

Polycopié de cours : Probabilités et Statistiques (Première Générale)

<https://exomaths.fr>

18 juin 2025

1 Introduction aux Probabilités et Statistiques

L'enseignement des probabilités et statistiques en classe de première générale vise à consolider les acquis de la classe de seconde et à introduire de nouvelles notions essentielles pour une compréhension approfondie du modèle probabiliste. Ce programme met en évidence la distinction entre modèle et réalité, notamment en ce qui concerne l'hypothèse d'indépendance. Il articule également les notions de statistique descriptive vues en seconde avec les probabilités, en faisant un lien entre des notions statistiques (sous-population, proportion, moyenne, écart type) et leurs analogues probabilistes (événement, probabilité, espérance, écart type).

2 Probabilités conditionnelles et indépendance

Cette partie du programme introduit la notion de probabilité conditionnelle, fondamentale en modélisation probabiliste, et aborde la problématique de l'inversion des conditionnements. La notion d'indépendance entre événements est également formalisée.

2.1 Probabilité conditionnelle

Définition

Soient A et B deux événements, avec $P(A) \neq 0$. La probabilité de B sachant A , notée $P_A(B)$, est la probabilité que l'événement B se réalise sachant que l'événement A est déjà réalisé.

$$P_A(B) = \frac{P(A \cap B)}{P(A)}$$

Propriété

Règle du produit : Pour deux événements A et B tels que $P(A) \neq 0$, on a $P(A \cap B) = P(A) \times P_A(B)$.

Remarque 1. Les probabilités conditionnelles peuvent être calculées à partir d'un tableau croisé d'effectifs.

Point de méthode

Calcul d'une probabilité conditionnelle à partir d'un tableau croisé d'effectifs

1. Identifier la population totale et la sous-population correspondant à l'événement conditionnant.
2. Identifier le nombre d'individus dans la sous-population qui satisfont l'événement dont on cherche la probabilité.
3. Calculer le rapport entre ces deux nombres.

Exemple 1. Un sondage est réalisé auprès de 200 personnes. 120 sont des femmes et 80 sont des hommes. Parmi les femmes, 70 aiment le chocolat. Parmi les hommes, 30 aiment le chocolat. On choisit une personne au hasard dans ce groupe.

On organise les données dans un tableau :

	Aiment le chocolat (C)	N'aiment pas le chocolat ($\bar{C}$)	Total
Femmes (F)	70	50	120
Hommes (H)	30	50	80
Total	100	100	200

Quelle est la probabilité qu'une personne aime le chocolat, sachant que c'est une femme ? C'est $P_F(C)$. D'après le tableau, il y a 120 femmes au total, et 70 d'entre elles aiment le chocolat. Donc $P_F(C) = \frac{70}{120} = \frac{7}{12} \approx 0,583$.

En utilisant la formule : $P(F) = \frac{120}{200} = \frac{3}{5}$. $P(F \cap C) = \frac{70}{200} = \frac{7}{20}$. $P_F(C) = \frac{P(F \cap C)}{P(F)} = \frac{7/20}{3/5} = \frac{7}{20} \times \frac{5}{3} = \frac{35}{60} = \frac{7}{12}$.

2.2 Arbres pondérés et formule des probabilités totales

Propriété

Arbres pondérés : Les arbres pondérés sont des outils visuels utiles pour représenter des situations impliquant des probabilités conditionnelles et calculer des probabilités.

- Chaque branche est associée à une probabilité.
- La somme des probabilités des branches issues d'un même nœud est égale à 1.
- La probabilité d'un chemin est le produit des probabilités des branches qui le composent (règle du produit).
- La probabilité d'un événement est la somme des probabilités des chemins qui mènent à cet événement (règle de la somme).

