

'e
é
é
é
é
é
é

Probabilités et Statistiques

Cours, Exemples et Exercices Corrigés

Fiche de révision

Table des matières

1	La Fonction de Répartition $F(x)$	3
2	Définitions du cours et Exemples	3
2.1	Effectifs, fréquences et cumuls	3
2.2	Caractéristiques de position	4
2.3	Caractéristiques de dispersion	4
2.4	Caractéristiques de forme	4
3	Représentations Graphiques	5
3.1	Diagramme à bandes (Variable qualitative)	5
3.2	Diagramme en bâtons (Variable discrète)	5
3.3	Histogramme (Variable continue)	5
3.4	Courbe cumulative $F(x)$ (Fréquences cumulées)	6
4	Exercice 1 : Données continues (Temps de trajet)	6
4.1	Énoncé	6
4.2	Corrigé de l'exercice 1	7
5	Exercice 2 : Données discrètes (Jours d'absence)	8
5.1	Énoncé	8
5.2	Corrigé de l'exercice 2	9

1 La Fonction de Répartition $F(x)$

Définition de la fonction $F(x)$

La **fonction de répartition**, notée $F(x)$ (ou fréquence cumulée), est définie par :

$$F(x) = \frac{N_i}{n}$$

où N_i est l'effectif cumulé croissant jusqu'à la valeur x , et n l'effectif total.

Interprétation : $F(x)$ représente la **proportion (ou le pourcentage)** d'individus dont la valeur de la variable est **inférieure ou égale à x** .

Propriétés essentielles :

- $F(x)$ est une fonction **croissante**.
- Elle varie de **0 à 1** (ou de 0% à 100%).
- Elle permet de lire graphiquement les caractéristiques de position :
 - **Médiane (M_e)** : valeur pour laquelle $F(M_e) = 0,5$.
 - **1er Quartile (Q_1)** : valeur pour laquelle $F(Q_1) = 0,25$.
 - **3ème Quartile (Q_3)** : valeur pour laquelle $F(Q_3) = 0,75$.

2 Définitions du cours et Exemples

2.1 Effectifs, fréquences et cumuls

Effectif (n_i) et Fréquence (f_i)

Effectif (n_i) : Nombre d'individus ayant la modalité x_i . La somme $\sum n_i = n$.

Fréquence (f_i) : Proportion $f_i = \frac{n_i}{n}$. La somme $\sum f_i = 1$.

Exemples

- *Exemple 1* : Dans un sondage de 100 personnes, 30 préfèrent le thé. $n_{\text{thé}} = 30$ et $f_{\text{thé}} = 0,30$ (30%).
- *Exemple 2* : Dans une classe de 25 élèves, 5 ont eu 15/20. $n_{15} = 5$ et $f_{15} = 5/25 = 0,20$ (20%).

Effectif cumulé (N_i) et Fréquence cumulée (F_i)

Effectif cumulé (N_i) : Somme des effectifs jusqu'à la modalité x_i ($N_i = n_1 + \dots + n_i$).

Fréquence cumulée (F_i) : $F_i = \frac{N_i}{n}$.

Exemples

- *Exemple 1* : Notes : 5 élèves ont 14, 8 ont 15, 4 ont 16. L'effectif cumulé pour 15 est $N_{15} = 5 + 8 = 13$.
- *Exemple 2* : La fréquence cumulée est $F_{15} = 13/25 = 0,52$. Cela signifie que 52% des élèves ont eu 15 ou moins.

2.2 Caractéristiques de position

Moyenne ($\bar{x}$)

- **Moyenne** : $\bar{x} = \frac{1}{n} \sum n_i x_i$ (ou $\sum f_i x_i$). Pour des classes, on utilise le centre c_i .
- **Mode** (M_o) : Valeur (ou classe) d'effectif maximal.
- **Médiane** (M_e) : Valeur qui partage la série ordonnée en deux groupes de même effectif ($F(M_e) = 0,5$).
- **Quartiles** (Q_1, Q_3) : $F(Q_1) = 0,25$ et $F(Q_3) = 0,75$.

Exemples

- *Moyenne* : 3 personnes gagnent 2000€, 2 personnes gagnent 3000€. $\bar{x} = \frac{3 \times 2000 + 2 \times 3000}{5} = 2400$ €.
- *Mode* : Dans la série {2, 3, 3, 4, 5, 3}, le mode est $M_o = 3$.
- *Médiane* : Série {1, 3, 5, 7, 9}. $M_e = 5$. Série {1, 3, 4, 6, 8, 9}. $M_e = (4+6)/2 = 5$.

