Chapitre 6
Loi binomiale – Échantillonnage

I. Succession d'épreuves indépendantes

1) Univers d'une succession d'épreuves

Définition :

Lorsqu'une expérience aléatoire se compose d'une succession de n épreuves indépendantes E_1, E_2, E_3, … , E_n, **l'univers des issues possibles** est le produit cartésien $\Omega_1 \times \Omega_2 \times \Omega_3 \times \ldots \times \Omega_n$, où Ω_i désigne l'univers de l'épreuve E_i pour i allant de 1 à n.

Une **issue** de la succession d'épreuves est donc un n-uplet $(i_1 ; i_2 ; i_3 ; \ldots ; i_n)$ où i_p est une issue de E_p.

Remarque :

On représente cette situation par un arbre dans lequel un chemin correspond à une issue.

Exemple :

On considère l'expérience aléatoire qui consiste à lancer une pièce équilibrée et à noter sa face, puis à piocher une carte dans un jeu de huit cartes et noter sa hauteur.

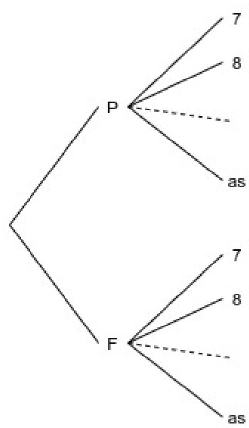
L'univers de la première épreuve est U = {P ; F}.

L'univers de la seconde épreuve est V = {7 ; 8 ; 9 ; 10 ; valet ; dame ; roi ; as}.

Les deux épreuves sont indépendantes l'une de l'autre, car le résultat de la hauteur de la carte ne dépend pas de la face obtenue avec la pièce.

L'univers Ω de la succession des deux épreuves précédentes est le produit cartésien des univers des expériences :

Ω = U × V = {(P ; 7) ; (P ; 8) ; … ; (P ; as) ; (F ; 7) ; (F ; 8) ; … ; (F ; as)}



Il y a 16 issues possibles. C'est le nombre de branches finales de l'arbre.

On lit les issues en parcourant les chemins de l'arbre.

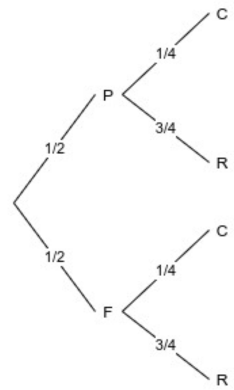
2) Calcul de probabilités

Propriété :

Lors d'une succession de n épreuves indépendantes, la probabilité d'une issue (i_1 ; i_2 ; i_3 ; … ; i_n) est égale au produit des probabilités de chacune des issues du n-uplet.

Exemple :

On considère une expérience aléatoire qui consiste à tirer une pièce équilibrée à Pile (P) ou Face (F), puis à choisir un jeton dans une urne contenant un jeton carré (C) et trois jetons ronds (R).



La probabilité de l'issue (F ; C) est alors égale à $\frac{1}{2}\times\frac{1}{4}=\frac{1}{8}$.

En utilisant la formule des probabilités totales, la probabilité d'obtenir un jeton carré est égale à $p(\mathrm{C}) = p(\mathrm{P}\ ;\ \mathrm{C}) + p(\mathrm{F}\ ;\ \mathrm{C}) = \frac{1}{2}\times\frac{1}{4}+\frac{1}{2}\times\frac{1}{4}=\frac{1}{4}$.

II. Schéma de Bernoulli

1) Épreuve de Bernoulli

Définition :

Une **épreuve de Bernoulli** est une expérience aléatoire à deux issues appelées « Succès » et « Échec ». On dit qu'une épreuve de Bernoulli est de paramètre p si la probabilité de l'issue « Succès » est p.

Exemple :

On lance un dé équilibré à six faces et on considère comme un succès d'obtenir un 1.

Cette expérience aléatoire est une épreuve de Bernoulli de paramètre $\frac{1}{6}$.

Remarque :

On utilisera communément la lettre q pour désigner la probabilité d'un échec.

