

Chapitre 7

Statistique descriptive

Echantillonnage statistique

On appelle **population**, l'ensemble des individus auxquels on s'intéresse

On appelle **échantillon**, une fraction de la population

L'échantillonnage est l'ensemble des opérations qui ont pour objet de prélever un certain nombre d'individus dans la population.

L'échantillonnage aléatoire simple consiste à prélever au hasard et de façon indépendante n individus dans une population

Caractère statistique

On appelle **caractère statistique** simple toute application

$$X : P \rightarrow R$$

P est un ensemble fini appelé population

$$\omega \rightarrow X(\omega)$$

ω est un individu de la population

On appelle **modalité** toute valeur

$$x_i \in X(P)$$

$$\text{telle que : } X(P) = \{x_1, x_2, \dots, x_i, \dots, x_k\}$$

Caractère statistique

1. Caractères qualitatifs

Caractère mesuré dans une échelle : nominale (type de globules blancs)
ordinaire (stade d'une maladie)

2. Caractères quantitatifs discrets

Caractère qui ne peut prendre que certaines valeurs sur un intervalle

3. Caractères quantitatifs continus

Caractère pouvant prendre toutes les valeurs sur un intervalle donné

Tableau statistique

On appelle **effectif** de la modalité x_i , le nombre n_i d'individus

tel que: $X(\omega) = x_i$

On appelle **fréquence** de la modalité x_i , le nombre f_i

tel que: $f_i = \frac{n_i}{n}$

On appelle **fréquence cumulée** en x_i , le nombre f_i cum

tel que: $f_i \text{ cum} = \sum_{p=1}^i f_p$

Tableau statistique

1. Caractère quantitatif discret

Caractère X : x_i : nombre de galles par feuille	0	1	2	3	4	5	6	7	8	9	10
n_i : nombre de feuilles portant x_i galles	182	98	46	28	12	5	2	1	0	1	0
f_i : fréq. relative	0,485	0,261	0,123	0,075	0,032	0,013	0,005	0,003	0	0,003	0
f_i cum. : fréq. relative cumulée	0,485	0,746	0,869	0,944	0,976	0,989	0,994	0,997	0,997	1	1

$n = 375$ feuilles de hêtre examinées

Tableau statistique

2. Caractère quantitatif continu

Règle de Sturge : $1 + (3,3 \log 50) = 6,60$

Règle de Yule : $2,5 \sqrt[4]{50} = 6,64$

$$AC = \frac{174 - 140}{6,6} = 5,15 \text{ arrondi à } 5 \text{ mm}$$

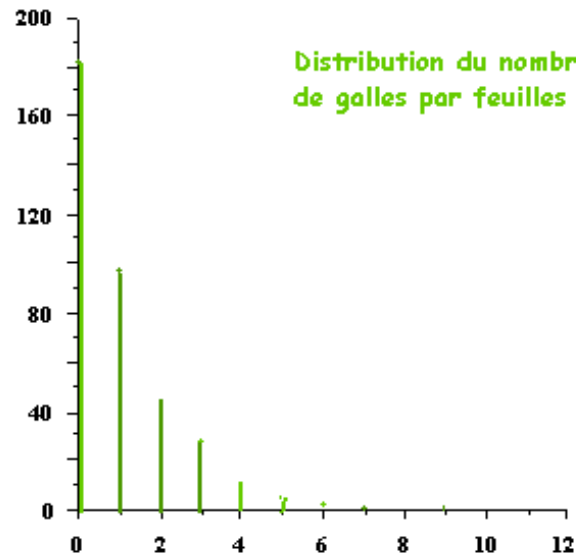
Caractère X : x_i : longueur de la rectrice bornes des classes	[140-145[	[145-150[	[150-155[	[155-160[	[160-165[	[165-170[	[170-175[
Valeur médiane des classes, x'_i	142,5	147,5	152,5	157,5	162,5	167,5	172,5
n_i : nombre d'individu par classe de taille x	1	1	9	17	16	3	3
f_i : fréquence relative	0,02	0,02	0,18	0,34	0,32	0,06	0,06
$f_{i \text{ cum.}}$: fréquence relative cumulée	0,02	0,04	0,22	0,56	0,88	0,94	1

n = 50 rectrices examinées

Représentation graphique

1. Caractère discret

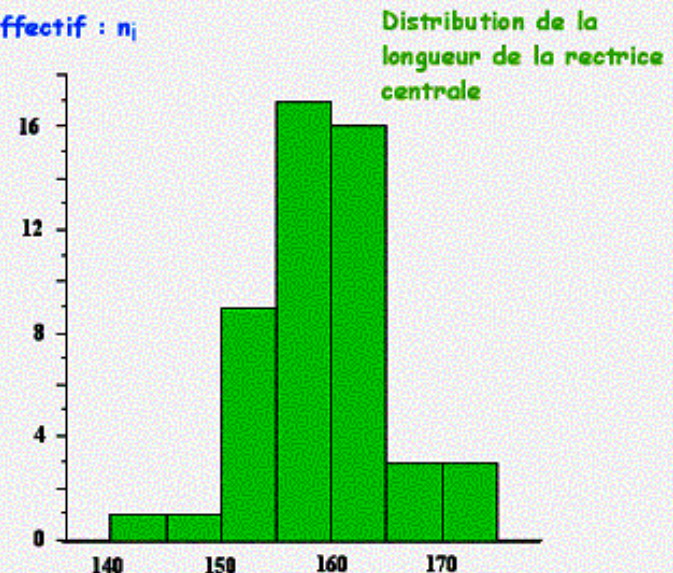
Effectif : n_i



X : nombre de galles par feuille

2. Caractère continu

Effectif : n_i



X : longueur de la rectrice centrale en mm

Indicateurs statistiques

1. Indicateurs de position

▪ La moyenne arithmétique

Soit un échantillon de n valeurs observées $x_1, x_2, \dots, x_i, \dots, x_n$ d'un caractère quantitatif X

Données non groupées

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

Données groupées pour caractère discret

$$\bar{x} = \frac{1}{n} \sum_{i=1}^k n_i x_i$$

Données groupées pour caractère continu

$$\bar{x} = \frac{1}{n} \sum_{i=1}^k n_i x_i'$$

Indicateurs statistiques

1. Indicateurs de position

▪ La médiane

La médiane est la valeur du caractère pour laquelle la fréquence cumulée est de 0,5. Soit n , le nombre de valeurs de la série:

Si n est impair, la médiane est la valeur au rang $(n+1)/2$

Si n est pair, la médiane est la **moyenne** des valeurs aux rangs $n/2$ et $(n/2) + 1$

▪ Le mode

Le mode est la valeur du caractère la plus fréquente

Indicateurs statistiques

2. Indicateurs de dispersion

▪ La variance

Soit un échantillon de n valeurs observées $x_1, x_2, \dots, x_i, \dots, x_n$ d'un caractère quantitatif X

Données non groupées
$$s^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{1}{n} \sum_{i=1}^n x_i^2 - \bar{x}^2$$

Données groupées pour caractère discret
$$s^2 = \frac{1}{n} \sum_{i=1}^k n_i (x_i - \bar{x})^2 = \frac{1}{n} \sum_{i=1}^k n_i x_i^2 - \bar{x}^2$$

Données groupées pour caractère continu
$$s^2 = \frac{1}{n} \sum_{i=1}^k n_i (x'_i - \bar{x})^2 = \frac{1}{n} \sum_{i=1}^k n_i x'^2_i - \bar{x}^2$$