

Statistiques et Probabilités

0.1 Objectifs du programme

Dans cette partie du programme, nous allons consolider et prolonger les notions d'information chiffrée (effectifs, fréquences, proportions, pourcentages, taux d'évolution) et introduire de nouveaux indicateurs statistiques. Nous aborderons également la modélisation du hasard et le calcul des probabilités.

0.2 Utiliser l'information chiffrée et statistique descriptive

0.2.1 Proportions et Pourcentages

Au collège, vous avez manipulé les notions de proportion et de pourcentage. Nous allons ici approfondir ces concepts, notamment dans le cas de pourcentages de pourcentages.

Définition

Proportion : La proportion d'une sous-population A au sein d'une population E est le rapport entre l'effectif de A et l'effectif de E .

$$p(A) = \frac{\text{effectif de } A}{\text{effectif de } E}$$

Pourcentage : Un pourcentage est une proportion exprimée pour 100. Si p est la proportion, le pourcentage correspondant est $p \times 100$.

Point de méthode

Pour exploiter la relation entre effectifs, proportions et pourcentages, il est souvent utile de construire un tableau ou d'utiliser la règle de trois.

0.2.2 Pourcentages de Pourcentages

Il s'agit de situations où une proportion est calculée au sein d'une sous-population, elle-même proportion d'une population plus grande.

Propriété

Si A est une sous-population de E représentant $p_1\%$ de E , et B est une sous-population de A représentant $p_2\%$ de A , alors B représente $p_3\%$ de E , avec $p_3 = p_1 \times p_2 / 100$.

Exemple commenté : Dans une entreprise de 200 employés, 60% sont des hommes. Parmi ces hommes, 30% sont en télétravail. Quel pourcentage des employés de l'entreprise sont des hommes en télétravail ?

Étape 1 : Calculer le nombre d'hommes dans l'entreprise. Nombre d'hommes = $200 \times \frac{60}{100} = 120$ hommes.

Étape 2 : Calculer le nombre d'hommes en télétravail. Nombre d'hommes en télétravail = $120 \times \frac{30}{100} = 36$ hommes.

Étape 3 : Calculer le pourcentage d'hommes en télétravail par rapport à l'ensemble des employés. Pourcentage = $\frac{36}{200} \times 100 = 18\%$.

Alternative (en utilisant la propriété des pourcentages de pourcentages) : Pourcentage = $60 \times \frac{30}{100} = 18\%$.

0.2.3 Évolutions

Une évolution décrit la variation d'une grandeur dans le temps. On distingue la variation absolue et la variation relative.

Définition

Variation absolue : Si une grandeur passe d'une valeur initiale V_i à une valeur finale V_f , la variation absolue est $V_f - V_i$. **Variation relative (ou taux d'évolution)** : La variation relative est $\frac{V_f - V_i}{V_i}$. Elle est souvent exprimée en pourcentage. **Coefficient multiplicateur** : Le coefficient multiplicateur associé à une évolution est $CM = \frac{V_f}{V_i} = 1 + \text{taux d'évolution}$.

Point de méthode

Pour passer d'une valeur à une autre suite à une évolution, on peut utiliser le coefficient multiplicateur.

Exemple commenté : Le prix d'un article passe de 50€ à 55€. *Calcul de la variation absolue* : $55 - 50 = 5$. *Calcul du taux d'évolution* : $\frac{55-50}{50} = \frac{5}{50} = 0,1 = 10\%$. *Calcul du coefficient multiplicateur* : $\frac{55}{50} = 1,1$. On remarque que $1,1 = 1 + 0,1$.

0.2.4 Évolutions Successives et Réciproques

Propriété

Évolutions successives : Si une grandeur subit plusieurs évolutions successives de coefficients multiplicateurs $CM_1, CM_2, \dots, CM_n$, le coefficient multiplicateur global est le produit des coefficients multiplicateurs individuels : $CM_{global} = CM_1 \times CM_2 \times \dots \times CM_n$. **Évolution réciproque** : Si une grandeur a subi une évolution de coefficient multiplicateur CM , le coefficient multiplicateur de l'évolution réciproque (pour revenir à la valeur initiale) est $1/CM$.

Exemple commenté : Un article coûte 100€. Son prix augmente de 10%, puis diminue de 20%.

Étape 1 : Calculer les coefficients multiplicateurs de chaque évolution. Augmentation de 10% : $CM_1 = 1 + \frac{10}{100} = 1,1$. Diminution de 20% : $CM_2 = 1 - \frac{20}{100} = 0,8$.

Étape 2 : Calculer le coefficient multiplicateur global. $CM_{global} = CM_1 \times CM_2 = 1,1 \times 0,8 = 0,88$.

