2 = a^2\sigma_x^2$

ch.3

- $\mu_x = x \sum_{x=0}^{n} p(X = x_i)$
- $\sigma_x^2 = E[x^2] - \mu_x^2$
- Are x and y indep? $p(X=x, Y=y) \overset{??}{=} p(X=x) \cdot p(Y=y)$ ⇝ check at any point you want.

  $p(A \cap B) = p(A \cap B)$
- if they are indep. $E[xy] = E[x] \cdot E[y]$

correlation coefficient

- $\rho_{xy} = \dfrac{E[xy] - \mu_x \mu_y}{\sigma_x \sigma_y}$, indep $\xrightarrow{\text{must Be}}$ un Carolated
- Covariance $= E[xy] - \mu_x \mu_y$
- $\rho_{xy} = 0 \rightarrow$ unCarolated , $\rho_{xy} = \mp 1 \rightarrow$ fully Carolated.
- $E[xy] = \sum \sum x y \, p(X=x, Y=y)$
- if $z = x + y$ then $E[z] = E[x+y] = E[x] + E[y]$
- if $Y = a_1 x_1 + a_2 x_2$

  ↳ $\mu_y = a_1 \mu_{x_1} + a_2 \mu_{x_2}$

  ↳ $\sigma_y^2 = a_1^2 \sigma_{x_1}^2 + a_2^2 \sigma_{x_2}^2 + 2a_1 a_2 \sigma_{x_1} \sigma_{x_2} \rho_{x_1 x_2}$ ⇝ equals covariance

  ↳ $\sigma_y^2 = a_1^2 \sigma_{x_1}^2 + a_2^2 \sigma_{x_2}^2 + 2a_1 a_2 [E[xy] - \mu_{xy}]$.

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

  Covariance $= \sigma_{x_1} \sigma_{x_2} \rho_{x_1 x_2}$

Two Continuous Random Variable

- to find any value in f_{xy} you should do double integration

  $\int_{x_1}^{x_2} \int_{y_1}^{y_2} f_{xy} \, dy \, dx$ or $\int_{y_1}^{y_2} \int_{x_1}^{x_2} f_{xy} \, dx \, dy$

- $p(X \leq x_0)$ → f_{xy}

  ↳ margine function

  $f_x(x) = \int_{y_1}^{y_2} f_{xy}(x, y) \, dy$ $f_y(y) = \int_{x_1}^{x_2} f_{xy}(x, y) \, dx$

* $f_{xy} = f_x(x) \cdot f_y(y)$

↳ valid if they are independent

Conditional pdf :-

1. $f_{Y/X}(y) = \dfrac{f_{XY}(x,y)}{f_X(x)}$, 2. $f_{X|Y}(x) = \dfrac{f_{XY}(x,y)}{f_Y(y)}$

* $f_{X/Y} = f_X(x)$ and $f_{Y/X} = f_Y(y)$

then they are statistically indep.

* $0 < y < x < 2$
  - $\rightarrow 0 < x < 2$
  - $\rightarrow 0 < y < 2$
  - $\rightarrow y < x$

Ch. 4

- Sample Mean $\hat{\mu}_x = \frac{1}{n}\sum_{i=1}^{n} x_i$
- Sample Variance $\hat{\sigma}_x^2$
  - when μ_x is known $\hat{\sigma}_x^2 = \frac{1}{n}\sum(x_i - \mu_x)^2$
  - when μ_x is unknown $\hat{\sigma}_x^2 = \frac{1}{n-1}\sum(x_i - \hat{\mu}_x)^2$
  - $\hat{\sigma}_x^2 = \dfrac{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}{n(n-1)}$

- Sample Standard diviation :- $\hat{\sigma}_x = \sqrt{\hat{\sigma}_x^2}$
- Sample Covariance between x and y
  - $C_{xy} = \frac{1}{n-1}\sum(x_i - \hat{\mu}_x)(y_i - \hat{\mu}_y)$
  - $C_{xy} = \dfrac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n(n-1)}$

- Sample Correlation Coefficient $\hat{\rho}_{xy} = \dfrac{\hat{\sigma}_{xy}}{\hat{\sigma}_x \hat{\sigma}_y}$ → C_{xy}

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

regression techniques

- $\epsilon = \sum_{i=1}^{n}[Y_i - (\alpha x_i + \beta)]^2$ → least square errors.
- $\alpha = \dfrac{C_{xy}}{\hat{\sigma}_{xy}^2}$, $\beta = \hat{\mu}_y - \alpha\hat{\mu}_x$

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

∴ fitting a polynomial by the method of least squares

$$Y = \beta_1 + \beta_2 x + \beta_3 x^2$$

$$\begin{bmatrix} n & \sum x_i & \sum x_i^2 \\ \sum x_i & \sum x_i^2 & \sum x_i^3 \\ \sum x_i^2 & \sum x_i^3 & \sum x_i^4 \end{bmatrix}\begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum x_i y_i \\ \sum x_i^2 y_i \end{bmatrix}$$

Theorem: independent Gaussian Radom variables

- $E[\hat{\mu}_x] = \mu_x$
- $Var[\hat{\mu}_x] = \dfrac{\sigma_x^2}{n}$

$* \; y = X_1 + X_2 + \dots + X_n$

$\hookrightarrow E[y] = \mu_{x_1} + \mu_{x_2} + \dots + \mu_{x_n} = n\mu_x$

$\hookrightarrow Var[y] = n\sigma_x^2$

$* \; \hat{\mu}_x = y = \frac{1}{n} \Sigma X_i$

- $E[\hat{\mu}_x] = \mu_x$
- $Var[\hat{\mu}_x] = \dfrac{\sigma_x^2}{n}$
- $STD = \dfrac{\sigma_x}{\sqrt{n}}$

Y gaussian

$\hookrightarrow$ standard $\hookrightarrow$ var $=1$, mean $=0$

$\hookrightarrow$ non-standard $\longrightarrow \Phi\left(\dfrac{x - \mu_x}{\sigma_y}\right)$

• استعلاء .. ثمنه التعب

ربنا تقبل منا إنك

أنت السميع العليم ..