« Succès » et « Échec » étant des événements contraires, on a donc $q=1-p$.

Dans l'exemple précédent, un échec a la probabilité $q=1-\frac{1}{6}=\frac{5}{6}$.

2) Loi de Bernoulli

Définition :

La variable aléatoire qui prend la valeur 1 en cas de succès et 0 en cas d'échec est appelé **variable aléatoire de Bernoulli**.

La loi de probabilité de cette variable aléatoire est appelé **loi de Bernoulli de paramètre *p***.

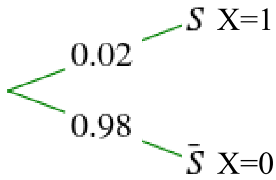
Exemple :

Dans une usine, la probabilité qu'un article fabriqué présente un défaut est 0,02.

Soit X la variable aléatoire qui vaut 1 si l'article présente un défaut et 0 sinon.

X suit une loi de Bernoulli de paramètre $p=0{,}02$.

x_i	0	1
$p(X=x_i)$	0,98	0,02



Propriété :

Soit X une variable aléatoire de Bernoulli de paramètre p.

- **L'espérance** de X est $E(X)=p$.
- La **variance** de X est $V(X)=p(1-p)$.
- **L'écart-type** est $\sigma(X)=\sqrt{p(1-p)}$.

Démonstrations :

$E(X)=0\times(1-p)+1\times p=p$ et $V(X)=0^2\times(1-p)+1^2\times p-p^2=p-p^2=p(1-p)$.

3) Schéma de Bernoulli

Définition :

Soit n un entier naturel non nul et p un nombre réel appartenant à l'intervalle $[0;1]$.

Un **schéma de Bernoulli** est une expérience consistant à répéter n fois, de manière **indépendante**, la même épreuve de Bernoulli.

Un schéma de Bernoulli a deux paramètres : n le nombre de répétitions de l'épreuve et p le paramètre de l'épreuve répétée.

Propriété :

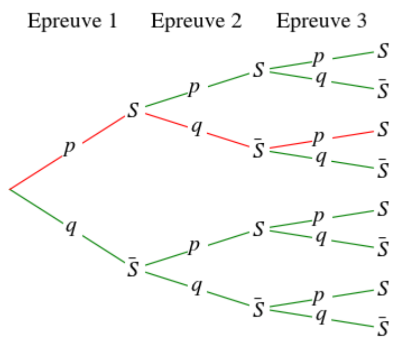
On peut représenter un schéma de Bernoulli de paramètre n et p par un **arbre de probabilité** à 2^n branches.

Les issues sont des n-uplets dont les n termes sont S pour « Succès » ou $\bar{S}$ pour « Échec ».

Exemple :

Ci-contre un schéma de Bernoulli pour $n=3$.

L'issue correspondant au chemin rouge peut être notée $(S, \overline{S}, S)$.



Propriété :

La probabilité d'une issue d'un schéma de Bernoulli s'obtient en faisant le produit des probabilités des issues obtenues à chaque épreuve de Bernoulli.

Exemple :

Si un schéma de Bernoulli a pour paramètre $n=4$ et $p=0{,}3$, alors l'issue $(S, \overline{S}, \overline{S}, S)$ a pour probabilité $0{,}3\times 0{,}7\times 0{,}7\times 0{,}3=0{,}0441$.

Remarque :

Si l'on définit X comme la variable aléatoire égale au nombre de succès dans ce schéma, X prend ses valeurs dans l'ensemble {0 ; 1 ; 2 ; … ; *n*}.

III. Loi binomiale

1) Définition

Définition :

Une expérience suit un schéma de Bernoulli de paramètre *n* et *p*.

k est un entier naturel tel que $0\leqslant k\leqslant n$.

On associe à l'expérience la variable aléatoire X qui donne le nombre total de succès.

La loi de probabilité de X est appelée **loi binomiale** de paramètre *n* et *p*.

On la note $\mathscr{B}(n\ ;p)$.

