

LOI BINOMIALE

I Schéma de Bernoulli

Définition : une épreuve de Bernoulli est une expérience aléatoire qui admet exactement deux issues que l'on peut nommer « succès » et « échec » et que l'on peut noter S et $\bar{S}$.

On appelle alors loi de Bernoulli une loi de probabilité pour laquelle la probabilité d'obtenir un succès est $p(S) = p$ et la probabilité d'un échec est $p(\bar{S}) = 1 - p$.

On dit alors que p est le paramètre de cette loi.

Ex : on lance un dé et s'intéresse à la sortie du nombre 6.

Cette expérience a deux issues : S « Le résultat est 6 » et $\bar{S}$.

C'est donc une épreuve de Bernoulli de paramètre $p = \frac{1}{6}$

Associons le nombre 1 au succès et le nombre 0 à l'échec. La variable aléatoire X qui suit une loi de Bernoulli de paramètre p peut alors être représentée par le tableau suivant :

x_i	1	0
p_i	p	$1 - p$

Propriété : Soit X une variable aléatoire suivant une loi de Bernoulli de paramètre p .

On a $E(X) = p$ et $V(X) = p(1 - p)$.

Preuve : $E(X) = p \times 1 + (1 - p) \times 0$

$$= p$$

$$V(X) = \sum_{i=1}^n p_i (x_i - E(X))^2 \quad (\text{somme pondérée des carrés des écarts à la moyenne})$$

$$= (1 - p)^2 \times p + (0 - p)^2 \times (1 - p)$$

$$= (1 - p)^2 \times p + p^2 \times (1 - p)$$

$$= p(1 - p)(1 - p + p)$$

$$= p(1 - p)$$

Définition : un schéma de Bernoulli d'ordre n est la répétition de n épreuves de Bernoulli identiques (de même paramètre p notamment) et indépendantes (la réalisation de l'épreuve suivante ne dépend pas du résultat des épreuves précédentes).

II Loi binomiale

Définition : considérons un schéma de Bernoulli de paramètres n et p et notons X la variable aléatoire égale au nombre de succès au cours des n épreuves de Bernoulli. La loi de probabilité de X est appelée loi binomiale de paramètres n (nombre d'expériences) et p (probabilité d'un succès)

On note alors $X \hookrightarrow \mathcal{B}(n, p)$

Ex : On lance 3 fois de suite une pièce truquée dont la probabilité d'obtenir pile est $p \in [0 ; 1]$. On s'intéresse au nombre de fois où l'on obtient Pile.

Soit X la variable aléatoire correspondant au nombre de fois où Pile est obtenu au cours de ces 3 épreuves.

$X \hookrightarrow \mathcal{B}(3 ; p)$ car les lancers de pièces sont identiques et indépendants.

► Pour $(X = 0)$, 1 chemin est possible.

D'où $p(X = 0) = (1 - p) \times (1 - p) \times (1 - p) = (1 - p)^3$

► Pour $(X = 1)$, 3 chemins sont possibles.

Or, ces trois chemins ont tous la même probabilité :

$$p \times (1 - p) \times (1 - p) = p(1 - p)^2$$

D'où $p(X = 1) = 3p(1 - p)^2$

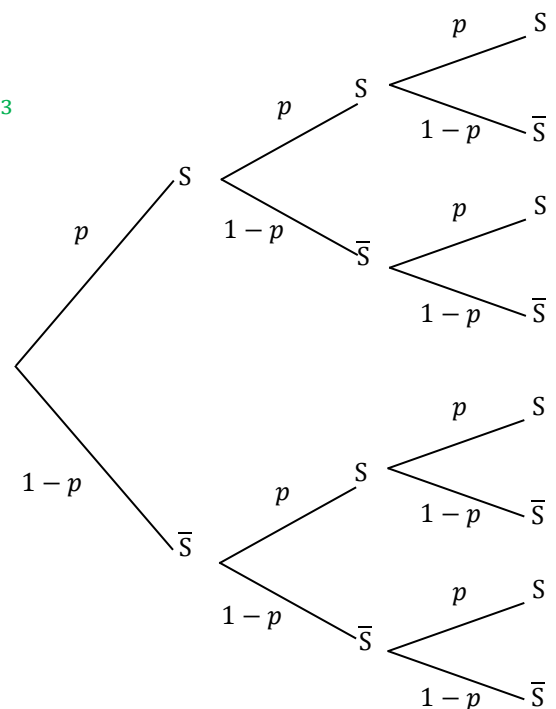
► Pour $(X = 2)$, 3 chemins sont possibles.

