Terminale Spé Maths – Chapitre P-02

LOI BINOMIALE

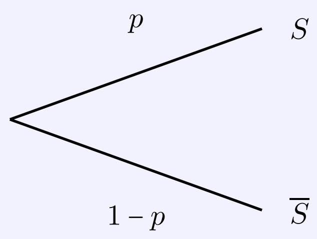


Table des matières

I Présentation

1) Espérance, variance et écart-type d'une variable aléatoire

DÉFINITION

Soit n un entier naturel et soit X une variable aléatoire qui prend les valeurs x_1, x_2, ..., x_n.
On pose, pour tout entier naturel i compris entre 1 et n, $p_i = P(X = x_i)$.

- L'espérance mathématique de X est $E(X) = \sum_{i=1}^{n} p_i x_i = p_1x_1 + p_2x_2 + \ldots + p_nx_n$.
- La variance de X est $V(X) = \sum_{i=1}^{n} p_i(x_i - E(X))^2$.
- L'écart-type de X est $\sigma(X) = \sqrt{V(X)}$.

REMARQUE

La formule de König-Huyghens vue en classe de Première donne une autre expression de la variance :
$V(X) = E(X^2) - E(X)^2$

PROPRIÉTÉS

Soient a et b deux réels et X une variable aléatoire. Alors :

$$E(aX + b) = aE(X) + b$$

$$V(aX + b) = a^2V(X)$$

$$\sigma(aX + b) = |a|\sigma(X)$$

REMARQUE

Ces propriétés ont été démontrées en Première.

PROPRIÉTÉ admise

Soient X et Y deux variables aléatoires. Alors :

$$E(X + Y) = E(X) + E(Y)$$

Si de plus X et Y sont **indépendantes**, alors :

$$V(X + Y) = V(X) + V(Y)$$

REMARQUES

- Ces deux formules seront démontrées au prochain chapitre.
- Ces deux formules peuvent être étendues à n variables (indépendantes pour la variance).

2) Problématique

On lance vingt fois de suite, dans les mêmes conditions, un dé bien équilibré à 6 faces.

1) Quelle est la probabilité d'obtenir 20 fois la face 6 sur les 20 lancers ?

2) Quelle est la probabilité d'obtenir 0 fois la face 6 sur les 20 lancers ?

3) Quelle est la probabilité d'obtenir 4 fois la face 6 sur les 20 lancers ?

Correction :

1) Soit X la variable aléatoire égale au nombre de fois qu'apparait la face 6 sur les 20 lancers.

Un seul chemin donne 20 fois la face 6 : celui du haut. Donc $P(X=20)=\left(\dfrac{1}{6}\right)^{20}$.

2) Un seul chemin donne 0 fois la face 6 : celui du bas. Donc $P(X=0)=\left(\dfrac{5}{6}\right)^{20}$.

3) Tous les chemins dans l'arbre donnant exactement 4 fois la face 6 sur les 20 lancers ont la même probabilité, car ils contiennent autant de branches qui vont vers le haut (4 branches car 4 succès) que de branches qui vont vers le bas (16 branches car 16 succès) : cette probabilité vaut $\left(\dfrac{1}{6}\right)^{4}\times\left(\dfrac{5}{6}\right)^{16}$.

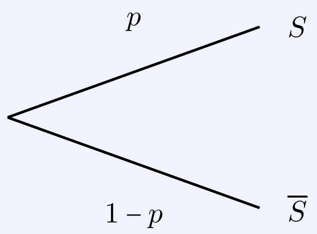
Le nombre de ces chemins correspond au nombre de combinaisons de 4 éléments parmi 20, soit $\binom{20}{4}$.

Ainsi, $P(X=4)=\binom{20}{4}\times\left(\dfrac{1}{6}\right)^{4}\times\left(\dfrac{5}{6}\right)^{16}=4845\times\left(\dfrac{1}{6}\right)^{4}\times\left(\dfrac{5}{6}\right)^{16}\approx 0,2022$.

II Épreuve, loi et schéma de Bernoulli

1) Épreuve de Bernoulli

DÉFINITION



Soit p un réel appartenant à l'intervalle $[0;1]$.