Propriété :

Si une variable aléatoire X suit une loi binomiale $\mathscr{B}(n\ ;p)$, alors pour tout entier *k* compris entre 0 et *n*, la probabilité que X soit égal à *k* est :

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

Démonstration :

On sait que tous les chemins comportant k succès sont équiprobables car, en faisant le produit des probabilités des n issues de chaque épreuve de Bernoulli, on obtient k facteurs p (pour k succès) et $n-k$ facteurs $(1-p)$ (pour $n-k$ échecs). Leur probabilité est donc $p^k(1-p)^{n-k}$.

Il suffit ensuite de compter les chemins menant à k succès : il y en a $\binom{n}{k}$. On obtient donc :

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

Exemple :

On lance 3 fois un dé équilibré à six faces et on considère comme un succès d'obtenir un 6.

En nommant X la variable aléatoire donnant le nombre de succès, X suit une loi binomiale $\mathscr{B}(n,p)$.

On peut calculer $p(X=1)=\binom{3}{1}\times\left(\frac{1}{6}\right)^1\times\left(\frac{5}{6}\right)^2=3\times\frac{1}{6}\times\frac{25}{36}=\frac{75}{216}=\frac{25}{72}$.

Nombre de chemins avec un seul succès

Probabilité d'un succès

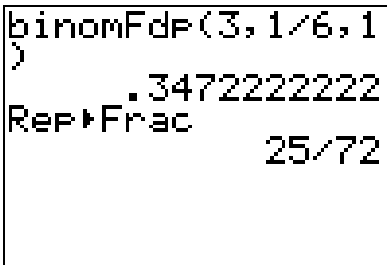
Probabilité de deux échecs

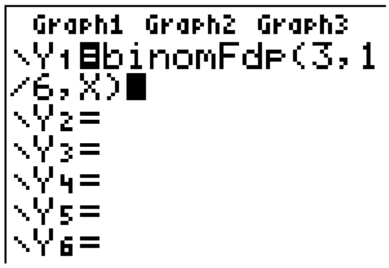
La probabilité d'obtenir exactement un 6 en trois lancers est $\frac{25}{72}$.

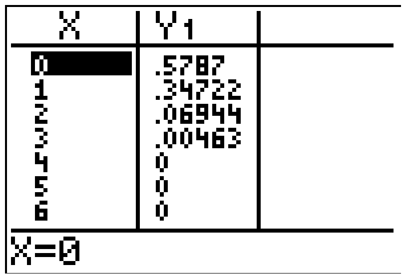
On peut de même déterminer la loi de probabilité de X :

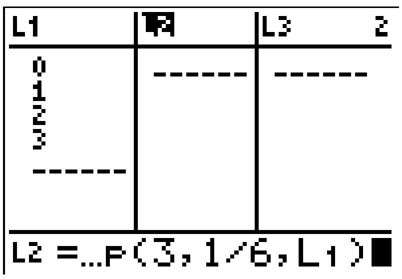
x_i	0	1	2	3
$p(X=x_i)$	$\frac{125}{216}$	$\frac{75}{216}$	$\frac{15}{216}$	$\frac{1}{216}$

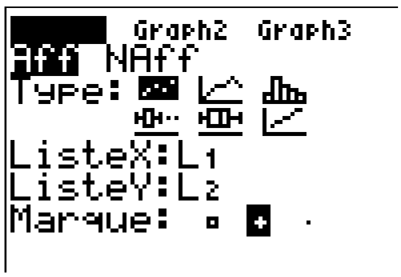
Calculatrice :

