1ère S - Chapitre 9 : LOI BINOMIALE. ÉCHANTILLONNAGE.

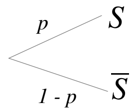
Textes officiels (30 septembre 2010) :

CONTENUS	CAPACITÉS ATTENDUES	COMMENTAIRES
Probabilités Épreuve de Bernoulli, loi de Bernoulli. Schéma de Bernoulli, loi binomiale (loi du nombre de succès). Coefficients binomiaux, triangle de Pascal.	• Reconnaître des situations relevant de la loi binomiale. • Calculer une probabilité dans le cadre de la loi binomiale.	La représentation à l'aide d'un arbre est privilégiée : il s'agit ici d'installer une représentation mentale efficace. On peut ainsi : faciliter la découverte de la loi binomiale pour des petites valeurs de n $(n \leqslant 4)$; introduire le coefficient binomial $\binom{n}{k}$ comme nombre de chemins de l'arbre réalisant k succès pour n répétitions ; établir enfin la formule générale de la loi binomiale.

1ère S - Chapitre 9 : LOI BINOMIALE. ÉCHANTILLONNAGE.

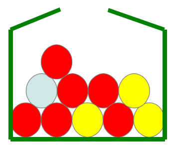
I. Épreuve de Bernouilli

Définition :	Une **épreuve de Bernoulli** est une **expérience aléatoire à deux issues** : • l'une que l'on nomme « **SUCCÈS** », que l'on note S, et dont la probabilité d'apparition est p ; • l'autre nommée « **ÉCHEC** », que l'on note $\overline{S}$ et dont la probabilité d'apparition est $1 - p$.



Exemple : Une urne contient 6 boules rouges, 3 boules jaunes et 1 boule bleue, toutes indiscernables. Avant de jouer, on mise un euro. On tire une boule au hasard et on obtient :

– 0 euro si elle est rouge ;
– 1 euro si elle est jaune ;
– 5 euros si elle est bleue.



On peut donc définir comme **SUCCÈS** le fait de tirer la boule bleue.

Il s'agit donc d'une épreuve de Bernoulli avec pour S : « tirer une boule bleue » et pour $\overline{S}$: « tirer une boule qui n'est pas bleue » (donc une boule rouge ou une jaune).

Définition :	Dans une épreuve de Bernoulli de paramètre p, la variable aléatoire X, prenant la valeur 1 si S se produit et la valeur 0 sinon, suit la loi de probabilité ci-contre : Son espérance est $E(X) = p$, sa variance est $V(X) = p(1 - p)$. On dit alors que X est une variable de Bernoulli de paramètre p ou encore que X suit la loi de Bernoulli de paramètre p.

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

Exemple : Dans l'exemple précédent, l'épreuve de Bernoulli a pour loi de paramètre $p = \frac{1}{10}$.

Donc la loi de Bernoulli de paramètre $p = \frac{1}{10}$ est définie par le tableau suivant :

Issue	S	$\overline{S}$
Probabilité	$\frac{1}{10}$	$\frac{9}{10}$

II. Schéma de Bernoulli

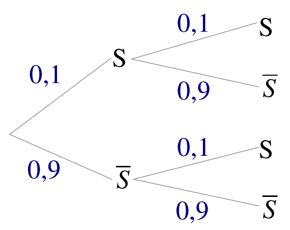
Définition :	Un **schéma de Bernoulli** est la répétition d'une même épreuve de Bernoulli dans des conditions d'indépendance : c'est-à-dire que l'issue d'une épreuve n'influe pas sur les autres.

Exemple 1 : On reprend l'exemple du paragraphe 1 et après avoir tiré une boule, on la replace dans l'urne avant d'en choisir une seconde.

On répète alors plusieurs fois cette expérience.

Le fait de replacer la boule dans l'urne assure l'indépendance entre deux tirages.

Le schéma de Bernoulli peut être illustré par un arbre comportant autant d'étapes qu'il y a de tirages.



Avec deux tirages

Exemple 2 : Dans l 'exemple précédent, on effectue trois fois le tirage d'une boule.

Comme on l'a vu précédemment, chaque tirage est une épreuve de Bernoulli.