Définition

Un ensemble d'événements $A_1, A_2, \dots, A_n$ forme une **partition de l'univers** (ou un système complet d'événements) si :

- ils sont deux à deux disjoints (c'est-à-dire $A_i \cap A_j = \emptyset$ pour $i \neq j$) ;
- leur union est égale à l'univers Ω (c'est-à-dire $A_1 \cup A_2 \cup \dots \cup A_n = \Omega$) ;
- la probabilité de chacun d'eux est non nulle ($P(A_i) \neq 0$ pour tout i).

Propriété

Formule des probabilités totales : Si $A_1, A_2, \dots, A_n$ forment une partition de l'univers, alors pour tout événement B , on a :

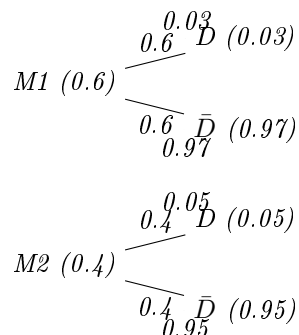
$$P(B) = P(B \cap A_1) + P(B \cap A_2) + \dots + P(B \cap A_n)$$

Ce qui peut aussi s'écrire :

$$P(B) = P(A_1) \times P_{A_1}(B) + P(A_2) \times P_{A_2}(B) + \dots + P(A_n) \times P_{A_n}(B)$$

Exemple 2. Dans une usine, deux machines $M1$ et $M2$ produisent des pièces. La machine $M1$ produit 60% des pièces et la machine $M2$ produit 40% des pièces. Le taux de pièces défectueuses de $M1$ est de 3% et celui de $M2$ est de 5%.

On peut représenter cette situation par un arbre pondéré :



Quelle est la probabilité qu'une pièce soit défectueuse ? Les événements $M1$ et $M2$ forment une partition de l'univers. On applique la formule des probabilités totales : $P(D) = P(M1 \cap D) + P(M2 \cap D)$ $P(D) = P(M1) \times P_{M1}(D) + P(M2) \times P_{M2}(D)$ $P(D) = 0.6 \times 0.03 + 0.4 \times 0.05$ $P(D) = 0.018 + 0.020$ $P(D) = 0.038$ La probabilité qu'une pièce soit défectueuse est de 3,8%.

2.3 Indépendance de deux événements

Définition

Deux événements A et B sont dits **indépendants** si la réalisation de l'un n'influence pas la probabilité de réalisation de l'autre. Mathématiquement, cela se traduit par l'une des conditions équivalentes suivantes (si les probabilités sont non nulles) :

- $P(A \cap B) = P(A) \times P(B)$
- $P_A(B) = P(B)$
- $P_B(A) = P(A)$

Remarque 2. Il est crucial de distinguer $P_A(B)$ et $P_B(A)$, notamment dans les situations de type "faux positifs".

Propriété

Succession de deux épreuves indépendantes : Lors d'une succession de deux épreuves indépendantes, la probabilité de la réalisation de deux événements A et B (où A est lié à la première épreuve et B à la seconde) est le produit de leurs probabilités. Cela peut être représenté par un arbre ou un tableau.

3 Variables aléatoires réelles

Cette section formalise la notion de variable aléatoire comme une fonction numérique définie sur un univers, permettant d'affecter des probabilités aux valeurs possibles de la variable. Les notions d'espérance, de variance et d'écart type sont introduites. Le programme se limite aux univers finis et aux variables aléatoires réelles.

3.1 Définition et loi de probabilité

Définition

Une **variable aléatoire réelle** X est une fonction qui associe un nombre réel à chaque issue d'une expérience aléatoire.

Définition

La **loi de probabilité** d'une variable aléatoire X est la donnée de toutes les valeurs x_i que X peut prendre, et de leurs probabilités associées $P(X = x_i)$.

Point de méthode

Déterminer la loi de probabilité d'une variable aléatoire

1. Lister toutes les valeurs possibles que la variable aléatoire peut prendre.
2. Pour chaque valeur x_i , calculer la probabilité $P(X = x_i)$.
3. Vérifier que la somme de toutes les probabilités est égale à 1.