Étape 3 : Calculer le prix final. Prix final = $100 \times 0,88 = 88$. Le taux d'évolution global est $0,88 - 1 = -0,12$, soit une diminution de 12%.

Exemple d'évolution réciproque : Le prix d'un produit a augmenté de 25%. Quel est le taux de diminution nécessaire pour revenir au prix initial ?

Étape 1 : Calculer le coefficient multiplicateur de l'augmentation. $CM = 1 + \frac{25}{100} = 1,25$.

Étape 2 : Calculer le coefficient multiplicateur réciproque. $CM_{reciproque} = \frac{1}{1,25} = 0,8$.

Étape 3 : Déduire le taux d'évolution réciproque. Taux réciproque = $0,8 - 1 = -0,2 = -20\%$. Il faut donc une diminution de 20%.

0.2.5 Indicateurs Statistiques

Au-delà de la moyenne et de l'étendue vues au collège, nous introduisons la moyenne pondérée, l'écart interquartile et l'écart type pour mieux décrire une série statistique.

Définition

Moyenne pondérée : Pour une série de données $(x_1, x_2, \dots, x_k)$ avec des effectifs (ou poids) $(n_1, n_2, \dots, n_k)$, la moyenne pondérée est :

$$\bar{x} = \frac{n_1x_1 + n_2x_2 + \dots + n_kx_k}{n_1 + n_2 + \dots + n_k}$$

Écart interquartile : C'est la différence entre le troisième quartile (Q_3) et le premier quartile (Q_1). $EI = Q_3 - Q_1$. Il mesure l'étendue de la série où se situent 50% des valeurs centrales.

Écart type : Il mesure la dispersion des valeurs autour de la moyenne. Sa formule est :

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2}$$

où N est la taille de la population.

Propriété

Linéarité de la moyenne : Si $y_i = ax_i + b$, alors $\bar{y} = a\bar{x} + b$.

Exemple commenté (Moyenne pondérée) : Notes d'une classe : 8 (coeff 2), 12 (coeff 3), 15 (coeff 1). Moyenne pondérée = $\frac{8 \times 2 + 12 \times 3 + 15 \times 1}{2 + 3 + 1} = \frac{16 + 36 + 15}{6} = \frac{67}{6} \approx 11,17$.

Exemple commenté (Écart interquartile) : Série de données triées : 5, 7, 8, 10, 12, 13, 15, 18, 20. Il y a 9 valeurs. Q_1 est la valeur pour laquelle au moins 25% des données sont inférieures ou égales. $0,25 \times 9 = 2,25$. Q_1 est la 3ème valeur, soit 8. Q_3 est la valeur pour laquelle au moins 75% des données sont inférieures ou égales. $0,75 \times 9 = 6,75$. Q_3 est la 7ème valeur, soit 15. Écart interquartile = $15 - 8 = 7$.

Les calculatrices et les logiciels sont des outils précieux pour calculer l'écart type et les quartiles pour

0.3 Modéliser le hasard, calculer des probabilités

0.3.1 Vocabulaire des Probabilités

Nous allons formaliser le langage des probabilités, en nous basant sur les notions intuitives acquises au collège. L'ensemble des issues est fini.

Définition

Expérience aléatoire : Une expérience dont le résultat ne peut pas être prédit avec certitude. **Issue (ou résultat)** : Un résultat possible d'une expérience aléatoire. **Univers** : L'ensemble de toutes les issues possibles d'une expérience aléatoire. Il est noté Ω . **Événement** : Une partie de l'univers, c'est-à-dire un ensemble d'issues. **Événement élémentaire** : Un événement qui ne contient qu'une seule issue. **Événement impossible** : L'événement qui ne contient aucune issue ($\emptyset$). Sa probabilité est 0. **Événement certain** : L'événement qui contient toutes les issues (Ω). Sa probabilité est 1.

0.3.2 Opérations sur les Événements

On utilise les notions ensemblistes pour décrire les relations entre événements.

Définition

Soient A et B deux événements : **Intersection** ($A \cap B$) : L'événement dont les issues appartiennent à la fois à A et à B . Il se lit "A et B". **Réunion** ($A \cup B$) : L'événement dont les issues appartiennent à A ou à B (ou aux deux). Il se lit "A ou B". **Complémentaire** ($\bar{A}$ ou A^c) : L'événement dont les issues n'appartiennent pas à A . Il se lit "non A".

Propriété

Probabilité d'une réunion : Pour deux événements A et B , la probabilité de leur réunion est donnée par la formule :

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Probabilité du complémentaire : $P(\bar{A}) = 1 - P(A)$.