On appelle **épreuve de Bernoulli** toute expérience aléatoire n'admettant que deux issues, appelées généralement **succès** S et **échec** $\overline{S}$, et de probabilités respectives p et $q=1-p$.

EXEMPLES

- Lancer une pièce de monnaie équilibrée et savoir si PILE est obtenu est une épreuve de Bernoulli de succès S : « Pile a été obtenu », de probabilité $p=0,5$. L'échec est donc $\overline{S}$: « Face a été obtenu ».
- Interroger une personne dans la rue en France et lui demander si elle est gauchère est une épreuve de Bernoulli de succès S : « La personne est gauchère », de probabilité $p\approx 0,13$.

2) Loi de Bernoulli

DÉFINITION

Soit p un réel de l'intervalle $[0;1]$ et X une variable aléatoire.
On réalise une épreuve de Bernoulli dont le succès S a pour probabilité p.
On dit que X est une **variable aléatoire de Bernoulli** lorsqu'elle est à valeurs dans $\{0;1\}$, où la valeur 1 est attribuée au succès.
On dit alors que X suit la **loi de Bernoulli** de paramètre p.
Autrement dit, on a $P(X=1)=p$ et $P(X=0)=1-p$.
On peut résumer la loi de Bernoulli par le tableau suivant :

x_i	1	0
$P(X=x_i)$	p	$1-p$

PROPRIÉTÉ

Soit p un réel de l'intervalle $[0;1]$, et soit X une variable aléatoire qui suit la loi de Bernoulli de paramètre p.
L'espérance mathématiques de X est $E(X) = p$.
La variance de X est $V(X) = p(1-p)$.

DÉMONSTRATION

$$\begin{aligned} E(X) &= 1 \times P(X=1) + 0 \times P(X=0) \\ &= 1 \times p + 0 \times (1-p) \\ &= p. \end{aligned}$$

$$\begin{aligned} V(X) &= P(X=1) \times (1-E(X))^2 + P(X=0) \times (0-E(X))^2 \\ &= p(1-p)^2 + (1-p)(0-p)^2 \\ &= p(1)p)^2 + p^2(1-p) \\ &= p(1-p)(1-p+p) \\ &= p(1-p). \end{aligned}$$

REMARQUE

L'écart-type de X est alors $\sigma(X) = \sqrt{p(1-p)}$.

EXEMPLE

Un jeu de dé est tel que le joueur gagne lorsque le 6 sort et perd dans le cas contraire.
Soit S l'événement « le 6 sort » ; alors si le dé n'est pas pipé, $P(S) = \dfrac{1}{6}$ et $P(\overline{S}) = 1 - \dfrac{1}{6} = \dfrac{5}{6}$.
La variable aléatoire qui prend la valeur 1 si le 6 sort et la valeur 0 dans les cinq autres cas suit une loi de Bernoulli :

x_i	1	2
$P(X = x_i)$	$\frac{1}{6}$	$\frac{5}{6}$

3) Schéma de Bernoulli

DÉFINITION

Soit n un entier naturel non nul.
Un **schéma de Bernoulli** d'ordre n est la répétition de n épreuves de Bernoulli **identiques** et **indépendantes**.

EXEMPLE

L'expérience présentée en introduction du chapitre est un schéma de Bernoulli d'ordre 20.

III Loi binomiale

1) Définition de la loi binomiale

DÉFINITION

Soient n un entier naturel non nul et p un réel de l'intervalle $[0;1]$.
Soit X la variable aléatoire comptant le nombre de succès obtenu lors la répétition de n épreuves de Bernoulli identiques et indépendantes et dont p est la probabilité du succès.
On dit alors que X suit la **loi binomiale** de paramètres n et p. On peut la noter $\mathcal{B}(n;p)$.

PROPRIÉTÉ

Soient n un entier naturel non nul et p un réel de l'intervalle $[0;1]$.
Soit X une variable aléatoire qui suit la loi binomiale de paramètres n et p.
Alors pour tout entier k compris entre 0 et n, on a :

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

DÉMONSTRATION

L'événement $X = k$ est associé à l'ensemble des chemins dans l'arbre pour lesquels il y a exactement k succès et donc $n-k$ échecs. Chacun de ces chemins a une probabilité égale au produit des probabilités inscrites sur les branches qui constituent ce chemin, c'est-à-dire $p^k(1-p)^{n-k}$. Or il y a $\binom{n}{k}$ chemins de ce type. D'où $P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$.

