Statistique :

• Caractère quantitatif continu :

• Le mode :

• Classe modale : $[l_1, l_2]$

$$M_o = l_1 + \frac{d_1}{d_1 + d_2} \times a_{m[l_1, l_2]}$$

d_1 : l'effectif de la classe modale - l'eff précé

d_2 : " - l'eff suiv

Interpretation :

Le mode est le caractère le plus fréquent

• La médiane :

• Cas discret :

$$Med = \begin{cases} \dfrac{x_{\frac{n}{2}} + x_{\frac{n}{2}+1}}{2} & \dots \text{ n pair} \\ x_{\frac{n+1}{2}} & \dots \text{ n impair} \end{cases}$$

• Cas Continu :

• Classe médiane : $[l_1, l_2]$

(1° classe dont l'effectif cumulé $\geq \frac{n}{2}$)

$$Med = l_1 + \frac{\frac{n}{2} - \tilde{n}_{précé}}{n_{[l_1, l_2]}} \times a_{m[l_1, l_2]}$$

Interprétation :

50% des valeurs sont $\leq$ Med

50% des valeurs sont $\geq$ Med.

• Les quantils :

• Cas discret :

$$Q_\alpha = \begin{cases} \dfrac{x_{n\alpha} + x_{n\alpha+1}}{2} & \dots \ n\alpha \in \mathbb{N} \\ x_{E(n\alpha+1)} & \dots \ n\alpha \notin \mathbb{N} \end{cases}$$

• Cas Continu : $(0 < \alpha < 1)$

• Classe $[l_1, l_2]$: la 1° classe dont l'effectif cumulé $\geq n.\alpha$

$$Q_\alpha = l_1 + \frac{n\alpha - \tilde{n}_{précé}}{n_{[l_1, l_2]}} \times a_{m[l_1, l_2]}$$

• Le graphe

La fonction de répartition ou courbe cumulative

Interpretation :

l'intervalle inter-quartilles $[Q_1, Q_3]$ contient 50% des observations

• La moyenne aréthmetique

• $\overline{X} = \frac{1}{n} \sum_{i=1}^{k} n_i x_i$ (Cas discret)

• $\overline{X} = \frac{1}{n} \sum_{i=1}^{k} n_i C_i$ (Cas Continu)

• La variance

$$V(x) = \frac{1}{n} \sum_{i=1}^{k} n_i (x_i - \overline{X})^2$$

$$\Rightarrow \frac{1}{n} \sum n_i x_i^2 - \overline{X}^2$$

• L'écart type

$$\sigma(X) = \sqrt{(V(x))}$$

Remarque

$Mo = Med = \overline{X}$. Symétrique

$Mo > Med > \overline{X}$ étalement à gauche

$Mo < Med < \overline{X}$. " à droite

• Changement de variable.

$Y = aX + b$

$\Rightarrow \overline{Y} = a\overline{X} + b$

$\Rightarrow V(Y) = a^2 . V(X)$

application :

$$Y = \frac{X - C_i}{a_m}$$ → classe médiane

• moment centré

$$\mu_r = \frac{1}{n} \sum_{i=1}^{k} n_i (x_i - \overline{X})^r$$

$r = 2$ $\mu_2(X) = V(X)$

• Coefficient d'asymétrie

$$\gamma = \frac{\mu_3(X)}{\sigma^3}$$

• Coefficient d'applatissement

$$\beta = \frac{\mu_4(X)}{\sigma^4} - 3$$

• $\gamma = 0 \Rightarrow$ symétrie, $\gamma > 0 \Rightarrow$ à droite

$\gamma < 0 \Rightarrow$ à gauche

• $\beta > 0 \Rightarrow$ faible, $\beta = 0 \Rightarrow$ moyen,

$\beta < 0 \Rightarrow$ fort

Statistique double :

- **Les distributions marginales :**

$n_{11} + n_{12} + n_{13} + \dots + n_{1k} = \sum_{j=1}^{k} n_{1j} = n_{1.}$

$n_{12} + n_{21} + n_{31} + \dots + n_{k1} = \sum_{i=1}^{k} n_{i1} = n_{.1}$

- La colonne $n_{i.}$ = distribution marginale de X
- La colonne $n_{.j}$ = distribution marginale de Y

