



## Module 4/5

# Outils statistiques





**Définition**

**Paramètres de position**

**Paramètres de dispersion**

**incertitude**

**échantillonnage**

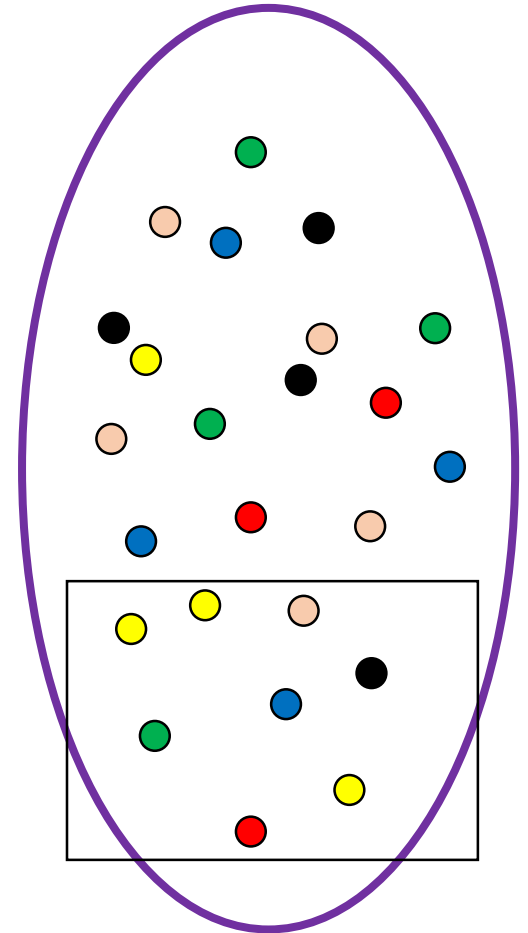


## population

$N$  : ensemble des éléments du jeu de données

## Échantillon

$n$  : sous-ensemble d'éléments prélevés



# Paramètre de position (dans un ensemble de valeurs)



## Moyenne arithmétique

$\bar{X}$  moyenne pour un échantillon **n**

$\mu$  moyenne pour une population **N**

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

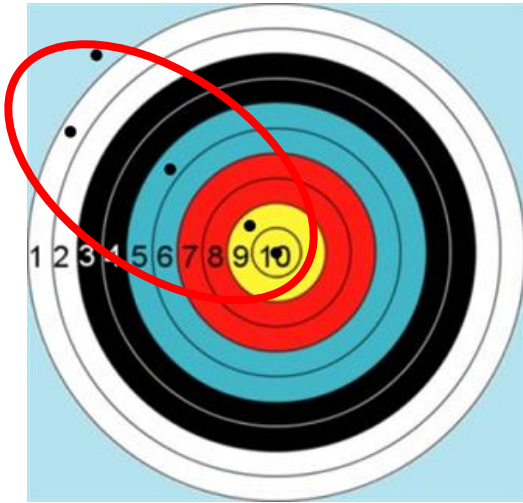
$$\mu = \frac{\sum_{i=1}^N X_i}{N}$$

Exemple :

moyenne des nombres 1,2 ; 1,0 ; 1,4 ; 1,4 , 1,3 sur un échantillon

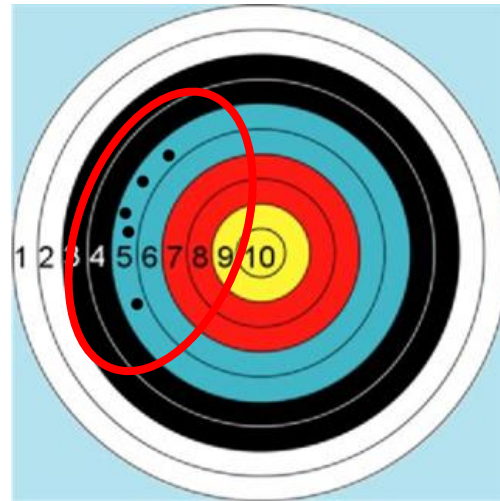
$$\bar{X} = \frac{1,2 + 1 + 1,4 + 1,4 + 1,3}{5} = 1,26$$