2.3 Caractéristiques de dispersion

Étendue

- **Étendue** (e) : $e = x_{\max} - x_{\min}$.
- **Variance** (V) : $V(X) = \frac{1}{n} \sum n_i (x_i - \bar{x})^2 = \frac{1}{n} \sum n_i x_i^2 - \bar{x}^2$ (König-Huygens).
- **Écart-type** (σ) : $\sigma = \sqrt{V(X)}$. Même unité que X .
- **Coefficient de variation** (CV) : $CV = \frac{\sigma}{|\bar{x}|}$. Sans unité, permet de comparer l'homogénéité.

Exemples

- *Écart-type* : Si $V = 9$ (années²), alors $\sigma = 3$ ans.
- *CV* : Série A ($\bar{x} = 100, \sigma = 10 \Rightarrow CV = 10\%$). Série B ($\bar{x} = 10, \sigma = 2 \Rightarrow CV = 20\%$). La série A est plus homogène.

2.4 Caractéristiques de forme

Asymétrie et Aplatissement

- **Asymétrie (Pearson)** : $\frac{\bar{x} - M_o}{\sigma}$. > 0 étalée à droite, $= 0$ symétrique, < 0 étalée à gauche.
- **Aplatissement** : Mesure l'étalement par rapport à la loi normale.

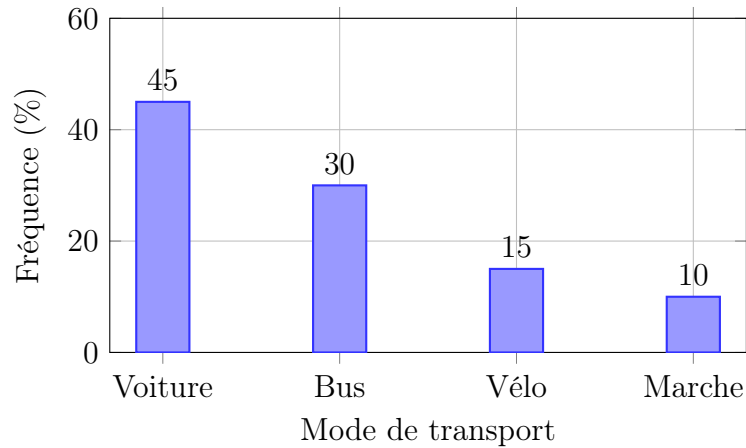
Exemples

- *Asymétrie à droite* : Revenus (beaucoup de petits salaires, quelques très gros). Moyenne $>$ Médiane.
- *Symétrique* : Taille des adultes. Moyenne $\approx$ Médiane.

3 Représentations Graphiques

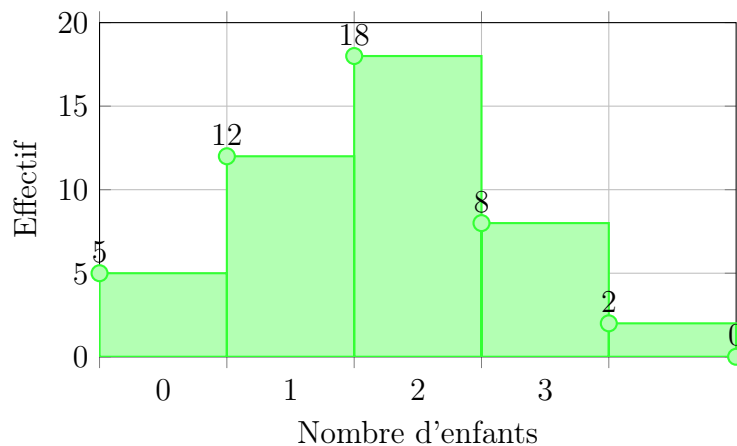
3.1 Diagramme à bandes (Variable qualitative)

Pour les variables qualitatives (catégories non ordonnées). Les rectangles sont séparés et leur hauteur est proportionnelle à l'effectif ou la fréquence.