- **La moyenne :**

$\overline{X} = \frac{1}{n} \sum_i n_{i.} \times x_i$

$\overline{Y} = \frac{1}{n} \sum_i n_{.j} \times y_j$

- **Les fréquences**

$f_{ij} = \frac{n_{ij}}{n}$, $f_{i.} = \frac{n_{i.}}{n}$

$f_{.j} = \frac{n_{.j}}{n}$

- **Les fréquences conditionnelles**

$\frac{n_{ij}}{n_{i.}}$ la fréquence conditionnelle de $Y = y_j$ sachant $X = x_i$ elle est noté :

- $f_{Y=y_j / X=x_i} = \frac{n_{ij}}{n_{i.}} = \frac{f_{ij}}{f_{i.}}$

- $f_{X=x_i / Y=y_j} = \frac{n_{ij}}{n_{.j}} = \frac{f_{ij}}{f_{.j}}$

- **La covariance :**

$Cov(X,Y) = \frac{1}{n} \sum_i \sum_j n_{ij}(x_i - \overline{X})(y_j - \overline{Y})$

$= \frac{1}{n} \sum_i \sum_j n_{ij} x_i y_j - \overline{X} \cdot \overline{Y}$

- **Le coefficient de corrélation linéaire :**

$r(X,Y) = \frac{Cov(X,Y)}{\sigma_x \times \sigma_y}$

• La variance

$$V(X) = \frac{1}{n}\sum_i n_{i.} x_i^2 - \overline{X}^2$$

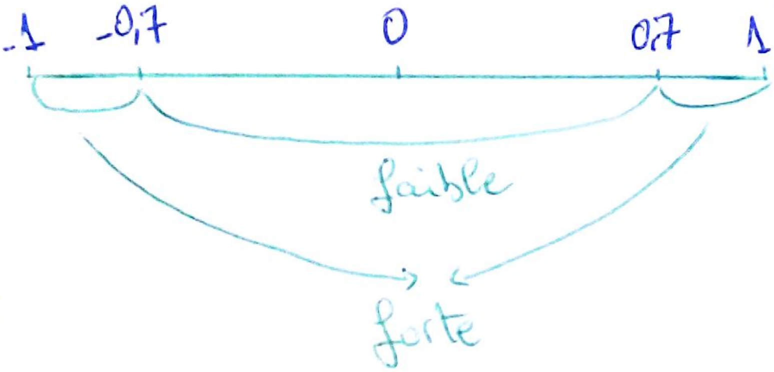
$$V(Y) = \frac{1}{n}\sum_j n_{.j} y_j^2 - \overline{Y}^2$$

• L'écart type :

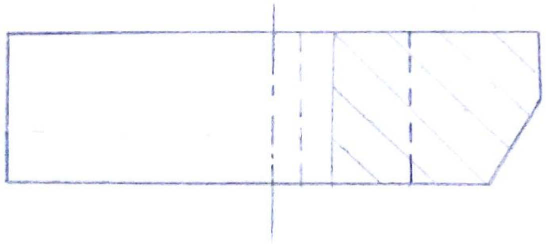
$$\sigma(x) = \sqrt{V(x)} \ , \ \sigma(Y) = \sqrt{V(Y)}$$

N.B :

$$-1 \le r(X,Y) \le 1$$



à approximative du 0,7 et -0,7 on calcule le $r^2(X,Y)$ et on utilise la table de décision



• Droite de regression :

$$Y = aX + b \quad \begin{cases} a = \dfrac{Cov(X,Y)}{V(Y)} \\ b = \overline{Y} - a\overline{X} \end{cases}$$

$$X = \alpha Y + \beta \quad \begin{cases} \alpha = \dfrac{Cov(X,Y)}{V(X)} \\ \beta = \overline{X} - \alpha\overline{Y} \end{cases}$$