# Paramètre de dispersion



$$\bar{x} = 5$$

étendue



$$\bar{x} = 5$$

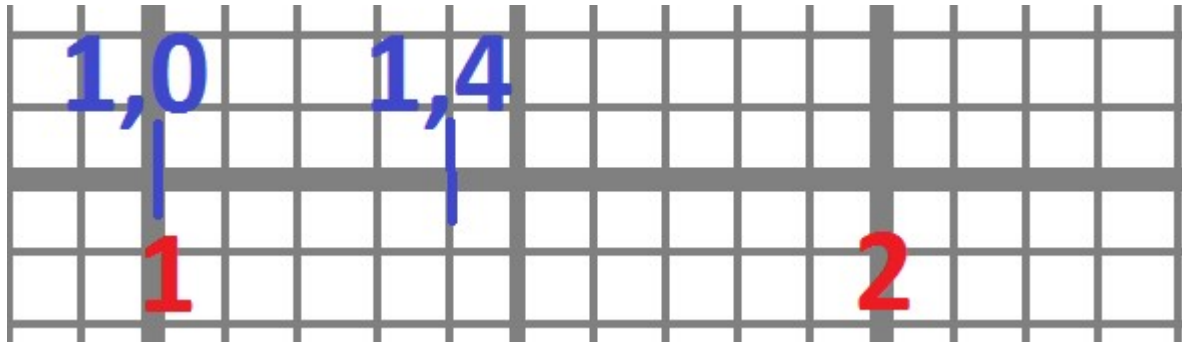
variance

Écart-type

# Etendue



Étendue de l'échantillon : 1,2 ; 1,0 ; 1,4 ; 1,4 ; 1,3



$$\text{Etendue} = 1,4 - 1,0 = 0,4$$

# Variance



$$v = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}$$

**v** : variance

**n** : nombre d'échantillons

**n-1** : on enlève 1 à l'effectif afin de corriger une sous estimation de l'hétérogénéité du fait de l'échantillonnage

Exemple :

**Variance** de l'échantillon 1,2 ; 1,0 ; 1,4 ; 1,4 , 1,3

$$v = \frac{[(1,2 - 1,26)^2 + (1,0 - 1,26)^2 + (1,4 - 1,26)^2 + (1,4 - 1,26)^2 + (1,3 - 1,26)^2]}{5-1} = 0,028$$



$$s = \sqrt{v}$$

**s** : écart-type

**v** : variance

**n** : nombre d'échantillons

Exemple :

Ecarts de l'échantillon 1,2 ; 1,0 ; 1,4 ; 1,4 , 1,3

$$s = \sqrt{v} = \sqrt{0,028} = 0.167$$





## définition

**Incertitude sur une moyenne d'un échantillon**

**Incertitude sur un pourcentage**



$$\Delta L = 0,5\text{m}$$

$$L = 12\text{m} \pm 0,5\text{m}$$

**Niveau de confiance**

**95% de chance que  $12\text{m} - 0,5\text{m} \leq \text{mesure} \leq 12\text{m} + 0,5\text{m}$**

# Incertitude sur une moyenne d'un échantillon



$$\Delta\bar{X} = t \frac{s}{\sqrt{n}}$$

$$X = \bar{X} \pm \Delta\bar{X}$$

**s** : écart-type

**n** : taille de l'échantillons

**t** : coefficient de Student

n	2	3	4	5	6	7	8	9	10
t (95%)	12,7	4,3	3,18	2,78	2,57	2,45	2,37	2,31	2,26
t (99%)	63,7	9,93	5,84	4,6	4,03	3,71	3,5	3,36	3,25
n	12	14	16	18	20	30	50	100	∞
t (95%)	2,2	2,16	2,13	2,11	2,09	2,04	2,01	1,98	1,96
t (99%)	3,11	3,01	2,95	2,9	2,86	2,76	2,68	2,63	2,57

# Incertitude sur un pourcentage



n	2	3	4	5	6	7	8	9	10
t (95%)	12,7	4,3	3,18	2,78	2,57	2,45	2,37	2,31	2,26
t (99%)	63,7	9,93	5,84	4,6	4,03	3,71	3,5	3,36	3,25
n	12	14	16	18	20	30	50	100	∞
t (95%)	2,2	2,16	2,13	2,11	2,09	2,04	2,01	1,98	1,96
t (99%)	3,11	3,01	2,95	2,9	2,86	2,76	2,68	2,63	2,57

Exemple :

Taux d'exhaustivité :  $e = 92\%$

Nombre d'échantillons : 50

Niveau de confiance : 95%

Coefficient de Student : 2,04

$$\Delta \bar{X} = t \frac{s}{\sqrt{n}}$$

$$\Delta P = 2.04 \times \frac{\sqrt{0,92 \times (1 - 0,92)}}{50} = 0,08 = 8\%$$



**Évaluation des éléments conformes/non conformes**

**Évaluation d'une grandeur mesurable**

# Evaluation des éléments conformes/non conformes



Niveau de confiance = 95 %			$p_0 = \text{LAQ}$ (limite d'acceptation de la qualité)					
Taille de la population (N)		Taille de l'échantillon	0,5 %	1,0 %	2,0 %	3,0 %	4,0 %	5,0 %
De	à	(n)	Limite de rejet					
1	8	Toutes	1	1	1	1	1	1
9	50	8	1	1	1	2	2	2
51	90	13	1	1	2	2	2	3
91	150	20	1	2	2	3	3	4
151	280	32	1	2	3	3	4	4
281	400	50	2	3	3	4	5	6
401	500	60	2	3	4	5	6	7
501	1200	80	3	3	5	6	7	8
1 201	3 200	125	3	4	6	8	10	11
3 201	10 000	200	4	6	8	11	14	16
10 001	13 000	315	5	7	12	16	20	23
13 001	15 000	500	6	10	16	23	28	34
15 001	50 000	800	9	14	24	33	42	51
> 50 000		1250	12	20	34	49	63	76