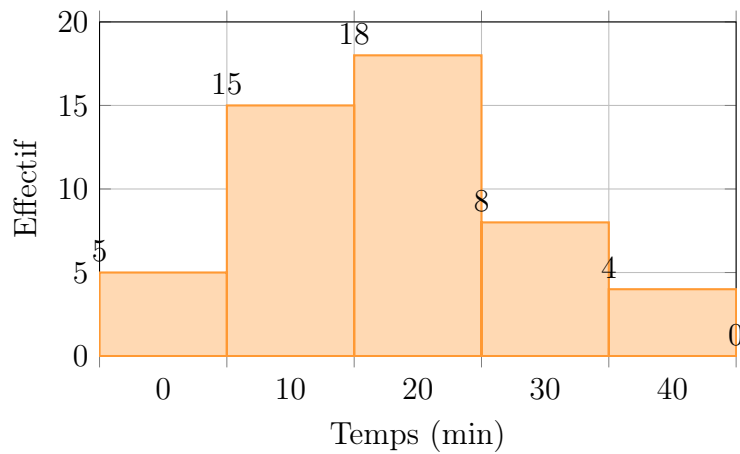
3.2 Diagramme en bâtons (Variable discrète)

Pour les variables discrètes (valeurs isolées). Des segments verticaux (bâtons) sont placés aux valeurs exactes.



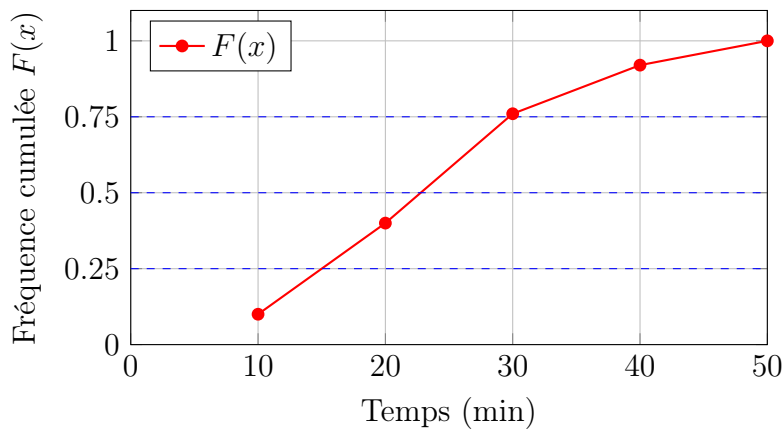
3.3 Histogramme (Variable continue)

Pour les variables continues regroupées en classes. L'**aire** du rectangle est proportionnelle à l'effectif. Si les amplitudes sont égales, la hauteur représente l'effectif.



3.4 Courbe cumulative $F(x)$ (Fréquences cumulées)

La courbe monte progressivement de 0 à 1. On peut y lire la médiane et les quartiles à l'aide de lignes horizontales.



4 Exercice 1 : Données continues (Temps de trajet)

4.1 Énoncé

Le temps de trajet (en minutes) pour se rendre au travail a été relevé sur un échantillon de **50 employés**. Les données sont regroupées en classes :

Temps (min)	[0, 10[	[10, 20[	[20, 30[	[30, 40[	[40, 50[
Effectif (n_i)	5	15	18	8	4

Travail demandé :

1. Compléter le tableau avec les centres de classe, les effectifs cumulés (N_i) et les fréquences cumulées (F_i).
2. Calculer la **moyenne** du temps de trajet.
3. Déterminer la **classe modale** et calculer le **mode** approché.
4. Calculer la **médiane** (M_e) et les **quartiles** (Q_1, Q_3) par interpolation linéaire.
5. Interpréter la médiane et l'écart interquartile.

4.2 Corrigé de l'exercice 1

1. Tableau statistique complet

L'amplitude de chaque classe est $h = 10$. Le centre $c_i = (a_i + a_{i+1})/2$.

Classe	n_i	Centre c_i	$n_i \times c_i$	N_i (Cumul)	$F_i = N_i/50$
$[0, 10[$	5	5	25	5	0,10
$[10, 20[$	15	15	225	20	0,40
$[20, 30[$	18	25	450	38	0,76
$[30, 40[$	8	35	280	46	0,92
$[40, 50[$	4	45	180	50	1,00
Total	50	-	1160	-	-

2. Calcul de la Moyenne

$$\bar{x} = \frac{\sum n_i c_i}{n} = \frac{1160}{50} = \mathbf{23,2 \text{ minutes}}$$

3. Mode

Classe modale : C'est la classe d'effectif maximal. Ici, c'est $[20, 30[$ (avec 18 employés).