Les deux droites passent par le point $G(\overline{X}, \overline{Y})$

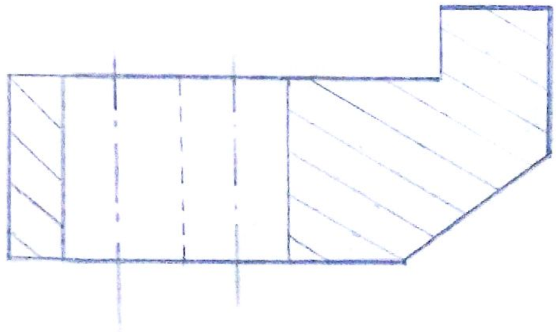
• Changement de variable :

$$Y' = \frac{X - C_i}{am}$$

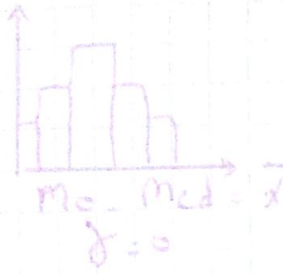
$$\overline{Y'} = a\overline{X} + b$$

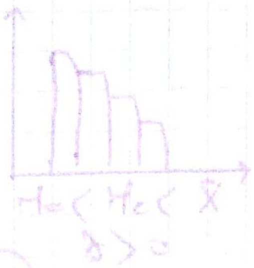
$$V(Y') = a^2 \times V(X)$$

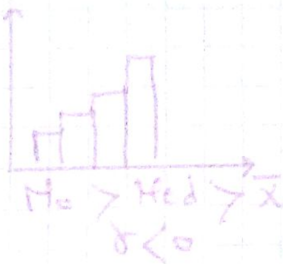
$$Cov(X,Y') = \frac{1}{am} \times Cov(X,Y)$$

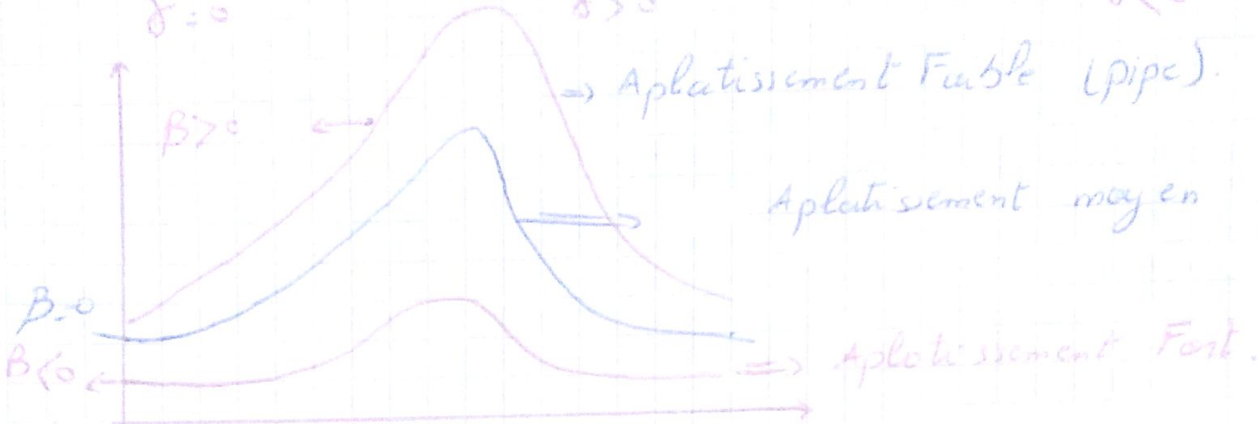


Symétrie | Étalement à droit | étalement à gauche









$$\Rightarrow \mu_3(x) = \frac{1}{n} \sum_{i} n_i (c_i - \bar{x})^3 = 0$$

$$\Rightarrow \gamma = \frac{\mu_3(x)}{\sigma^3} = 0 \quad \text{(symétrique).}$$

$$\mu_4(x) = \frac{1}{n} \sum_{i=1}^{k} n_i (c_i - \bar{x})^4 = \frac{8000}{50} = 160$$

$\Rightarrow$ le Coefficient d'aplatissement

$$\beta = \frac{\mu_4(x)}{\sigma^4} - 3 = \frac{160}{(2,7)^4} - 3 = \frac{160}{(7,36)^2} - 3 = 0,01 .$$

Aplatissement moyen $\simeq 0$.