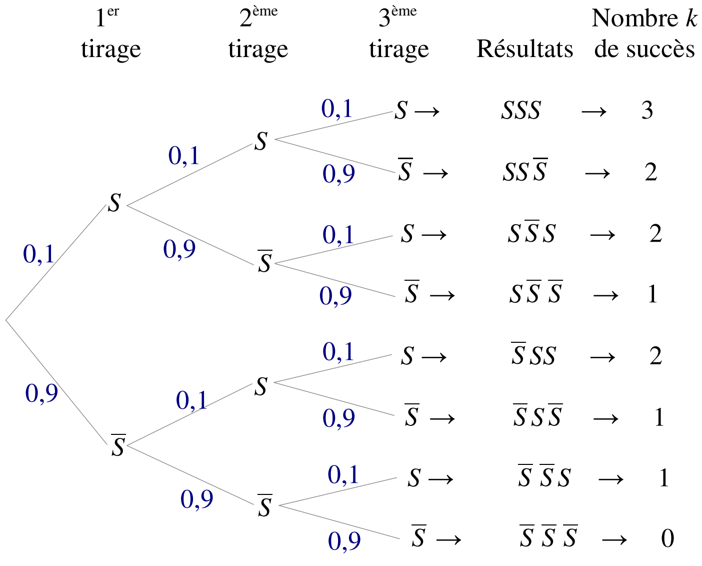
La probabilité d'obtenir la liste $SS\overline{S}$ est :

$$P\left(SS\,\overline{S}\right)=P\left(S\right)\times P\left(S\right)\times P\left(\overline{S}\right)$$
$$=0{,}1\times 0{,}1\times 0{,}9$$
$$=0{,}09.$$

On note X la variable aléatoire qui compte le nombre de succès sur 3 tirages.

La probabilité d'obtenir deux succès est :

$$\begin{aligned} P(X=2) &= P\left(SS\,\overline{S}\right)+P\left(S\,\overline{S}\,S\right)+P\left(\overline{S}\,SS\right)\\ &= 0{,}009 \quad + \quad 0{,}009 \quad + \quad 0{,}009\\ &= 0{,}027. \end{aligned}$$



Avec trois tirages

On peut alors construire de la même manière le tableau donnant la loi de probabilité de la variable aléatoire X :

Nombre k de succès	0	1	2	3
Probabilité P(X = k)	0,729	0,243	0,027	0,001

Remarque : Comme toute loi de probabilité, la somme des probabilités est égale à 1.

Ici : P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) = 0,729 + 0,243 + 0,027 + 0,001 = 1.

De plus : $E\left(X\right)=0\times 0{,}729+1\times 0{,}243+2\times 0{,}027+3\times 0{,}001=0{,}3$.

Cela signifie qu'en répétant un grand nombre de fois l'expérience (de trois lancers), on obtient en moyenne 0,3 fois le succès.

Définition :	On considère un **schéma de Bernoulli** de ***n* épreuves,** $n\in\mathbb{N}^*$, représenté par un arbre. Pour tout entier k tel que $0\leqslant k\leqslant n$, on note $\binom{n}{k}$ le nombre de chemins de l'arbre réalisant k succès lors des n répétitions. Par convention, on pose $\binom{0}{0}=1$. $\binom{n}{k}$ est aussi appelé un « coefficient binomial » et se lit « k parmi n ».

Exemple : À l'aide de l'arbre réalisé précédemment, on obtient :

- $\binom{3}{0}$ nombre de chemins amenant à 0 succès S en 3 répétitions ; il y en a un seul donc $\binom{3}{0}=1$;
- $\binom{3}{1}$ nombre de chemins amenant à 1 succès S en 3 répétitions ; il y en a un seul donc $\binom{3}{1}=3$;
- $\binom{3}{2}$ nombre de chemins amenant à 2 succès S en 3 répétitions ; il y en a un seul donc $\binom{3}{2}=3$;
- $\binom{3}{3}$ nombre de chemins amenant à 3 succès S en 3 répétitions ; il y en a un seul donc $\binom{3}{3}=1$.