Calcul du mode (M_o) par interpolation : $a = 20$, $h = 10$, $d_1 = 18 - 15 = 3$, $d_2 = 18 - 8 = 10$.

$$M_o = a + h \times \frac{d_1}{d_1 + d_2} = 20 + 10 \times \frac{3}{3 + 10} = 20 + \frac{30}{13} \approx \mathbf{22,3 \text{ minutes}}$$

4. Médiane et Quartiles (Interpolation linéaire)

La formule générale est : $Val = a + h \times \frac{\text{Position cible} - N_{\text{préc}}}{n_{\text{classe}}}$

Médiane (M_e) : On cherche $F(x) = 0,5$ (soit $N = 25$). La classe médiane est $[20, 30[$.

$$M_e = 20 + 10 \times \frac{25 - 20}{18} = 20 + \frac{50}{18} \approx \mathbf{22,78 \text{ minutes}}$$

1er Quartile (Q_1) : On cherche $F(x) = 0,25$ (soit $N = 12,5$). La classe est $[10, 20[$.

$$Q_1 = 10 + 10 \times \frac{12,5 - 5}{15} = 10 + \frac{75}{15} = \mathbf{15 \text{ minutes}}$$

3ème Quartile (Q_3) : On cherche $F(x) = 0,75$ (soit $N = 37,5$). La classe est $[20, 30[$.

$$Q_3 = 20 + 10 \times \frac{37,5 - 20}{18} = 20 + \frac{175}{18} \approx \mathbf{29,72 \text{ minutes}}$$

5. Interprétation

- **Médiane :** 50% des employés mettent moins de 22,78 minutes pour venir travailler, et 50% mettent plus.
- **Écart interquartile ($Q_3 - Q_1$) :** $29,72 - 15 = 14,72$ minutes. Cela signifie que

50% des employés (la partie centrale) ont un temps de trajet compris dans une plage de 14,72 minutes (entre 15 et 29,72 min).

5 Exercice 2 : Données discrètes (Jours d'absence)

5.1 Énoncé

Voici le nombre de jours d'absence sur une année de **40 employés** :

Nombre de jours (x_i)	0	1	2	3	4	5
Effectif (n_i)	8	12	10	6	3	1

Travail demandé :

1. Compléter le tableau avec les fréquences (f_i), les effectifs cumulés (N_i) et les fréquences cumulées (F_i).
2. Calculer la **moyenne** du nombre de jours d'absence.
3. Déterminer le **mode**.
4. Calculer la **médiane** (M_e) et les **quartiles** (Q_1, Q_3).
5. Calculer l'**écart interquartile** et l'interpréter.

5.2 Corrigé de l'exercice 2

1. Tableau statistique complet

x_i	n_i	$f_i = n_i/40$	N_i (Cumul)	$F_i = N_i/40$
0	8	0,20	8	0,20
1	12	0,30	20	0,50
2	10	0,25	30	0,75
3	6	0,15	36	0,90
4	3	0,075	39	0,975
5	1	0,025	40	1,00
Total	40	1,00	-	-

2. Calcul de la Moyenne

$$\bar{x} = \frac{\sum n_i x_i}{n} = \frac{0 \times 8 + 1 \times 12 + 2 \times 10 + 3 \times 6 + 4 \times 3 + 5 \times 1}{40}$$
$$\bar{x} = \frac{0 + 12 + 20 + 18 + 12 + 5}{40} = \frac{67}{40} = 1,675 \text{ jours}$$

3. Mode

L'effectif maximal est $n_1 = 12$. Le mode est donc :

$$M_o = 1 \text{ jour}$$

C'est le nombre de jours d'absence le plus fréquent parmi les employés.

4. Médiane et Quartiles

Pour les données discrètes, on cherche directement dans le tableau des effectifs cumulés N_i .

Médiane (M_e) : On cherche la position $n/2 = 40/2 = 20$.

En regardant N_i , on voit que $N_1 = 20$. La 20ème valeur est donc $x_1 = 1$.

$$M_e = 1 \text{ jour}$$

1er Quartile (Q_1) : On cherche la position $n/4 = 40/4 = 10$.

En regardant N_i , on voit que $N_1 = 20 \geq 10$. La 10ème valeur est donc $x_1 = 1$.

$$Q_1 = 1 \text{ jour}$$

3ème Quartile (Q_3) : On cherche la position $3n/4 = 3 \times 40/4 = 30$.

En regardant N_i , on voit que $N_2 = 30$. La 30ème valeur est donc $x_2 = 2$.

$$Q_3 = 2 \text{ jours}$$

5. Écart interquartile et Interprétation

$$IQR = Q_3 - Q_1 = 2 - 1 = \mathbf{1 \text{ jour}}$$

Interprétation : 50% des employés (la partie centrale de la distribution) ont entre 1 et 2 jours d'absence dans l'année. L'écart interquartile de 1 jour montre que les absences sont assez concentrées autour de 1-2 jours pour la moitié des employés.