Théorème

Loi (ou distribution) de probabilité : Une loi de probabilité sur un univers fini $\Omega = \{e_1, e_2, \dots, e_n\}$ est l'attribution à chaque issue e_i d'un nombre p_i tel que :

- $0 \leq p_i \leq 1$ pour tout i .
- $\sum_{i=1}^n p_i = p_1 + p_2 + \dots + p_n = 1$. La probabilité d'un événement est la somme des probabilités des issues qui le composent.

Il est crucial de distinguer le modèle probabiliste abstrait et la situation réelle.

Une loi de probabilité est une hypothèse du modèle choisi, elle ne se démontre pas.

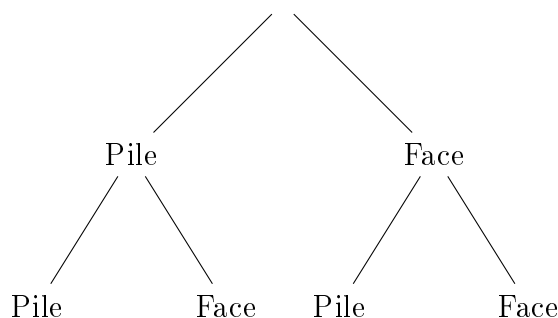
Exemple commenté : On lance un dé équilibré à six faces. L'univers est $\Omega = \{1, 2, 3, 4, 5, 6\}$. Puisque le dé est équilibré, la probabilité de chaque issue est $\frac{1}{6}$. C'est une situation d'équiprobabilité. Soit l'événement A : "obtenir un nombre pair". $A = \{2, 4, 6\}$. $P(A) = P(2) + P(4) + P(6) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{3}{6} = \frac{1}{2}$. Soit l'événement B : "obtenir un nombre supérieur à 4". $B = \{5, 6\}$. $P(B) = P(5) + P(6) = \frac{1}{6} + \frac{1}{6} = \frac{2}{6} = \frac{1}{3}$.

Calcul de $A \cap B$: $A \cap B = \{6\}$. $P(A \cap B) = \frac{1}{6}$. *Calcul de $A \cup B$* : $A \cup B = \{2, 4, 5, 6\}$. En utilisant la formule : $P(A \cup B) = P(A) + P(B) - P(A \cap B) = \frac{1}{2} + \frac{1}{3} - \frac{1}{6} = \frac{3}{6} + \frac{2}{6} - \frac{1}{6} = \frac{4}{6} = \frac{2}{3}$. Par dénombrement direct : $P(A \cup B) = \frac{\text{nombre d'issues dans } A \cup B}{\text{nombre total d'issues}} = \frac{4}{6} = \frac{2}{3}$.

0.3.3 Dénombrement à l'aide de Tableaux et d'Arbres

Pour les expériences aléatoires à plusieurs épreuves, les tableaux et les arbres sont des outils graphiques efficaces pour visualiser l'univers et dénombrer les issues.

Exemple commenté (Diagramme en arbre) : On lance une pièce de monnaie deux fois de suite.



L'univers est $\Omega = \{PP, PF, FP, FF\}$. Chaque issue a une probabilité de $\frac{1}{4}$.

Exemple commenté (Tableau à double entrée) : On lance deux dés équilibrés.

Dés	1	2	3	4	5	6
1	(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
2	(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
3	(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
4	(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
5	(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
6	(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

Il y a $6 \times 6 = 36$ issues possibles. Si on s'intéresse à l'événement "obtenir une somme de 7", les issues sont (1,6), (2,5), (3,4), (4,3), (5,2), (6,1). Il y a 6 issues favorables. La probabilité est $\frac{6}{36} = \frac{1}{6}$.

0.4 Échantillonnage

L'échantillonnage nous permet d'explorer le lien entre la théorie des probabilités et l'observation de phénomènes réels, notamment à travers la loi des grands nombres.

Définition

Échantillon aléatoire : Un sous-ensemble d'une population choisi de manière aléatoire. Sa taille est notée n . **Fréquence observée :** Dans un échantillon, c'est la proportion d'une certaine issue ou événement.

Théorème

Version vulgarisée de la loi des grands nombres : Lorsque la taille de l'échantillon n est grande, la fréquence observée d'un événement dans cet échantillon est généralement proche de la probabilité théorique de cet événement.

La loi des grands nombres est fondamentale pour comprendre comment les statistiques issues d'échantillons peuvent nous informer sur les probabilités d'une population.

Point de méthode

Pour observer la loi des grands nombres, il est recommandé d'utiliser des simulations informatiques (par exemple avec Python ou un tableur) qui permettent de répéter un grand nombre de fois une expérience aléatoire.