<u>Methode (2)</u> : changement de variable.

les class c_i		$\tilde{n}_i$	n_i	y_i	$n_i y_i$	$(y_i-\bar{y})^2$	$n_i(y_i-\bar{y})^2$	$n_i(y_i-\bar{y})^3$	$n_i(y_i-\bar{y})^4$
[40,42[	41	2	2	-3	-6	9	18	-54	162
[42,44[	43	7	5	-2	-10	4	20	-40	80
[44,46[	45	15	8	-1	-8	1	8	-8	
[46,48[	47	35	20	0	0	0	0	0	0
[48,50[	49	43	8	1	8	1	8	8	8
[50,52[	51	48	5	2	10	4	20	40	80
[52,54[	53	50	2	3	6	9	18	54	162
total		x	50	x	0	x	92	0	500

• La variance :

$$V(X) = \frac{1}{n} \sum_i n_{i.}\, x_i^2 - \bar{X}^2$$

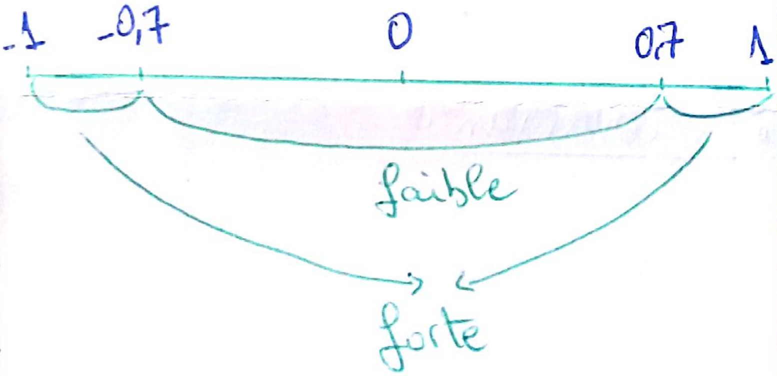
$$V(Y) = \frac{1}{n} \sum_j n_{.j}\, y_j^2 - \bar{Y}^2$$

• L'écart type :

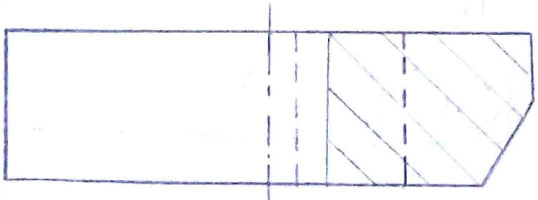
$$\sigma(x) = \sqrt{V(x)} \;,\; \sigma(Y) = \sqrt{V(Y)}$$

N.B :

$$-1 \leq r(X,Y) \leq 1$$



à approximative du 0,7 et -0,7 on calcule le $r^2(x,Y)$ et on utilise du table de décision



La dépendance entre X et Y :

$f_{ij} = f_{i.} \times f_{.j}$ $\forall i$ $\forall j$ dépendants

$f_{ij} \neq f_{i.} \times f_{.j}$ $\forall$ indépendants

• Droite de regression :

$Y = aX + b$ (Y en X)

$$\begin{cases} a = \dfrac{Cov(X,Y)}{V(Y)} \\ b = \bar{Y} - a\bar{X} \end{cases}$$

$X = \alpha Y + \beta$ (X en Y)

$$\begin{cases} \alpha = \dfrac{Cov(X,Y)}{V(X)} \\ \beta = \bar{X} - \alpha\bar{Y} \end{cases}$$

Les deux droites passent par le point $G(\bar{X}, \bar{Y})$

• Changement de variable

$$Y' = \frac{X - C_i}{a_m}$$

$$\bar{Y}' = a\bar{X} + b$$

$$V(Y') = a^2 \times V(X)$$

$$Cov(X,Y') = \frac{1}{a_m} \times Cov(X,Y)$$