EXEMPLE

Une société spécialisée dans l'audience des médias estime que 19,8% des Français vont regarder le premier match de la France pour l'Euro 2016 de Football. Enzo contacte 20 de ses amis et on estime que le nombre d'amis d'Enzo est assez grand pour assimiler ce tirage à un tirage avec remise de 20 amis.
Soit X la variable aléatoire qui donne parmi ces 20 personnes, le nombre de celles qui vont regarder le premier match de l'Euro 2016. (On donnera des valeurs approchées à 10^{-6} près)

1) Justifier que X suit une loi binomiale dont on précisera les paramètres.

2) Déterminer la probabilité qu'une personne contactée sur deux va regarder le match.

3) Déterminer la probabilité qu'au moins 2 personnes sur les 20 vont regarder le match.

4) Déterminer, à l'aide de la calculatrice, la probabilité qu'entre 5 et 15 personnes vont regarder le match.

2) Espérance et variance de la loi binomiale

PROPRIÉTÉ

Soient n un entier naturel non nul et p un réel de l'intervalle $[0;1]$.
Si X est une variable aléatoire qui suit la loi binomiale $\mathcal{B}(n;p)$, alors :

$$E(X) = np \qquad \text{et} \qquad Var(X) = np(1-p)$$

DÉMONSTRATION

X est égale au nombre de succès obtenus au cours de la répétition des n épreuves de Bernoulli.
Pour tout i allant de 1 à n, on note X_i la variable aléatoire qui prend pour valeur 1 si on obtient un succès à la i-ème épreuve et 0 sinon.
On a donc $X = X_1 + X_2 + ... + X_n = \sum_{i=1}^{n} X_i$.

Les variables aléatoires X_i suivent toutes la loi de Bernoulli de paramètre p, donc pour tout $1 \leqslant i \leqslant n$, $E(X_i) = p$ et $V(X_i) = p(1-p)$.

Ainsi, $E(X) = E(X_1 + X_2 + ... + X_n) = E(X_1) + E(X_2) + ... + E(X_n) = n \times p = np$

De plus, les variables X_1, X_2, ..., X_n étant indépendantes, on a :
$V(X) = V(X_1 + X_2 + ... + X_n) = V(X_1) + V(X_2) + ... + V(X_n) = n \times p(1-p) = np(1-p)$.

REMARQUE

L'écart-type de X est alors $\sigma(X) = \sqrt{np(1-p)}$.

EXEMPLE

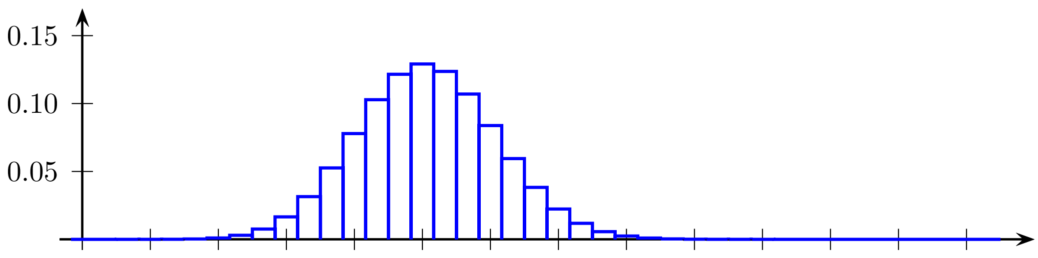
On reprend l'exemple précédent.
Calculer l'espérance de X et interpréter le résultat obtenu.

IV Introduction à l'échantillonnage

1) Représentation graphique d'une loi binomiale

Soit X une variable aléatoire suivant la loi binomiale de paramètres $n \in \mathbb{N}^*$ et $p \in [0;1]$.
On peut représenter la loi binomiale par un **diagramme en bâtons**, où les valeurs entières k prises par X sont représentées sur l'axe des abscisses, et la hauteur de chaque bâton est proportionnelle à la probabilité $P(X = k)$.