III. Propriété des $\binom{n}{k}$

Propriété :	Pour tout entier $n \geqslant 0$, $\binom{n}{0} = 1$ et $\binom{n}{n} = 1$.

Démonstration : Un seul chemin mène à 0 succès lors de n répétitions : c'est $\overline{S}\,\overline{S}\ldots\overline{S}$;
Un seul chemin mène à n succès lors de n répétitions : c'est $SS\ldots S$.

Propriété :	Pour tous entiers naturels n et k tels que $0 \leqslant k \leqslant n$, $\binom{n}{k} = \binom{n}{n-k}$.

Démonstration : Si $n=0$, $0 \leqslant k \leqslant n$ implique $k=0$ et donc l'égalité est vérifiée ;

Si $n>0$, alors sur l'arbre du schéma de n épreuves de Bernoulli, $\binom{n}{k}$ désigne le nombre de chemins menant à k succès, donc aussi le nombre de chemins réalisant $n-k$ échecs parmi n. C'est-à-dire $\binom{n}{n-k}$.

Propriété :	Pour tous entiers naturels n et k tels que $0 \leqslant k \leqslant n-1$, $\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$

Démonstration : Sur l'arbre du schéma de n épreuves de Bernoulli, les chemins qui conduisent à k succès sont :

- ceux qui conduisent à k - 1 succès lors des n - 1 premières répétitions et à un succès lors de la n-ième répétition. Il y en a $\binom{n-1}{k-1}$;
- ceux qui conduisent à k succès lors des n - 1 premières répétitions et à un échec lors de la n-ième répétition. Il y en a $\binom{n-1}{k}$. d'où : $\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$.

IV. Loi binomiale

Définition :	On considère un **schéma de Bernoulli** constitué de ***n* épreuves** dont la probabilité de succès est p. On désigne par X la variable aléatoire associée au nombre de succès lors de ces n épreuves. Alors : P(X = k) = $\binom{n}{k} p^k (1-p)^{n-k}$ où k prend les valeurs 0, 1, 2, … , n. Son espérance est $E(X) = np$. Sa variance est $V(X) = np(1-p)$ et son écart-type est $\sigma(X) = \sqrt{np(1-p)}$. La loi de probabilité de la variable aléatoire X est appelée **loi binomiale de paramètre *n* et *p*** et on la note B(n ; p).

Exemple : Comme on l'a vu dans l'exemple 2, les trois tirages consécutifs constituent un schéma de Bernoulli et on peut alors associer une loi binomiale à ces trois tirages consécutifs : B(n = 3 ; p = 0,1).

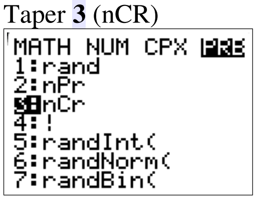
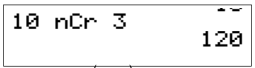
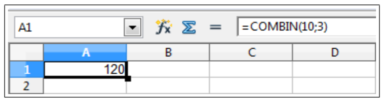
On a donc $E(X) = np = 3 \times 0{,}1 = 0{,}3$.

Et $\sigma(X) = \sqrt{np(1-p)} = \sqrt{3 \times 0{,}1 \times 0{,}9} = \sqrt{0{,}27} \approx 0{,}52$.

Annexes TICE : Avec la calculatrice ou un tableur

• Calculer un coefficient binomial (combinaison)

On veut calculer le coefficient binomial $\binom{10}{3}$, c'est-à-dire le nombre de **combinaisons** de « 3 parmi 12 », ou encore, le nombre de manières différentes de choisir « 3 parmi 12 ».

