Série statistique à 1 variable (x): discrète

I – Les données (fixées et déduites)

- Tableau d'effectifs (x_i ; n_i), de fréquence f_i ou $f(x)$, de fréquences cumulées croissantes FCC ou $F(x)$

x_i	4	6	7	9	11	13	16
n_i	1	2	2	3	2	2	1
f_i	0,08	0,15	0,15	0,23	0,15	0,15	0,08
FCC	0,08	0,23	0,38	0,62	0,77	0,92	1

- Paramètres (discrète ou continue) :

Effectif total : $N = \Sigma\, n_i$; fréquence: $f_i = n_i / N$ et donc $\Sigma f_i = 1$

Somme des x : $\Sigma\, x = \Sigma\, n_i \times x_i$ et des carrés : $\Sigma\, x^2 = \Sigma\, n_i \times x_i^2$

Position: Moyenne ($\bar{x}$ ou μ): $\mu = (1/N) \times \Sigma\, n_i \times x_i = \Sigma f_i \times x_i$

Variance (V): $V = (1/N) \times \{ \Sigma\, n_i \times (x_i - \mu)^2 \} = \Sigma f_i \times x_i^2 - \mu^2$

Dispersion: Écart-type (s ou σ): $s = \sqrt{V} \Leftrightarrow s^2 = V$

Position: Médiane (Méd ou Q_2): $F(Q_2) = ½$

Quartiles (Q_1 et Q_3): $F(Q_1) = ¼$ et $F(Q_3) = ¾$

Dispersion: intervalle inter-quartile: $Q_3 - Q_1$

Mode: Valeur de x pour laquelle n ou f est maximale

- Pour l'exemple: $N = 13$; $f_{max} = 0,23$; $\bar{x} \approx 9,31$; $s \approx 3,27$; $Q_1 = 7$; $Q_2 = 9$; $Q_3 = 11$; Mode = 9

II – Représentation graphique

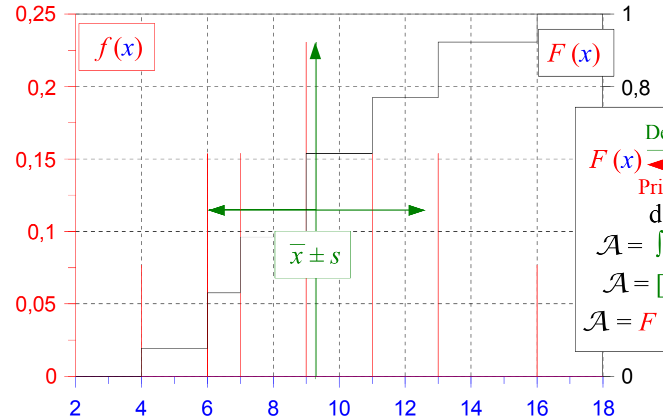
- $f(x)$: Distribution des (x) ou densité de probabilité:

A chaque valeur de x, un bâton de hauteur f associée.

- $F(x)$: Fonction de répartition:

A chaque valeur de x, une marche d'escalier de hauteur f.

- Moyenne ($\bar{x}$ ou μ): à $x = \mu$, une flèche de hauteur f_{max}
- Écart-type (s ou σ): à ½ f_{max} , flèche de largeur ± s.



Série statistique à 1 variable (x): continue

I – Les données (fixées et déduites)

- Tableau d'intervalles ou de classes ([x_{inf} ; x_{sup} [; n_i), de centre d'intervalle x_i, de fréquence f_i, de largeur d'intervalle l_i, de distribution d ou $f(x)$ telle que $d_i = f_i \div l_i$ et de fréquences cumulées croissantes FCC ou $F(x)$

Classe	[7;9[	[9;10[	[10;11[	[11;12[	[12;13[	[13;14[	[14;16[
n_i	40	43	63	72	47	32	27
x_i	8	9,5	10,5	11,5	12,5	13,5	15
f_i	0,12	0,13	0,19	0,22	0,15	0,1	0,08
l_i	2	1	1	1	1	1	2
$d_i = f_i / l_i$	0,06	0,13	0,19	0,22	0,15	0,1	0,04
FCC	0,12	0,26	0,45	0,67	0,82	0,92	1

- Paramètres (discrète ou continue): **formules identiques**

a) x est le **centre** de l'intervalle [x_{inf} ; x_{sup} [pour $\bar{x}$ et s

b) Les quartiles se déterminent graphiquement (les valeurs renvoyées par la calculatrice sont légèrement erronées)

- Ici: $N = 324$; $d_{max} = 0,22$; $\bar{x} \approx 11,24$; $s \approx 1,93$; $Q_1 \approx 9,95$; $Q_2 \approx 11,25$; $Q_3 \approx 12,5$; Mode = [11;12[

II – Représentation graphique

- $f(x)$: Densité de probabilité : Histogramme.