Exemple 3. On lance deux dés équilibrés à six faces et on s'intéresse à la somme des deux faces obtenues. Soit X la variable aléatoire qui représente cette somme. Les valeurs possibles de X vont de $1 + 1 = 2$ à $6 + 6 = 12$. L'univers est composé de $6 \times 6 = 36$ issues équiprobables.

Calculons les probabilités pour chaque valeur de X :

- $P(X = 2) = P((1, 1)) = \frac{1}{36}$
- $P(X = 3) = P((1, 2), (2, 1)) = \frac{2}{36}$
- $P(X = 4) = P((1, 3), (2, 2), (3, 1)) = \frac{3}{36}$
- $P(X = 5) = P((1, 4), (2, 3), (3, 2), (4, 1)) = \frac{4}{36}$
- $P(X = 6) = P((1, 5), (2, 4), (3, 3), (4, 2), (5, 1)) = \frac{5}{36}$
- $P(X = 7) = P((1, 6), (2, 5), (3, 4), (4, 3), (5, 2), (6, 1)) = \frac{6}{36}$
- $P(X = 8) = P((2, 6), (3, 5), (4, 4), (5, 3), (6, 2)) = \frac{5}{36}$
- $P(X = 9) = P((3, 6), (4, 5), (5, 4), (6, 3)) = \frac{4}{36}$
- $P(X = 10) = P((4, 6), (5, 5), (6, 4)) = \frac{3}{36}$
- $P(X = 11) = P((5, 6), (6, 5)) = \frac{2}{36}$
- $P(X = 12) = P((6, 6)) = \frac{1}{36}$

La loi de probabilité de X peut être présentée sous forme de tableau :

x_i	2	3	4	5	6	7	8	9	10	11	12
$P(X = x_i)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

3.2 Espérance, variance et écart type

Définition

Soit X une variable aléatoire dont la loi de probabilité est donnée par les valeurs $x_1, x_2, \dots, x_n$ et les probabilités associées $P(X = x_1), P(X = x_2), \dots, P(X = x_n)$.

L'espérance $E(X)$ de X est la valeur moyenne de X sur un grand nombre d'expériences. Elle est calculée par :

$$E(X) = \sum_{i=1}^n x_i P(X = x_i) = x_1 P(X = x_1) + x_2 P(X = x_2) + \dots + x_n P(X = x_n)$$

La variance $V(X)$ de X mesure la dispersion des valeurs de X autour de son espérance. Elle est calculée par :

$$V(X) = \sum_{i=1}^n (x_i - E(X))^2 P(X = x_i)$$

L'écart type $\sigma(X)$ de X est la racine carrée de la variance. Il est exprimé dans la même unité que la variable aléatoire, ce qui le rend plus facilement interprétable.

$$\sigma(X) = \sqrt{V(X)}$$

Point de méthode

Calculer l'espérance, la variance et l'écart type d'une variable aléatoire

1. Déterminer la loi de probabilité de la variable aléatoire.
2. Calculer l'espérance $E(X)$ en utilisant la formule.
3. Calculer la variance $V(X)$ en utilisant la formule.
4. Calculer l'écart type $\sigma(X)$ en prenant la racine carrée de la variance.

Exemple 4. Reprenons l'exemple du jet de deux dés et de la somme X . Nous avons la loi de probabilité suivante :

x_i	2	3	4	5	6	7	8	9	10	11	12
$P(X = x_i)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Calcul de l'espérance $E(X)$: $E(X) = 2 \times \frac{1}{36} + 3 \times \frac{2}{36} + 4 \times \frac{3}{36} + 5 \times \frac{4}{36} + 6 \times \frac{5}{36} + 7 \times \frac{6}{36} + 8 \times \frac{5}{36} + 9 \times \frac{4}{36} + 10 \times \frac{3}{36} + 11 \times \frac{2}{36} + 12 \times \frac{1}{36}$ $E(X) = \frac{1}{36}(2 + 6 + 12 + 20 + 30 + 42 + 40 + 36 + 30 + 22 + 12)$ $E(X) = \frac{252}{36} = 7$
L'espérance de la somme des deux dés est 7.