Par exemple, si X suit la loi binomiale de paramètres $n = 40$ et $p = 0,38$, on obtient le diagramme ci-dessous :



L'espérance de X est $E(X) = np = 40 \times 0,38 = 15,2$.
On remarque que les valeurs de X forment une représentation en forme de « cloche » qui est centrée sur l'espérance.
Plus l'écart-type est grand, plus les valeurs de X sont dispersées autour de $E(X)$: ainsi,le diagramme est plus étalé et « aplati ».

2) Intervalle de fluctuation

PROPRIÉTÉ

Soit n un entier naturel non nul, et soient α et p deux réels appartenant à l'intervalle à $[0\,;1]$.
Soit X une variable aléatoire qui suit la loi binomiale $\mathcal{B}(n\,;p)$.
Il existe un intervalle I non vide tel que $P(X \in I) \geqslant 1-\alpha$.

DÉMONSTRATION

L'intervalle $I=[0\,;n]$ convient. En effet, pour tout $\alpha \in [0\,;1]$, $1-\alpha \leqslant 1$, et $P(0 \leqslant X \leqslant n)=1$, donc on a bien $P(X \in I) \geqslant 1-\alpha$.

REMARQUES

- Cet intervalle n'est pas unique.
- Dans la pratique, on cherchera souvent un intervalle dont l'amplitude est la plus petite possible.

EXEMPLE

Soit X une variable aléatoire suivant la loi binomiale de paramètres $n=30$ et $p=0,23$.
Déterminer le plus petit entier a tel que $P(X \leqslant a) \geqslant 0,8$.

Correction :
On obtient à l'aide de la calculatrice $P(X \leqslant 8) \approx 0,76$ et $P(X \leqslant 9) \approx 0,87$.
Donc le plus petit entier a tel que $P(X \leqslant a) \geqslant 0,8$ est $a=9$.

DÉFINITION

Soit X une variable aléatoire suivant la loi binomiale de paramètres n et p, où n est un entier naturel non nul et p un réel de $[0\,;1]$.
Soient α un réel de $[0\,;1]$ et a et b des réels.

Un intervalle $[a\,;b]$ tel que $P(a \leqslant X \leqslant b) \geqslant 1-\alpha$ est appelé **intervalle de fluctuation au seuil de** $1-\alpha$ associé à X.

EXEMPLE

Dans l'exemple précédent, l'intervalle $[0\,;9]$ est un intervalle de fluctuation, au seuil de $0,2$ associé à X.

3) Un exercice d'application

Une société de vente en ligne de matériel de jardinage propose à ses clients des lots de 80 asperseurs.
Une étude à montré que 5% des asperseurs vendus sont défectueux.
On choisit un lot au hasard et on note X la variable aléatoire qui compte le nombre d'asperseurs défectueux sur les 80 du lot.

1) Modéliser la situation par une loi binomiale dont on précisera les paramètres.

2) **Cas particulier :** la société souhaite déterminer le plus petit entier naturel k tel que $P(X \leqslant k) \geqslant 0,95$.

- **a)** Déterminer cette valeur de k à l'aide de la calculatrice.
- **b)** Interpréter le résultat dans le contexte de l'exercice.

3) **Généralisation** : la société souhaite déterminer le plus petit entier naturel k tel que $P(X \leqslant k) \geqslant S$, où $S \in]0:1[$.

a) Justifier l'existence d'un entier k quelle que soit la valeur de S dans $]0;1[$.

b) Compléter le programme en Python suivant pour répondre au souhait de la société :

```python
from math import *
def seuil(S):
    k=0
    p=0.95**80
    while p<S:
        k=.........
        p=........
    return k
```

c) Que renvoie la fonction `seuil` pour $S = 0,99$? Pour $S = 0,999$?