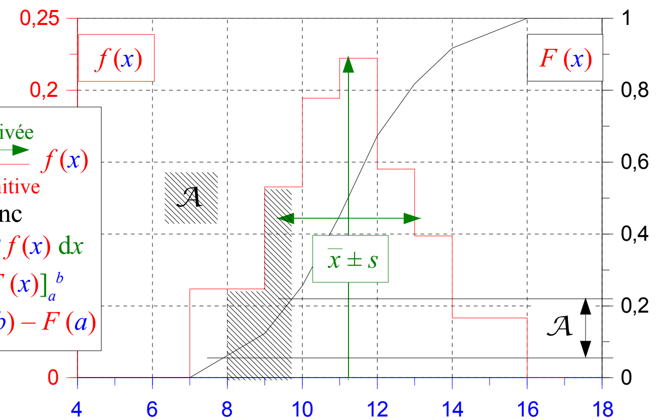
Sur chaque intervalle, un rectangle de hauteur d_i associée et donc par définition d'aire ou d'intégrale $d_i \times l_i = f_i$ (en u.a.)

- $F(x)$: Fonction de répartition : Affine par morceaux

A chaque intervalle, un segment de largeur l_i et de hauteur f_i donc de pente (ou de dérivée) $f_i / l_i = d_i$. donc $F'(x) = f(x)$

- Moyenne ($\bar{x}$ ou μ) et écart-type (s ou σ): Prendre d_{max}

(Auteur : S. Basnary – version 2013)



Série statistique à 2 variables (x ; y)

I – Les données (fixées et déduites)

- Tableau de **coordonnées** des points M_i (x_i ; y_i).

x_i	5	7	8	9	11	12	13	15
y_i	5,9	8,09	8,16	8,43	11,03	12,04	14,81	12,58
n_i	1	1	1	1	1	1	1	1

- Paramètres : (n_i peut être ≠ 1) Effectif total $N = \Sigma\, n$

Moyenne(s): $\bar{x}$ et $\bar{y}$; écart(s)-type(s): σ_x et σ_y

Covariance: $\mathrm{Cov}(x ; y) = (1/N) \times \Sigma\, n_i \times (x_i - \bar{x}) \times (y_i - \bar{y})$

$\mathrm{Cov}(x ; y) = (1/N) \times \{ \Sigma\, n_i \times x_i \times y_i \} - \bar{x} \times \bar{y}$

- Ici: $\bar{x} \approx 10$; $\sigma_x \approx 3,12$; $\bar{y} \approx 10,13$; $\sigma_y \approx 2,76$; Cov ≈ 7,99

II – Ajustement affine (D) $y = a\,x + b$ des points M_i

- Nuage de points M_i (x_i ; y_i) et point moyen G ($\bar{x}$; $\bar{y}$)
- (D): Droite d'ajustement ou droite des moindres carrés.

a) La droite (D) passe par le G ($\bar{x}$; $\bar{y}$) donc $b = \bar{y} - a\,\bar{x}$

b) A chaque M_i , on calcule la distance $d_i = y_i - (a\,x_i + b)$

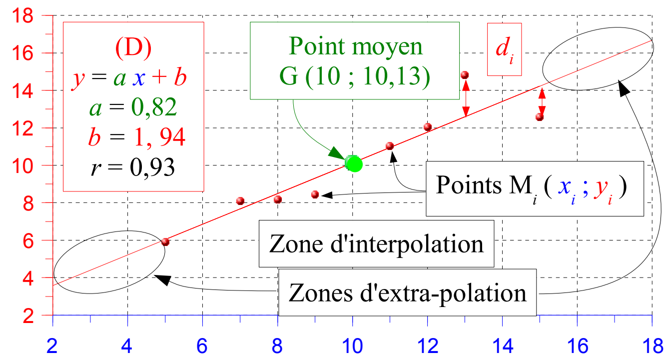
c) La droite (D) minimise la distance $D = \Sigma\, d_i^2$

d) On déduit : **(valeurs obtenues à la calculatrice)**

$a = \mathrm{Cov}(x ; y) / V_x$; $b = \bar{y} - a\,\bar{x}$ et $r = \mathrm{Cov}(x ; y) / (\sigma_x \times \sigma_y)$

r : coefficient de corrélation (même signe que la pente a)

Si $|r| \approx 1 \Leftrightarrow$ ajustement affine justifié ⇔ Points M_i alignés



III – Autres ajustements qui mènent à un ajustement affine