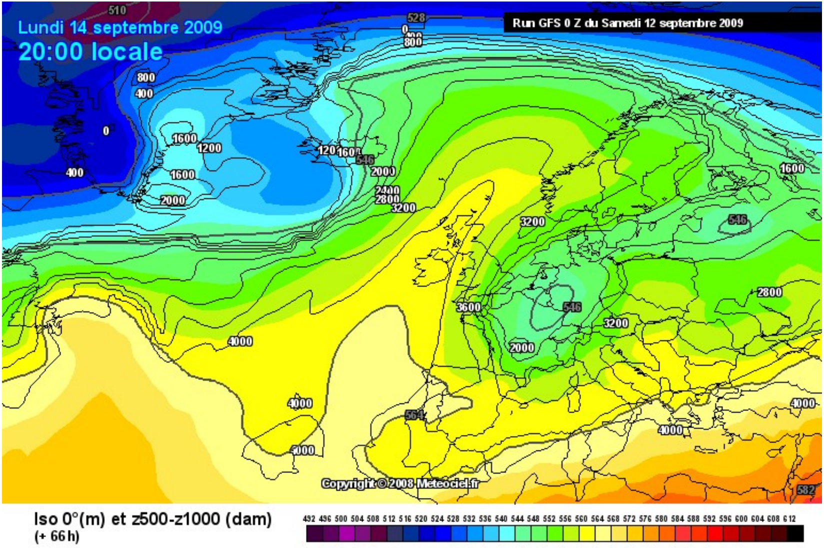
Statistique

MatheX

19 août 2020


Lundi 14 septembre 2009
20:00 locale
Run GFS 0 Z du Samedi 12 septembre 2009
Copyright © 2008 Meteociel.fr
Iso 0°(m) et z500-z1000 (dam)
(+ 66 h)
492 496 500 504 508 512 516 520 524 528 532 536 540 544 548 552 556 560 564 568 572 576 580 584 588 592 596 600 604 608 612

1. Vocabulaire

Définition 1 : (série statistique)

Une série statistique est la liste des valeurs (x_i) d'une caractéristique prise par les individus d'une population.

L' étendue est la différence entre la plus grande valeur et la plus petite.

L' effectif total (N) est le nombre total de valeurs dans la série.

Exemple :

Le nombre de frères et de soeurs des élèves de la classe :

2 0 0 0 6 0 0 1 1 1 6 1 2 2 2 2 2 2 1 2 2 3 3 3 3 3 4 4 4 2

Effectif total : $N = 30$

Étendue : $6 - 0 = 6$

Définition 2 : (effectif, fréquence)

Dans une série statistique pondérée, l'effectif (n_i) d'une valeur (x_i) (d'une classe) est le nombre de fois où l'on a cette valeur (cette classe)

La fréquence (f_i) d'une valeur x_i est le quotient de l'effectif (n_i) par l'effectif total (N) :

$$\boxed{f_i = \frac{n_i}{N}}$$

Exemple :

◦ Le nombre de frères et de soeurs des élèves de la classe :

Nombre (x_i)	0	1	2	3	4	5	6
Effectif (n_i)	5	5	10	5	3	0	2

La fréquence des élèves ayant 2 frères et soeurs :

$$f_3 = \frac{n_3}{N} = \frac{10}{30} = \frac{1}{3}$$

◦ La taille des élèves de la classe :

Taille en m (x_i)	[1,2 ;1,4[	[1,4 ;1,6[	[1,6 ;1,8[	[1,8 ;2[
Effectif (n_i)	1	20	7	2

2. Position et dispersion

Définition 3 : (moyenne pondérée)

La moyenne $\overline{x}$ d'une série statistique pondérée x est :

$$\boxed{\overline{x} = \frac{1}{N}\sum_{i=1}^{n} n_i\, x_i}$$

$$= \sum_{i=1}^{n} f_i\, x_i$$

$$= \frac{1}{N}\Big(n_1 \times x_1 + n_2 \times x_2 + \cdots + n_n \times x_n\Big)$$

Exemple :

∘ Le nombre de frères et de soeurs des élèves de la classe :

Nombre (x_i)	0	1	2	3	4	5	6
Effectif (n_i)	5	5	10	5	3	0	2

$$\overline{x} = \frac{1}{30}\left(0\times 5 + 1\times 5 + 2\times 10 + 3\times 5 + 4\times 3 + 5\times 0 + 6\times 2\right) = \frac{64}{30} \simeq 2,13$$

∘ La taille des élèves de la classe :

Taille en m (x_i)	[1,2;1,4[	[1,4;1,6[	[1,6;1,8[	[1,8;2[
Effectif (n_i)	1	20	7	2

$$\overline{x} = \frac{1}{30}\left(1\times 1,3 + 20\times 1,5 + 7\times 1,7 + 2\times 1,9\right) \simeq 1,57$$