Casio	TI	Tableur
• Taper 10 • Sélectionner OPTN • Taper F6 pour ▷ , puis F3 pour PROB • Choisir F3 pour nCr • Taper 3 puis EXE  • Donc $\binom{10}{3}$=120.	• Taper 10 • Sélectionner **MATH** • ▶ ▶ ▶ **PRB** • Taper **3** (nCR)   • Donc $\binom{10}{3}$=120.	• Fonction **COMBIN** Syntaxe : =COMBIN(n;k) 

Source : Manuel Indice Maths – 1ère ES-L – Éditions Bordas - 2011

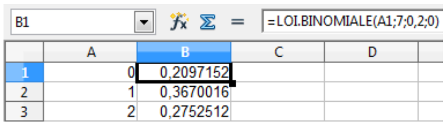
• Calcul pratique de P(X = k) et P(X ⩽ k)

	Casio	TI	OpenOffice/LibreOffice	Excel
Syntaxe	Touche **OPTN** puis choisir **STAT**, puis **DIST**, puis **BINM**, puis **Bpd** ou **Bcd**	Menu distrib (**2nde var**), puis choisir **binomFdp** (ou **binomFrép**)	Fonction **LOI.BINOMIALE**	
$P(X = k)$	BinomialPD(k,n,p)	binomFdp(n,p,k)	=LOI.BINOMIALE(k;n;p;0)	=LOI.BINOMIALE(k;n;p;FAUX)
$P(X \leqslant k)$	BinomialCD(k,n,p)	binomFrep(n,p,k)	=LOI.BINOMIALE(k;n;p;1)	=LOI.BINOMIALE(k;n;p;VRAI)

Source : Manuel Indice Maths – 1ère ES-L – Éditions Bordas - 2011

• Tableau donnant la loi de probabilité de X

On peut afficher un tableau donnant la loi binomiale de X.
Pour cela, on utilise les calculatrices en mode TABLE :
- on choisit pour x les entiers de 0 jusque n (« pas » de 1) ;
- pour Y1, on utilise les fonctions de lois binomiales en remplaçant k par X.

Casio	TI	Tableur
Dans le menu **TABLE**, entrer en Y1 : Y1=BinomialPD(*X*,*n*,*p*) BinomialPD est accessible par **SHIFT 4** (**CATALOG**) en prenant le soin de remplacer *n* et *p* par les valeurs voulues. Régler la table avec **SET** (**F5**) puis l'afficher avec **TABLE** (**F6**).	Dans le menu ***f*(*x*)**, entrer en Y1 : Y1=BinomFdp(*n*,*p*,*X*) BinomFdp est accessible par 2nde 0 (**catalog**) en prenant le soin de remplacer *n* et *p* par les valeurs voulues. Régler la table avec **déf table** (**2nde fenêtre**) puis l'afficher avec **table** (**2nde Graphe**).	On entre 0, 1, 2, … , jusque *n*, puis la formule calculant les probabilités dans la colonne suivante avec la fonction : =LOI.BINOMIALE(*Cellule*;*n*;*p*;FAUX)  ci-dessus, $n = 7$ et $p = 0{,}2$.

Source : Manuel Indice Maths – 1ère ES-L – Éditions Bordas - 2011

V. Échantillonnage

Activité p. :

1. Étudier une hypothèse à partir d'un échantillon

On pose une hypothèse : dans une population donnée de taille N, on **suppose** qu'un caractère est présent dans la proportion p.

Pour juger de la validité de cette hypothèse, on effectue un prélèvement, au hasard et avec remise, un échantillon de taille n , $n \leqslant N$, et on observe la fréquence f du caractère observé dans cet échantillon.

Population	Proportion théorique p	Taille N
Échantillon	Proportion dans l'échantillon f	Taille n

On cherche à savoir si la fréquence observée (appelée aussi fréquence empirique) f est suffisamment éloignée de p, dans un sens ou dans l'autre, pour **pouvoir rejeter l'hypothèse**.

2. Intervalle de fluctuation : rappel

On a déterminé en classe de seconde un intervalle de fluctuation d'une proportion d'un caractère à 95 %.

Propriété : (vue en 2nde)

Si p est la proportion d'un caractère dans une population, avec $0{,}2 \leqslant p \leqslant 0{,}8$, alors pour un échantillon de taille $n \geqslant 25$, la fréquence f du caractère dans l'échantillon appartient à l'intervalle $\left[p - \frac{1}{\sqrt{n}} \,;\, p + \frac{1}{\sqrt{n}}\right]$ avec une probabilité d'au moins 95 %.

On va améliorer ce résultat en utilisant une loi binomiale.

3. Intervalle de fluctuation à 95 %

La proportion de la population présentant le caractère étudié est noté p.

