1 Langage des statistiques descriptives

Exemple 1. Résultats au dernier contrôle d'une classe de 22 élèves.

< données de l'exercice 17 p. 257 sous forme de fonction>

On considère un ensemble à N éléments appelé la population, et une fonction appelée variable statistique de cet ensemble dans un ensemble fini de valeurs (numériques ou pas) x_1, x_2, ..., x_p :

< variable statistique>

Une série statistique est un résumé de la variable statistique, on ne s'intéresse pas à la correspondance complète, mais seulement au nombre d'antécédents de chaque valeur prise. On note n_i le nombre d'antécédents de x_i, et on l'appelle l'effectif de la valeur x_i. Le nombre $f_i = \dfrac{n_i}{N}$ est appelé la fréquence de la valeur x_i.

< données de l'exercice 17 p. 257 sous forme de série statistique>

On peut aussi regrouper les valeurs par classes :

< données de l'exercice 17 p. 257 regroupées par classes>

Exercice 1. 5 et 9 p. 256

Théorème 1

Une série statistique pour une population de N individus :

Valeur	x_1	x_2	...	x_p
Effectif	n_1	n_2	...	n_p
Fréquence	f_1	f_2	...	f_p

étant donnée, on a les relations :

$$\sum_{i=1}^{n} f_i = f_1 + f_2 + \cdots + f_p = 1$$

$$\sum_{i=1}^{n} n_i = n_1 + n_2 + \cdots + n_p = N$$

DÉMONSTRATION : On a :

$$f_1 + f_2 + \cdots + f_p = \frac{n_1}{N} + \frac{n_2}{N} + \cdots + \frac{n_p}{N} = \frac{n_1 + n_2 + \cdots + n_p}{N} = \frac{N}{N} = 1. \qquad \square$$

Exercice 2. 35 1. p. 259

2 Effectifs et fréquences cumulés croissants

Exemple 2. Exercice 17 p. 257.

Définition 1

Le 1er quartile Q_1 est la plus petite valeur de la série pour laquelle la fréquence cumulée croissante dépasse 25%. Une médiane Me est une valeur comprise entre les deux fréquences cumulées croissantes successives qui encadrent 50%. Le 3e quartile Q_3 est la plus petite valeur de la série pour laquelle la fréquence cumulée croissante dépasse 75%.

Exercice 3. 10, 11, 13 p. 257

Exercice 4. 51 p. 261.

Remarque 1. La calculatrice donne tout directement : voir p. 326 dans le manuel.

3 Indicateurs de tendance centrale et de dispersion

3.1 Moyenne

Définition 2

Une série statistique pour une population de N individus :

Valeur	x_1	x_2	...	x_p
Effectif	n_1	n_2	...	n_p
Fréquence	f_1	f_2	...	f_p

étant donnée, la moyenne $\overline{x}$ est la valeur qui donne le même total lorsqu'on l'attribue à tous les individus :

$$N\overline{x} = n_1x_1 + n_2x_2 + \cdots + n_px_p.$$

Théorème 2

Dans le contexte de la définition précédente, on a :

$$\overline{x} = \frac{n_1x_1 + n_2x_2 + \cdots + n_px_p}{N}$$
$$\overline{x} = f_1x_1 + f_2x_2 + \cdots + f_px_p$$

Démonstration : La première s'obtient en divisant par N les deux membres de la définition, la deuxième s'obtient à partir de la première :

$$\begin{aligned}\overline{x} &= \frac{n_1x_1 + n_2x_2 + \cdots + n_px_p}{N}\\ &= \frac{n_1x_1}{N} + \frac{n_2x_2}{N} + \cdots + \frac{n_px_p}{N}\\ &= f_1x_1 + f_2x_2 + \cdots + f_px_p.\end{aligned}$$ □

Exercice 5. 38, 41 p. 260

Théorème 3 (*Associativité de la moyenne*)

Si plusieurs population d'effectifs N_1, N_2, ..., N_P ont pour moyennes respectives $\overline{x_1}$, $\overline{x_2}$, et N_p, alors la réunion de ces populations, d'effectif $N = N_1 + N_2 + \cdots + N_p$ a pour moyenne

$$\overline{x} = \frac{N_1\overline{x_1} + N_2\overline{x_2} + \cdots + N_p\overline{x_p}}{N_1 + N_2 + \cdots + N_p}.$$

Démonstration : D'après la définition de la moyenne, on doit avoir :

$$(N_1 + N_2 + \cdots + N_p)\overline{x} = N_1\overline{x_1} + N_2\overline{x_2} + \cdots + N_p\overline{x_p}.$$ □

Exercice 6. 39 p. 260.

3.2 Médiane

de manière informelle, une médiane d'une série statistique est une valeur qui partage l'effectif total en deux.

3.3 Étendue