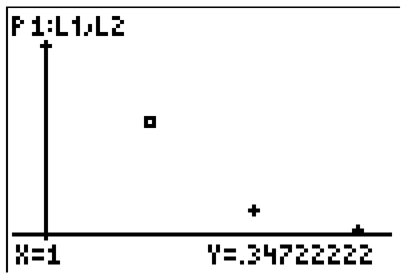


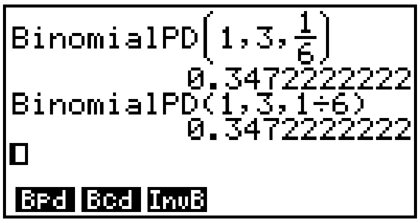


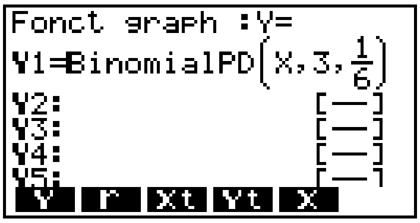


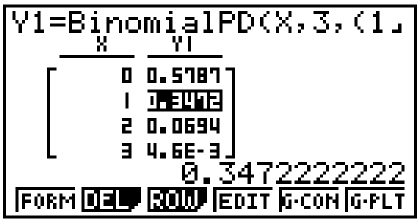


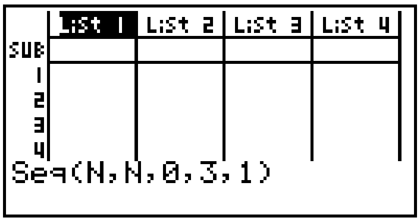


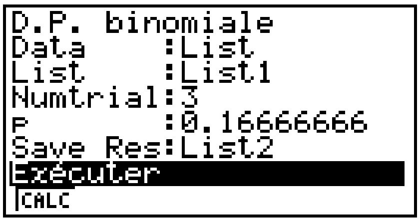


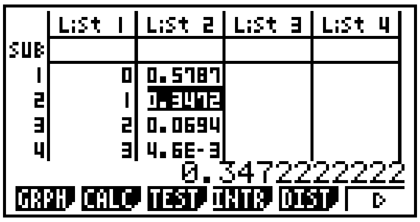


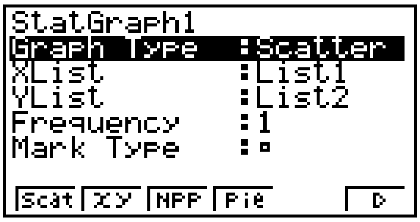


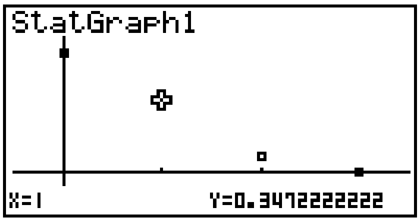






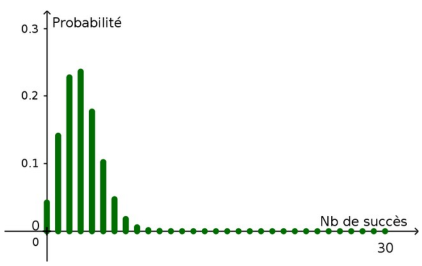




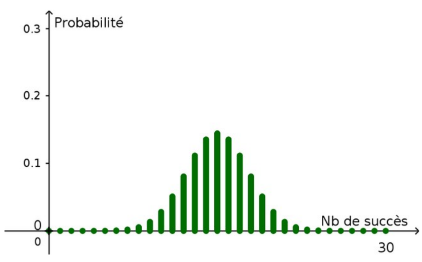


2) Représentation graphique

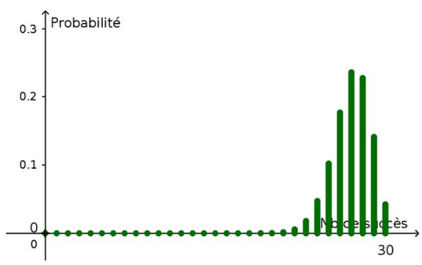
Quelques exemples de représentation de la loi binomiale pour $n = 30$.



$p = 0,1$



$p = 0,5$



$p = 0,9$

3) Indicateurs

Propriétés :

Soit une variable aléatoire X qui suit une loi binomiale $\mathscr{B}(n\ ;p)$.