Or, ces trois chemins ont tous la même probabilité :

$$p \times p \times (1 - p) = p^2(1 - p)$$

D'où $p(X = 2) = 3p^2(1 - p)$

► Pour $(X = 3)$, 1 chemin est possible. D'où $p(X = 3) = p \times p \times p = p^3$



Propriété : Soit X une variable aléatoire, $X \hookrightarrow \mathcal{B}(n, p)$. Pour tout entier k compris entre 0 et n , on a :

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

A
U Preuve : L'évènement $(X = k)$ est composé des chemins contenant k succès et $n - k$ échecs.

P
R Ces chemins sont au nombre de $\binom{n}{k}$.

O
G Et la probabilité de chacun d'eux est $p^k (1 - p)^{n-k}$.

R
A Donc $p(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$

M
M
E

Ex : Un test d'aptitude consiste à poser une série de 10 questions indépendantes. Pour chacune d'entre elles, 3 réponses sont proposées dont une seule est correcte. Eric répond à chaque fois au hasard.

On note X la variable aléatoire correspondant au nombre de bonnes réponses d'Eric.

0° Quelle est la loi suivie par X ?

1° Quelle est la probabilité qu'Eric obtienne 8 bonnes réponses ? (en %, au centième)

2° Quelle est la probabilité qu'Eric obtienne 4 ou moins de 4 bonnes réponses ? (en %, au centième)

3° Quelle est la probabilité qu'Eric obtienne au moins 8 bonnes réponses ? (en %, au centième)

0° $X \hookrightarrow \mathcal{B}\left(10; \frac{1}{3}\right)$ car les questions sont identiques et indépendantes.

$$1^\circ p(X = 8) = \binom{10}{8} \left(\frac{1}{3}\right)^8 \left(\frac{2}{3}\right)^2$$

$$\approx 0,0030 \approx 0,3 \%$$

⇒ La calculatrice permet d'effectuer directement (de vérifier) les calculs du type $p(X = k)$:



- Aller dans « distrib » (en appuyant sur « 2nde » puis « var »)
- Choisir « A : binomFdp(»
- Indiquer la valeur de n pour « nbreEssais », la valeur de p et la valeur de k pour « valeur de x ».
- Valider « Coller » en appuyant sur « entrer », et valider à nouveau.

2° La question posée revient à calculer $p(X \leq 4)$.

⇒ La calculatrice permet d'effectuer directement les calculs du type $p(X \leq k)$:



- Aller dans « distrib » (en appuyant sur « 2nde » puis « var »)
- Choisir « B : binomFRép(»
- Indiquer la valeur de n pour « nbreEssais », la valeur de p et la valeur de k pour « valeur de x ».
- Valider « Coller » en appuyant sur « entrer », et valider à nouveau.