Calcul de la variance $V(X)$: $V(X) = \sum (x_i - E(X))^2 P(X = x_i)$ $V(X) = (2-7)^2 \frac{1}{36} + (3-7)^2 \frac{2}{36} + (4-7)^2 \frac{3}{36} + (5-7)^2 \frac{4}{36} + (6-7)^2 \frac{5}{36} + (7-7)^2 \frac{6}{36} + (8-7)^2 \frac{5}{36} + (9-7)^2 \frac{4}{36} + (10-7)^2 \frac{3}{36} + (11-7)^2 \frac{2}{36} + (12-7)^2 \frac{1}{36}$
 $V(X) = \frac{1}{36}((-5)^2 \times 1 + (-4)^2 \times 2 + (-3)^2 \times 3 + (-2)^2 \times 4 + (-1)^2 \times 5 + 0^2 \times 6 + 1^2 \times 5 + 2^2 \times 4 + 3^2 \times 3 + 4^2 \times 2 + 5^2 \times 1)$
 $V(X) = \frac{1}{36}(25 + 32 + 27 + 16 + 5 + 0 + 5 + 16 + 27 + 32 + 25)$ $V(X) = \frac{210}{36} = \frac{35}{6} \approx 5.83$

Calcul de l'écart type $\sigma(X)$: $\sigma(X) = \sqrt{V(X)} = \sqrt{\frac{35}{6}} \approx \sqrt{5.833} \approx 2.415$

Remarque 3. L'espérance est un concept clé pour résoudre des problèmes, par exemple dans le cadre de jeux équitables où la "mise pour un jeu équitable" est liée à l'espérance de gain.

4 Expérimentations

Le programme encourage le travail expérimental de simulation d'échantillons pour faire percevoir le principe de l'estimation de l'espérance d'une variable aléatoire, ou de la moyenne d'une variable statistique dans une population, par une moyenne observée sur un échantillon.

Point de méthode**Simulation d'une variable aléatoire avec Python**

1. Importer la bibliothèque 'random' ou 'numpy'.
2. Utiliser des fonctions comme 'random.choice()' ou 'numpy.random.choice()' pour simuler les issues d'une variable aléatoire selon sa loi de probabilité.
3. Répéter la simulation un grand nombre de fois pour obtenir un échantillon.

Exemple 5. Simulation du jet d'un dé et calcul de la moyenne d'un échantillon

```
import random

def lancer_de():
    """Simule le lancer d'un dé équilibré à six faces."""
    return random.randint(1, 6)

def moyenne_echantillon(taille_echantillon):
    """Calcule la moyenne d'un échantillon de lancers de dé."""
    somme_resultats = 0
    for _ in range(taille_echantillon):
        somme_resultats += lancer_de()
    return somme_resultats / taille_echantillon

# Exemple d'utilisation
taille = 1000
moyenne_obs = moyenne_echantillon(taille)
print(f"Moyenne observée sur {taille} lancers : {moyenne_obs}")
```

Commentaires :

- La fonction `lancer_de()` simule le résultat d'un lancer de dé.
- La fonction `moyenne_echantillon()` génère un échantillon de la taille spécifiée et calcule la moyenne des résultats.
- En augmentant `taille_echantillon`, la moyenne observée devrait se rapprocher de l'espérance théorique d'un dé équilibré, qui est 3.5.

Remarque 4. Les expérimentations avec Python ou un tableur permettent d'étudier la distance entre la moyenne d'un échantillon simulé et l'espérance de la variable aléatoire. Il est possible de calculer la proportion des cas où l'écart entre la moyenne de l'échantillon (m) et l'espérance (μ) est inférieur ou égal à $2\sigma/\sqrt{n}$. Ces activités renforcent l'intuition des élèves sur la loi des grands nombres.