- $E(X)=np$
- $V(X)=np(1-p)$
- $\sigma(X)=\sqrt{np(1-p)}$

Exemple :

En reprenant l'exemple précédent, on peut calculer l'espérance grâce à la loi de probabilité de X.

$E(X)=0\times\frac{125}{216}+1\times\frac{75}{216}+2\times\frac{15}{216}+3\times\frac{1}{216}=\frac{108}{216}=\frac{1}{2}$. On vérifie que $n\times p=3\times\frac{1}{6}=\frac{1}{2}$.

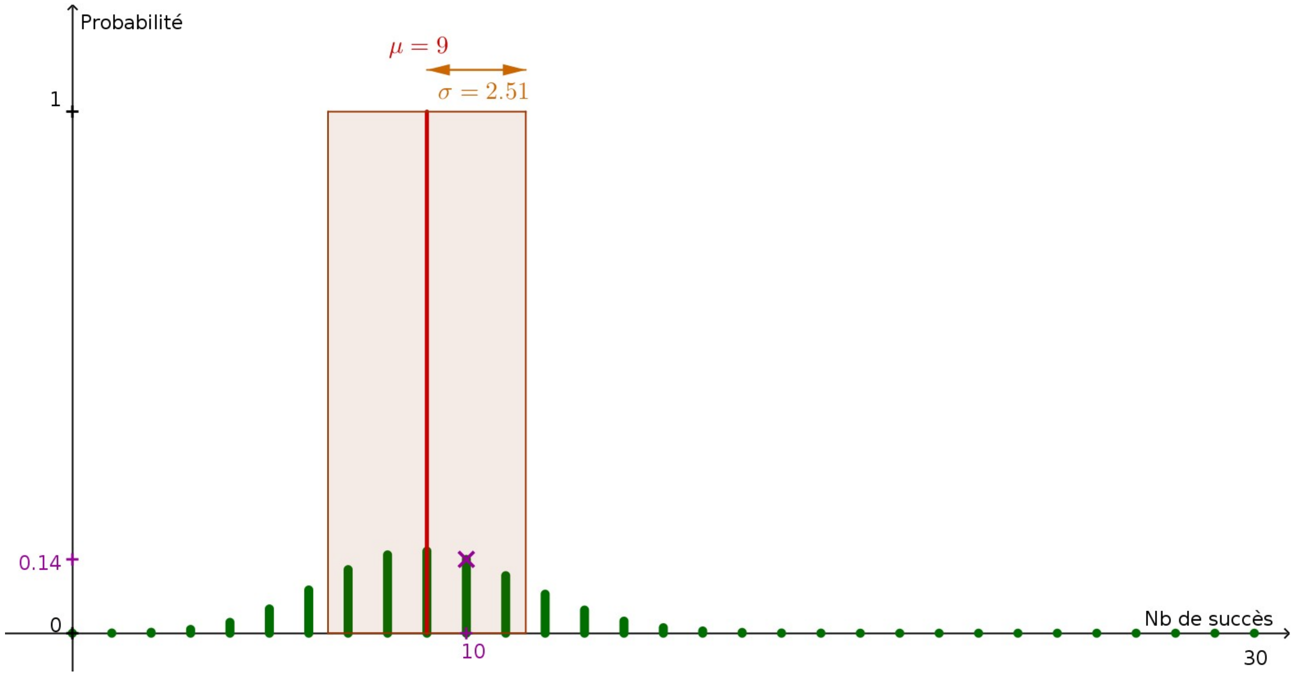
Remarque :

On peut remarquer que la formule de l'espérance peut s'expliquer sans calcul. En effet, chaque épreuve de Bernoulli a pour espérance de succès p donc, en la répétant n fois, on peut espérer obtenir en moyenne $n\times p$ succès.

Dans l'exemple précédent, le 6 sort avec la probabilité $\frac{1}{6}$, donc on peut espérer tripler ce résultat en triplant l'expérience, car ce sont des expériences successives et indépendantes et $E(X)=3\times\frac{1}{6}=\frac{1}{2}$.

Représentation graphique :

Exemple pour n = 30 et p = 0,3.