# Evaluation des éléments conformes/non conformes

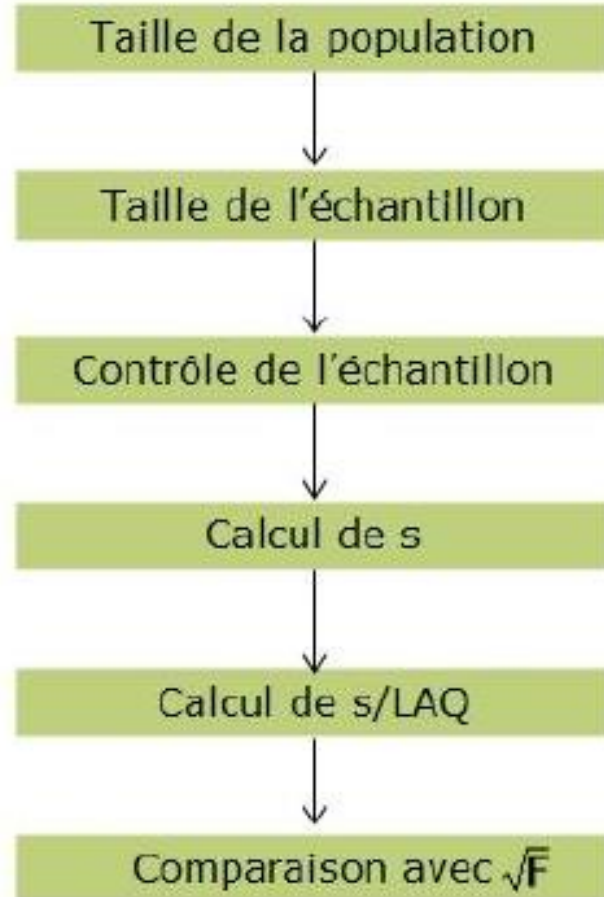


Nombre total : **2440 Bâtiments**      LAQ : **0,5%**

taille de l'échantillon: **125**      Limite de rejet: **3 bâtiments manquants**

Niveau de confiance = 95 %			$p_0 =$ LAQ (limite d'acceptation de la qualité)					
Taille de la population (N)		Taille de l'échantillon (n)	0,5%	1,0%	2,0%	3,0%	4,0%	5,0%
De	à		Limite de rejet					
1	8	Toutes	1	1	1	1	1	1
9	50	8	1	1	1	2	2	2
51	90	13	1	1	2	2	2	3
91	150	20	1	2	2	3	3	4
151	280	32	1	2	3	3	4	4
281	400	50	2	3	3	4	5	6
401	500	60	2	3	4	5	6	7
501	1200	80	3	3	5	6	7	8
1 201	3200	125	3	4	6	8	10	11
3 201	10 000	200	4	6	8	11	14	16
10 001	13 000	315	5	7	12	16	20	23
13 001	15 000	500	6	10	16	23	28	34
15 001	50 000	800	9	14	24	33	42	51
> 50 000		1250	12	20	34	49	63	76

# Évaluation d'une grandeur mesurable



Taille de la population (N)		Taille de l'échantillon (n)	$\sqrt{F}$
De	à		
26	50	5	1,54
51	90	7	1,45
91	150	10	1,37
151	280	15	1,3
281	400	20	1,26
401	500	25	1,23
501	1 200	35	1,2
1 201	3 200	50	1,16
3 201	10 000	75	1,13
10 001	35 000	100	1,12
35 001	150 000	150	1,09
150 001	500 000	200	1,08
> 500 000		200	1,08

Tableau de *Fisher* pour un niveau de signification de 95%



# Exemple



Taille de la population (N)		Taille de l'échantillon (n)	$\sqrt{F}$
De	à		
26	50	5	1,54
51	90	7	1,45
91	150	10	1,37
151	280	15	1,3
281	400	20	1,26
401	500	25	1,23
501	1 200	35	1,2
1 201	3 200	50	1,16
3 201	10 000	75	1,13
10 001	35 000	100	1,1
35 001	150 000	150	1,09
150 001	500 000	200	1,08
> 500 000		200	1,08

**REJECT**

N = 800 objets

**n = 35 objets**

LAQ = 10 cm

s = 17 cm

s/LAQ = 1,7

**s/LAQ > 1,2**



Notion de statistiques

moyenne

Écart type

incertitude

échantillonnage