Propriété : La variable aléatoire X qui compte le nombre d'individus de l'échantillon qui présente le caractère étudié, suit une loi binomiale de paramètre n et p : $\mathcal{B}(n\,;p)$.

Pour limiter l'influence des valeurs aberrantes, on va écarter les valeurs extrêmes et créer ainsi un nouvel intervalle : on partage l'intervalle de l'échantillon [0 ; n] en trois sous-intervalles :

$$[0\,; a-1]\ ,\ [a\,; b] \text{ et } [b+1\,; n],$$

de façon à ce que X prenne ses valeurs dans les deux intervalles extrêmes avec une probabilité proche de 2,5 % mais sans jamais la dépasser.

Définition : L'intervalle de fluctuation à 95 % d'une fréquence correspondant à la réalisation, sur un échantillon de taille n, d'une variable aléatoire X de loi binomiale $\mathcal{B}(n\,;p)$, est l'intervalle $\left[\frac{a}{n}\,;\frac{b}{n}\right]$ défini par :

- a est le plus petit entier tel que $P(X \leqslant a) > 2{,}5\ \%$;
- b est le plus petit entier tel que $P(X \leqslant b) \geqslant 97{,}5\ \%$.

On obtient alors que $P(a \leqslant X \leqslant b) \geqslant 95\ \%$.
L'intervalle de fluctuation est aussi appelé « **intervalle de confiance** ».

VI. Un exemple

La proportion des personnes ayant les yeux marrons dans la population française est 0,34.

On veut déterminer un intervalle de fluctuation de la fréquence f des personnes ayant les yeux marrons dans des échantillons de taille 100.

Réponse :

- On définit la variable aléatoire X égale au nombre de personnes ayant les yeux marrons dans un échantillon de taille 100. En assimilant le choix d'une personne au hasard dans cet échantillon à un tirage avec remise, on peut supposer que X suit une loi binomiale de paramètre $n = 100$ et $p = 0{,}34$. X prend alors les valeurs de 0 à 100.

- Avec la calculatrice ou un tableur, on dresse la liste de ces valeurs que prend X. On en profite pour faire calculer les valeurs cumulées (ci-contre dans la colonne C).

B2 f_x =LOI.BINOMIALE(A2;100;0,34;0)

	A	B	C	D	E
1	n	X			
2	0	9,00313E-19	9,00313E-19		
3	1	4,63798E-17	4,72801E-17		
4	2	1,18268E-15	1,22996E-15		

Avec Excel.

- On va partager l'ensemble des valeurs de X en trois parties :
  - A : valeurs comprises entre 0 et a – 1 (a entier) ;
  - B : valeurs comprises entre a et b (entier) ;
  - C : Valeurs comprises entre b + 1 et n.

25	23	0,00530263	1,2%
26	24	0,008764069	2,0%
27	25	0,013725063	3,4%

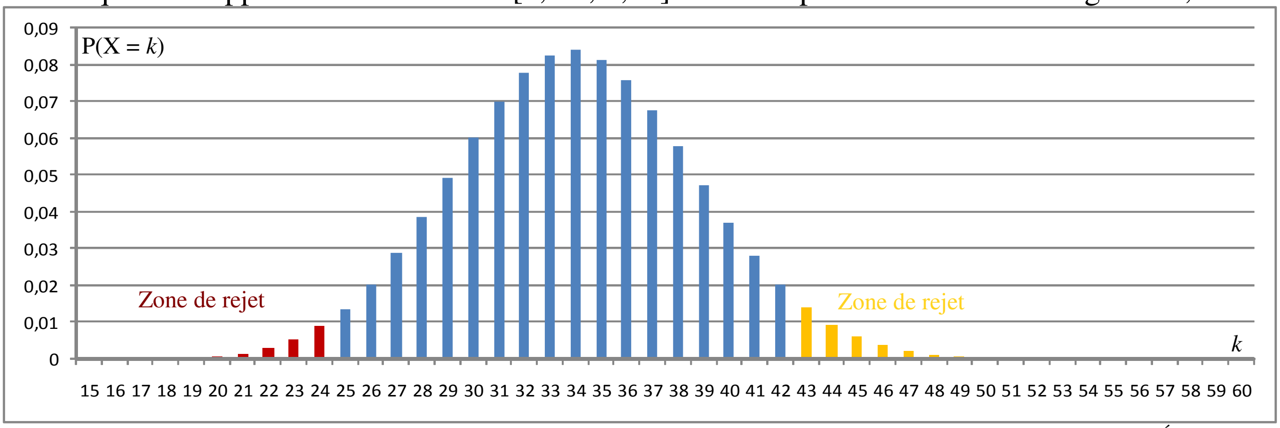
$P(X \leqslant k) \leqslant 2{,}5\,\%$ pour $k \leqslant 25$.