IV. Échantillonnage

1) Échantillon

Définition :

Un échantillon de taille n est obtenu en prélevant au hasard, **successivement** et **avec remise**, n éléments d'une population.

Remarques :

- Après prélèvement d'un échantillon, on s'intéresse à la valeur d'un caractère des éléments de la population. Nous nous restreignons ici à des populations dont le caractère étudié n'a que **deux valeurs possibles**.

- **Exemples de situations correspondants à un prélèvement d'échantillon** :
  - Prélever des pièces dans une production de manière identique et indépendante, noter à chaque fois si la pièce présente un défaut ou non et la remettre dans la production.
  - Lancer plusieurs fois un dé et noter à chaque fois si la face supérieure est un 6 ou non.
  - Lancer plusieurs fois de manière indépendante une pièce de monnaie et noter si elle affiche « pile » ou « face ».
  - Sortir au hasard de manière indépendante une boule dans une urne qui ne contient que des boules rouges et des boules d'autres couleurs et noter à chaque fois si elle est rouge ou non.
- Souvent, il n'y a pas de remise lors du prélèvement. Mais lorsque l'effectif total est très grand par rapport au nombre d'objets prélevés, on considère néanmoins que l'échantillon est constitué, au sens de la définition donnée, avec remise.

  Exemples de situations assimilées à un prélèvement d'échantillon :
  - Lors d'un prélèvement de pièces dans une production, après avoir constaté qu'une pièce a un défaut, il n'est pas envisageable de la remettre parmi l'ensemble des pièces produites.
  - Lors d'un sondage « sortie des urnes », on ne peut pas attendre que tout le monde ait voté et réunir toutes les personnes avant de choisir au hasard des individus pour obtenir un échantillon.

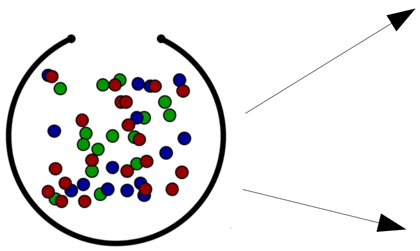
2) Intervalle de fluctuation

On souhaite donner les caractéristiques d'un échantillon à partir de celles connues de la population dans laquelle il a été prélevé. C'est ce qu'on nomme l'**échantillonnage**.

Exemple :

Urne avec 40 % de boules rouges.

Le caractère « rouge » a, dans la population, une proportion de $p=0{,}4$.



Premier échantillon de taille $n=15$.

Nombres de boules ayant le caractère « rouge » : $k=6$.

Fréquence observée : $f=\frac{6}{15}=0{,}4$.

Deuxième échantillon de taille $n=15$.

Nombres de boules ayant le caractère « rouge » : $k=4$.

Fréquence observée : $f=\frac{4}{15}\simeq 0{,}27$.

Propriété :

Soit une population dont une proportion ***p*** des éléments admet un caractère donné.

Dans un **échantillon de** taille ***n*** prélevé dans cette population, l'**effectif** des éléments qui présentent ce **caractère** est une variable aléatoire qui suit la **loi binomiale** de paramètres ***n*** et ***p***.

Démonstration :

- Le prélèvement d'un échantillon de taille n peut être assimilé à un tirage successif avec remise dès lors que l'effectif de la population est assez grand par rapport à n.

  On peut donc considérer que l'expérience est une répétition de n tirages identiques à deux issues : avoir ou ne pas avoir le caractère choisi.

- La variable aléatoire X donnant le nombre d'éléments qui ont le caractère choisi suit alors une loi binomiale de paramètre n (nombre de répétition d'un tirage) et p (probabilité de prélever un élément ayant le caractère choisi).

Remarque :

Pour cette loi binomiale de paramètre n et p, la somme des probabilités

$\sum_{k=0}^{n} p(X=k) = p(X=0) + p(X=1) + \ldots + p(X=n)$ vaut 1.

