

## Méthodes statistiques TD n°2

O. Dadoun [dadoun@in2p3.fr](mailto:dadoun@in2p3.fr) semaine du 21/09/2015

### Rappel de quelques définitions de statistiques

- L'ensemble sur lequel on travaille en statistique est appelé population. Si cet ensemble est trop vaste, on en restreint l'étude à une partie appelée échantillon. Un élément de cet ensemble est appelé individu.

- La particularité commune que l'on étudie est appelée caractère. L'effectif d'une "valeur" d'un caractère est le nombre d'individus ayant cette valeur. Les valeurs prises par le caractère sont aussi appelées les modalités.

### Exercice I (utilisation de $\Sigma$ )

1. Calculer les sommes suivantes:

$$\sum_{k=1}^{k=4} (2k - 1) = 16 \text{ et } \sum_{i=3}^{i=8} i = 33$$

2. Exprimer les deux sommes suivantes en utilisant le symbole  $\Sigma$

$$A = 1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{8} = \sum_{i=1}^{i=8} 1/i$$

$$B = 1 + 4 + 16 + \dots + (n + 1)^2 = \sum_{k=0}^{k=n} k^2$$

3. Réécrivez les expressions suivantes en utilisant le signe sigma:

a.  $n_{10} + n_{11} + n_{12} + n_{13} + n_{14} + n_{15} = \sum_{i=10}^{i=15} n_i$

b.  $(ax_0 - \bar{x}) + (ax_1 - \bar{x}) + (ax_2 - \bar{x}) + (ax_3 - \bar{x}) + (ax_4 - \bar{x}) + (ax_5 - \bar{x}) = \sum_{i=0}^{i=5} (ax_i - \bar{x})$

### Paramètres de position : la moyenne

Deux formules possibles:

a. Sur les données brutes :  $\bar{X} = \frac{1}{N} \sum_{i=1}^N x_i$

b. Sur les données regroupées par modalité :  $\bar{X} = \frac{1}{N} \sum_{i=1}^m n_i x_i$

### Exercice 2

1. Un professeur assure l'enseignement de math. dans une classe de 6ème. Lors d'un contrôle effectué en fin de trimestre les notes obtenues pour chaque élève sont:  
Classe A : 9+10+11+19+12+10+10+15+11+3+8+10+12+12+9+10+12+9+11+15  
Calculer la moyenne avec les deux formules :  
 $125/20 = (3*9+5*10+3*11+19+2*15+3+8+4*12) / 20 = 10.9$
2. La moyenne des notes de français de dix élèves est de 10.8. La note d'un élève a été effacée mais on distingue les 9 autres notes:  
8, 8, 7, 10, 10, 10, 17, 16, 11 et 11

### Exercice 3

On a filtré selon l'âge les données d'un fichier regroupant des informations relevées sur un échantillon d'enfants de 4 à 7 ans. Des données\* concernant les enfants de sexe masculin âgés de 5ans (60 mois) ont été extraites et constituent le tableau ci-dessous

Num.	Genre	Âge	Tailles	Poids
51	M	60	112	20
903			105.5	18
1079			105.5	17
1088			112	20
1101			110.5	20
1129			105.5	17
1138			112	20
1963			108	18
2339			112	21
2388			106	19
2463			108	20
2470			114	22.5
2471			107	17.5
2486			108	20

- chacune des lignes du tableau correspond à un enfant (= sujet = individu = unité statistique)
- chacune des colonnes correspond à une caractéristique (= variable) observée

A) A partir de ces données, établissez :

- dans le tableau 1 ci-dessous, la série des tailles triées (par ordre croissant) avec répétition,
- dans le tableau 2, la distribution des tailles triées sans répétition = distribution ordonnée

B) Tableau I : série des valeurs triées

- Donnez la valeur de n : effectif du groupe **n=14**
- Calculez l'étendue de la distribution ( $x_{max} - x_{min}$ ) donnez la signification concrète du résultat obtenu.  
**Il y a 8.5 cm de différence entre la taille du plus petit et du plus grand de ces enfants de 5 ans**
- Calculez la somme des tailles de la totalité des sujets du groupe  $\sum x_i = 1526$
- Calculez la taille moyenne des sujets de ce groupe **114 cm**

<i>i</i> (rang)	$x_i$ (taille)
1	105.5
2	105.5
3	105.5
4	106



5	107
6	108
7	108
8	108
9	110.5
10	112
11	112
12	112
13	112
14	114

$i$ (rang)	$x_i$ (valeurs de la variable taille)	$n_i$ (effectifs)	$n_i x_i$ (produits des valeurs par leur effectif)	$x_i - \bar{x}$ (écarts des valeurs à la moyenne)	$n_i(x_i - \bar{x})$ (produits des écarts par leur effectif)
1	105.5	3	316.5	-3.5	-10.5
2	106	1	106	-3.0	-3.0
3	107	1	107	-2.0	-2.0
4	108	3	324	-1.0	-3.0
5	110.5	1	110.5	1.5	1.5
6	112	4	448	3.0	12.
7	114	1	114	5.	5.

Tableau 2 : distribution des valeurs ordonnées

- Vérifiez que toutes les observations ont bien été comptabilisées dans le tableau de la distribution ordonnée
- Calculez  $\sum_{i=5}^{i=7} n_i$  ; écrivez la forme développée de cette expression ; dites ce que le résultat représente concrètement **6 enfants ont une taille > 110.5 cm**
- Calculez la somme des tailles de la totalité des sujets du groupe  $\sum n_i x_i = 1526$
- Comparez méthode et résultat avec ceux du calcul effectué pour les valeurs triées.
- Calculez la taille moyenne des sujets de ce groupe  $\bar{x} = 109$
- Complétez le tableau en introduisant les colonnes de calcul des écarts à la moyenne et des produits
- Calculez la somme des écarts des tailles à la moyenne

**La somme des écarts à la moyenne est tjs nulle.**

- Quelle conclusion tirez-vous du résultat ?

**La somme des écarts à la moyenne négatifs est égale à la somme des écarts positifs: la moyenne est le centre d'équilibre de la distribution**

**Exercice 4** (Ref. <http://oorraannqgee.free.fr/>)

On a consigné les primes de fin d'année attribuées aux salariés d'une entreprise dans le tableau suivant

Primes (centaines d'euros)	[ 0 ; 6 [	[ 6 ; 10 [	[ 10 ; 14 [	[ 14 ; 16 [
Effectifs	41	79	78	2
Milieux des classes	3	8	12	15

1. Quelle est la population étudiée ?  
*La population étudiée est l'ensemble des salariés de l'entreprise*
2. Quel est le caractère étudié ?  
*Il s'agit du montant de la prime de fin d'année*
3. Quelle est la nature de ce caractère ?  
*Qualitatif continue*
4. Pourquoi a-t-on regroupé les primes en classes ?  
*Le caractère étant continue, les modalités (valeurs prises par le caractères) sont regroupées en intervalle (les classes)*
5. Déterminer la moyenne ? on fait d'abord le milieu des classes puis:  
$$\bar{x} = (41 * 3 + 79 * 8 + 78 * 12 + 2 * 15) / (41 + 79 + 78 + 2) = 8.6 \text{ centaines d'euros}$$