4) Intervalle de fluctuation centré

DÉFINITION

Soit X une variable aléatoire suivant la loi binomiale de paramètres n et p, où n est un entier naturel non nul et p un réel de $[0;1]$.
Un **intervalle de fluctuation centré** de X **au seuil de** 95% est l'intervalle $[a;b]$ où :
a est le plus petit entier tel que $P(X \leqslant a) > 0,025$;
b est le plus petit entier tel que $P(X \leqslant b) \geqslant 0,975$

EXEMPLE

Soit X une variable aléatoire suivant la loi binomiale de paramètres $n = 100$ et $p = 0,25$.

1) Tabuler sur la calculatrice la loi binomiale correspondante.

2) Déterminer l'intervalle de fluctuation centré au seuil de 95%.

5) Prise de décision

La détermination d'un intervalle de fluctuation permet de prendre une décision lorsque l'on fait une hypothèse sur une proportion dans une population.
En effet, en faisant une hypothèse sur la proportion p d'un caractère dans une population, on peut déterminer un intervalle de fluctuation I à 95% de la fréquence de ce caractère dans un échantillon de taille n.
Ainsi, on établira une règle de décision : si la fréquence observée dans l'échantillon n'appartient pas à I, comme cela n'a qu'une probabilité de 0,05 de se produire, alors on rejettera l'hypothèse faite sur p, avec un risque de se tromper de 5%.
On en déduit la propriété suivante :

PROPRIÉTÉ

On considère une population dans laquelle on **suppose** que la proportion d'un caractère est p. Après expérience, on **observe** f comme fréquence de ce caractère dans un échantillon de taille n.
Soit l'hypothèse : « La proportion de ce caractère dans la population est p ».
Si I est l'intervalle de fluctuation à 95% de dans un échantillon de taille n, alors :
- Si $f \notin I$: on rejette cette hypothèse au seuil de risque de 5%.
- Sinon, on ne rejette pas cette hypothèse au seuil de risque de 5%.

EXEMPLE

Dans un article, un journaliste affirme que 42% des jeunes qui aiment la lecture préfèrent lire le soir.
On interroge au hasard 140 jeunes qui aiment lire et on assimile le sondage à un tirage successif avec remise : on observe que 49 jeunes déclarent lire le soir.
Peut-on mettre en doute, au seuil de 5%, l'affirmation du journaliste ?

Correction :
On fait l'hypothèse que la proportion de jeunes aimant lire le soir est $p = 0,42$.
On interroge 140 jeunes au hasard. On note X la variable aléatoire égale au nombre de jeunes préférant lire le soir. X suit une loi binomiale de paramètres $n = 140$ et $p = 0,42$.

Dans la table des probabilités cumulées de X, on recherche :
- Le plus petit entier a tel que $p(X \leqslant a) > 0,025$. On trouve $a = 47$.
- Le plus petit entier b tel que $p(X \leqslant b) \geqslant 0,975$. On trouve $b = 70$.

Comme la taille de l'échantillon est $n = 140$, l'intervalle de fluctuation à 95% de la fréquence correspondant à X est $\left[\frac{47}{140}; \frac{70}{140}\right]$ soit environ $[0,33; 0,5]$.

La fréquence observée est $f = \frac{49}{140} = 0,35$, donc f **appartient** à l'intervalle $[0,33; 0,5]$.

On ne peut donc pas remettre en question l'hypothèse que la proportion des jeunes lecteurs préférant lire le soir est $p = 0,42$. On ne met donc pas en doute l'affirmation du journaliste.

EXERCICE

On a lancé 100 fois de suite une pièce de monnaie et le côté PILE est apparu 65 fois. Peut-on estimer que la pièce est équilibrée ?

Correction :
Tester si la pièce est équilibrée c'est tester l'hypothèse que la sortie de PILE a une probabilité égale à 0,5.
On observe dans un tableur (=binompdf(100,0.5)) que le plus petit entier tel que $p(X \leqslant a) > 0,025$ est $a = 40$, et le plus petit tel que $p(X \leqslant b) \geqslant 0,975$ est $b = 60$.
Sachant que $n = 100$, l'intervalle de fluctuation cherché est donc $I = [0,4; 0,6]$.
La fréquence d'apparition de piles sur les 100 lancers effectués est de 0,65 qui n'appartient pas à I donc on rejette l'hypothèse que la pièce soit équilibrée avec une probabilité inférieure à 5% de se tromper.