D'où $p(X \leq 4) \approx 0,7869 \approx 78,69 \%$

3° La question posée revient à calculer $p(X \geq 8)$.

$$p(X \geq 8) = 1 - p(X < 8)$$

$$= 1 - p(X \leq 7)$$

$$\approx 0,0034 \approx 0,34 \%$$

Remarques : 1. On a $p(X \geq 8) = p(X = 8) + p(X = 9) + p(X = 10)$

$$= \binom{10}{8} \left(\frac{1}{3}\right)^8 \left(\frac{2}{3}\right)^2 + \binom{10}{9} \left(\frac{1}{3}\right)^9 \left(\frac{2}{3}\right)^1 + \binom{10}{10} \left(\frac{1}{3}\right)^{10} \left(\frac{2}{3}\right)^0$$

$$= 45 \times \frac{1}{3^8} \times \frac{4}{3^2} + 10 \times \frac{1}{3^9} \times \frac{2}{3} + 1 \times \frac{1}{3^{10}} \times 1 = \frac{201}{3^{10}}$$

2. La probabilité d'obtenir entre 3 et 7 bonnes réponses correspond à :

$$p(3 \leq X \leq 7) = p(X \leq 7) - p(X \leq 2)$$

Propriété : Soit X une variable aléatoire telle que $X \hookrightarrow \mathcal{B}(n, p)$. On a $E(X) = n \times p$.

$\Rightarrow$ C'est le nombre **moyen** de succès lors des n expériences (preuve plus tard).

Ex : dans l'exemple précédent, un candidat qui répond au hasard obtient $E(X) = 10 \times \frac{1}{3} \approx 3,33$ bonnes réponses en moyenne

Exercice : On estime que la probabilité qu'un professeur de mathématiques commette une erreur durant une heure de cours est de 1%. A combien de cours un élève doit-il assister pour que la probabilité de voir son professeur commettre au moins une erreur dépasse 90% ?

Soit n le nombre d'heures de cours auxquelles un élève assiste et soit X la variable aléatoire correspondant **au nombre d'erreurs observées durant ces n heures de cours**.

On peut dire que $X \hookrightarrow \mathcal{B}(n ; 0; 01)$.

On cherche n tel que $p(X \geq 1) \geq 0,9$

Etape facultative : $1 - p(X \leq 0) \geq 0,9$

$$1 - p(X = 0) \geq 0,9$$

$$1 - \binom{n}{0} 0,01^0 0,99^n \geq 0,9$$

$$1 - 0,99^n \geq 0,9$$

$$1 - 0,9 \geq 0,99^n \text{ c'est-à-dire } 0,1 \geq 0,99^n$$

La table de valeurs de la calculatrice indique que c'est le cas dès lors que $n \geq 230$.

$\Rightarrow$ On a besoin de la fonction logarithme népérien pour résoudre ce problème sans avoir besoin de la table de valeurs

III Une application aux statistiques

Prenons un exemple : la direction d'un casino annonce que ses machines à sous sont réglées pour que 6% des parties soient gagnantes.

Un joueur joue 210 fois et obtient 6 succès alors qu'en moyenne il pouvait en espérer **12 ou 13**.

Peut-il remettre en question les chiffres annoncés par le casino... ?

Notons X la variable aléatoire indiquant le nombre de succès lorsqu'on joue 210 parties.

$$X \hookrightarrow \mathcal{B}(210 ; 0,06)$$

On peut par exemple vérifier qu'il y a 80% de chances d'obtenir entre 9 et 17 succès :

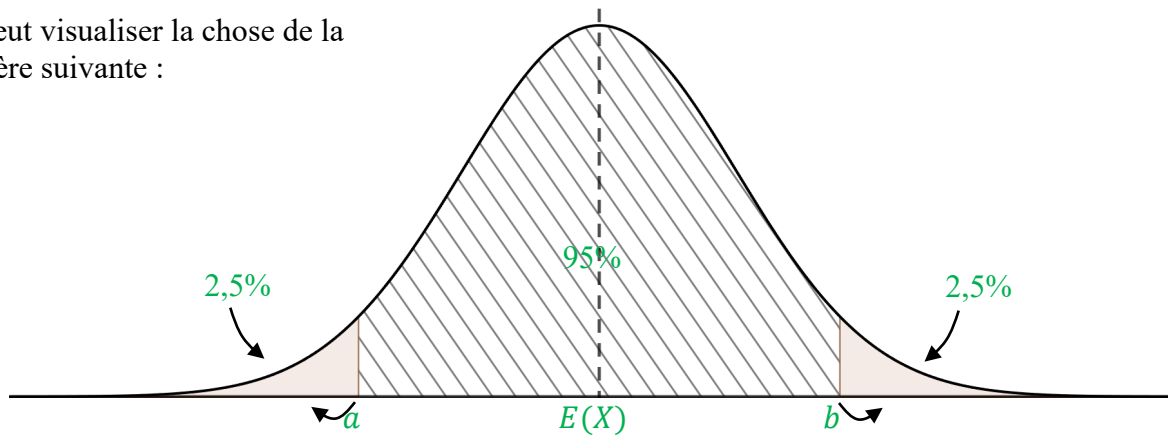
$$p(9 \leq X \leq 17) = p(X \leq 17) - p(X \leq 8) \approx 0,8056$$