Propriété 1 : (linéarité de la moyenne)

Soit a et b des nombres réels.
Soit x une série statistique de moyenne $\bar{x}$

Si une autre série statistique y est linéaire avec x alors leur moyenne suit la même linéarité :

$$y = ax + b \Longrightarrow \boxed{\bar{y} = a\bar{x} + b}$$

Exemple :

La taille des élèves de la classe :

Taille en m (x_i)	[1,2;1,4[	[1,4;1,6[	[1,6;1,8[	[1,8;2[
Effectif (n_i)	1	20	7	2

∘ Potion magique $\longrightarrow$ les élèves doublent de taille

Avec y la nouvelle taille et x l'ancienne taille, on a :
$y = 2x \Longrightarrow \overline{y} = 2\overline{x}$ et donc $\overline{y} \simeq 2 \times 1,57 = 3,14$

∘ Les élèves vont dans l'espace $\longrightarrow$ ils gagnent 5 cm

Avec y la nouvelle taille et x l'ancienne taille, on a :
$y = x + 0,05 \Longrightarrow \overline{y} = \overline{x} + 0,05$ et donc $\overline{y} \simeq 1,57 + 0,05 = 1,62$

Définition 4 : (variance, écart type)

La variance V d'une série statistique x est le nombre :

$$\boxed{V = \frac{1}{N}\sum_i n_i(x_i - \overline{x})^2}$$

$$= \frac{1}{N}\left(n_1(x_1 - \overline{x})^2 + n_2(x_2 - \overline{x})^2 + \cdots + n_n(x_n - \overline{x})^2\right)$$

L' écart-type $\sigma(X)$ d'une série statistique x est le nombre :

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

NB :
Ces indicateurs mesurent la dispersion autour de la moyenne.

Exemple :

Le nombre de frères et de soeurs des élèves de la classe :

Nombre (x_i)	0	1	2	3	4	5	6
Effectif (n_i)	5	5	10	5	3	0	2

$$\begin{aligned} V &= \frac{1}{N}\sum_i n_i(x_i-\overline{x})^2 \\ &= \frac{1}{30}\Big(5\times(0-\tfrac{64}{30})^2+5\times(1-\tfrac{64}{30})^2+10\times(2-\tfrac{64}{30})^2+5\times(3-\tfrac{64}{30})^2+3\times \end{aligned}$$
$$(4-\tfrac{64}{30})^2+0\times(5-\tfrac{64}{30})^2+2\times(6-\tfrac{64}{30})^2\Big)\simeq 2,45$$

$$\begin{aligned} \sigma &= \sqrt{V} \\ &\simeq \sqrt{2,45}\simeq 1,56 \end{aligned}$$

Définition 5 : (médiane)

Une médiane (M) d'une série statistique ordonnée est **une** valeur qui la sépare en deux sous séries de même effectif.

En pratique :

Effectif total (N) impair : $\boxed{M = x_{\frac{N+1}{2}}}$

Effectif total (N) pair : $\boxed{M = \dfrac{x_{\frac{N}{2}} + x_{\frac{N}{2}+1}}{2}}$

Exemple :

<u>N pair :</u>

i	x_i	
1	10	
2	12	
		$\rightarrow M = \frac{x_2+x_3}{2} = 13$
3	14	
4	20	

<u>N impair :</u>

i	x_i	
1	10	
2	12	
3	14	$\rightarrow M = x_3 = 14$
4	20	
5	20	

Définition 6 : (quartiles)

Le premier quartile (Q_1) d'une série statistique est **la** plus petite valeur de la série telle qu'au moins 25% des valeurs de la série soit inférieures ou égales à Q_1.

Le troisième quartile (Q_3) d'une série statistique est **la** plus petite valeur de la série telle qu'au moins 75% des valeurs de la série soit inférieures ou égales à Q_3.

L' écart interquartile est égale à $Q_3 - Q_1$.
C'est un indicateur de dispersion.

Exemple :

N pair :

i	x_i	
1	10	
2	12	$\rightarrow Q_1 = x_2 = 12$
3	12	
4	14	
		$\rightarrow M = \frac{x_4+x_5}{2} = 15$
5	16	
6	20	$\rightarrow Q_3 = x_6 = 20$
7	20	
8	20	

$Q_3 - Q_1 = 20 - 12 = 8$

N impair :

i	x_i	
1	10	
2	12	
3	14	$\rightarrow Q_1 = x_3 = 14$
4	14	
5	14	$\rightarrow M = x_5 = 14$
6	16	
7	18	$\rightarrow Q_3 = x_7 = 18$
8	19	
9	20	

$Q_3 - Q_1 = 18 - 14 = 4$