43	41	0,028007795	94,2%
44	42	0,020268268	96,2%
45	43	0,014083518	97,6%

$P(X \geqslant k) \geqslant 97{,}5\,\%$ pour $k \geqslant 44$.

- On détermine ensuite a et b de façon à ce que la probabilité que X appartienne à A et à C soit inférieure à 0,025 : on aura alors un intervalle [a ; b] tel que la probabilité que X soit compris entre a et b soit au moins égale à 0,95.

Ici, $a = 25$ et $b = 43$. Donc l'intervalle [a ; b] est [25 ; 43].
La fréquence f appartient à l'intervalle [0,25 ; 0,43] avec une probabilité au moins égale à 0,95.



Source : Manuel Indice Maths – 1ère ES-L – Éditions Bordas - 2011

4. Prendre des décisions

Règle de décision :

- si $f \notin \left[\frac{a}{n} ; \frac{b}{n}\right]$, alors **on rejette** l'hypothèse selon laquelle la proportion du caractère dans la population est p, *au risque de 5 %*.

- si $f \in \left[\frac{a}{n} ; \frac{b}{n}\right]$, alors **on ne peut pas rejeter** cette hypothèse.

Remarque : « *au risque de 5 %* » signifie que la probabilité de rejeter l'hypothèse, alors qu'elle est vraie, est inférieure à 5 %.

Exemple : Un laboratoire annonce qu'un médicament sauve 40 % des patients atteint d'une maladie rare.

Pour contrôler cette affirmation, l'autorité de surveillance demande le test sur 100 patients atteints de cette maladie.

Soit X le nombre de malades sauvés par ce médicament dans un échantillon aléatoire de malades et assimilé à un tirage avec remise de taille 100.

1. Quelle loi suit X ?

2. Déterminer les plus petits entiers a et b tels que $P\left(X \leqslant a\right) \geqslant 0{,}025$ et $P\left(X \leqslant b\right) \geqslant 0{,}975$.

3. Énoncer la règle de décision permettant de rejeter ou non l'hypothèse $p = 0{,}40$ selon la valeur de la fréquence f des malades sauvés dans l'échantillon.

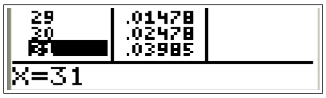
4. Sur les 100 malades auxquels on a administré ce traitement, on en a sauvé 30.
Au seuil de risque 5 %, que peut-on dire de l'annonce faite par le laboratoire ?

Réponse :

1. L'épreuve de Bernoulli « un patient est sauvé ou non par le médicament » est répétée 100 fois de manière identique et indépendante, donc X suit la loi binomiale de paramètres $n = 100$ et $p = 0{,}4$.

2. À l'aide de la calculatrice, on construit la table des valeurs $P\left(X \leqslant k\right)$

Avec une TI : Dans ***f(x)=*** 



et 

3. L'intervalle de fluctuation à 95 % de la fréquence dans les échantillons de taille 100 est [0,32 ; 0,50].
D'où la règle de décision : si f appartient à l'intervalle [0,32 ; 0,50], l'hypothèse $p = 40$ est acceptable, sinon elle est rejetée, au seuil de 5 %.

4. La fréquence observée est de 0,30 et $0{,}30 \notin \left[0{,}32\ ;\ 0{,}50\right]$, donc, au seuil de risque de 5 %, on rejette l'hypothèse selon laquelle le médicament sauve 40 % des malades.

Source : Manuel Indice Maths – 1ère ES-L – Éditions Bordas – 2011