Cherchons quelque chose de similaire mais avec une fiabilité d'au moins 95%.

Autrement dit, on cherche à déterminer a et b tels que $p(a \leq X \leq b) \geq 0,95$.

Pour obtenir ces 95%, il faut donc exclure 5% de résultats possibles ; et tant qu'à faire les 5% les plus rares. On va donc exclure 2,5% les moins fréquents « à gauche » et 2,5% moins fréquents « à droite » pour garder les 95% les plus centraux. Mais il serait étonnant qu'on tombe pile sur « 2,5 % »...

On peut visualiser la chose de la manière suivante :



On cherche donc le plus grand nombre a tel que $p(a \leq X) \leq 0,025$ et le plus petit nombre b tel que $p(X \geq b) \leq 0,025$, autrement dit tel que $p(X \leq b) \geq 0,975$

On utilise alors la table de valeur de la calculatrice avec $f(x) = \text{binomFRép}(210, 0,06, X)$

Ainsi, on a $p(X \leq 5) \approx 0,0119$ et $p(X \leq 6) \approx 0,029$

$\Rightarrow$ on peut donc dire que $a = 5$

Et de même, on a $p(X \leq 19) \approx 0,9715$ et $p(X \leq 20) \approx 0,9844$

$\Rightarrow$ on peut donc dire que $b = 20$

Par conséquent, $p(5 \leq X \leq 20)$ dépasse les 95%.

Vérification : $p(5 \leq X \leq 20) = p(X \leq 20) - p(X \leq 4) \approx 0,9803$

En fait, dans cet exemple $p(6 \leq X \leq 19)$ convenait, mais ce n'est pas toujours le cas...

Autrement dit, si la machine était vraiment réglée sur 6% de succès, on devrait obtenir entre 5 et 20 succès dans 95% des cas.

Morale de l'histoire : obtenir 6 succès est une valeur tout à fait normale, au sens où elle dans les valeurs attendues à 95%.

$\Rightarrow$ On parle d' « intervalle de fluctuation au seuil de 95% ». (et on pourrait modifier ce seuil)

Remarques : 1. Une fois la valeur de a déterminée, on peut obtenir une valeur approchée de b plus rapidement, par symétrie par rapport à $E(X)$.

Lorsqu'on a trouvé $a = 5$, sachant que $E(X) = 12,6$ alors $b \approx 12,6 + (12,6 - 5) = 20,2$.

2. On peut ajouter que pour une loi binomiale, $\sigma = \sqrt{V(X)} = \sqrt{np(1-p)}$.

Par ailleurs, on trouve environ 68% des valeurs dans l'intervalle $[E(X) - \sigma ; E(X) + \sigma]$ et 95% dans l'intervalle $[E(X) - 2\sigma ; E(X) + 2\sigma]$.

Dans notre exemple, $\sigma = \sqrt{210 \times 0,06 \times 0,94} \approx 3,44$.

Et $[E(X) - 2\sigma ; E(X) + 2\sigma] \approx [12,6 - 2 \times 3,44 ; 12,6 + 2 \times 3,44] \approx [5,72 ; 19,48]$