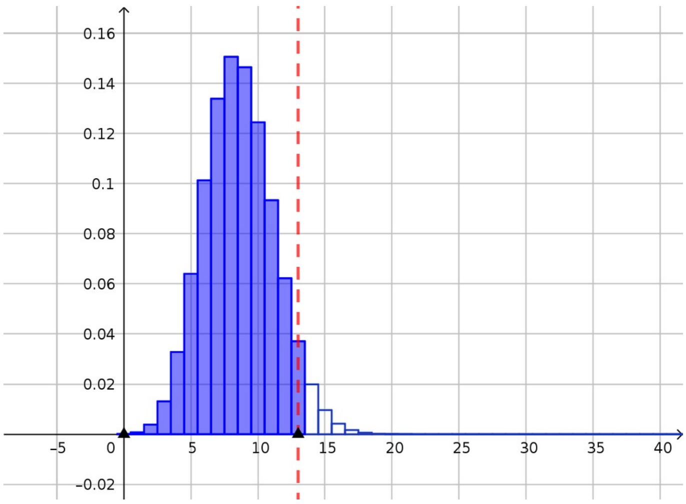
Définition :

Soit X une variable aléatoire suivant une loi binomiale, $\alpha \in]0\ ;\ 1[$ et a et b réels.

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

Exemple :

Pour X suivant la loi $\mathscr{B}(43\ ;\ 0{,}2)$ et $\alpha = 0{,}05$.



On a $p(X \leqslant 13) \approx 0{,}964$, donc $p(X \leqslant 13) \geqslant 0{,}95$.

Ainsi $[0\ ;\ 13]$ est un intervalle de fluctuation au seuil de $1 - \alpha = 0{,}95$ associé à X.

On est sûr, à au moins 95 %, qu'il n'y aura pas plus de 13 succès sur les 43 répétitions.

Remarque :

Lorsque l'on cherche à trouver un intervalle de fluctuation associé à une loi binomiale de paramètres n et p, suivant le contexte, on peut être amené à chercher des intervalles de la forme $[0\ ;\ b]$, $[a\ ;\ n]$ ou $[a\ ;\ b]$ « centré », de préférence les moins grands possibles.

Intervalle de fluctuation centré

Propriété :

Soit X une variable aléatoire suivant une loi binomiale, $\alpha \in]0\,;\,1[$ et a et b réels.

L'intervalle $[a\,;\,b]$ tel que a et b soient les plus petits entiers vérifiant respectivement :

$p(X \leqslant a) > \dfrac{\alpha}{2}$ et $p(X \leqslant b) \geqslant 1-\dfrac{\alpha}{2}$ est un **intervalle de fluctuation centré** (ou **bilatéral**) au seuil de $1-\alpha$ associé à X.

Remarque :

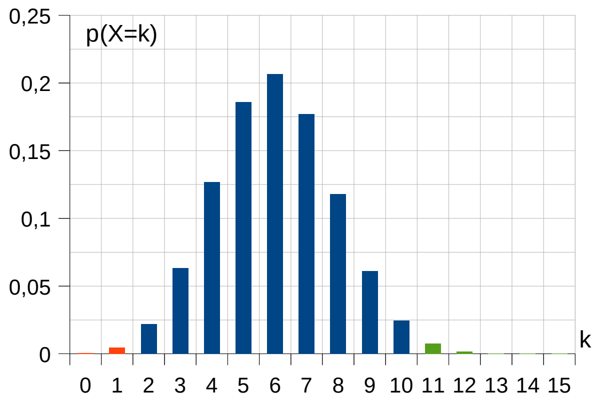
Pour un seuil de 0,95 :

Sur la représentation graphique ci-contre :

- La somme des probabilités des parties vertes est inférieure à 0,025.
- La somme des probabilités des parties rouges est inférieure à 0,025.

Pour que **la somme des probabilités des parties bleues soit supérieure où égale à 0,95 tout en étant la plus petite possible**, on détermine :

- a le plus petit entier tel que $p(X \leqslant a) > 0{,}025$.
- b le plus petit entier tel que $p(X \leqslant b) \geqslant 0{,}975$.



Loi binomiale de paramètre $n=15$ et $p=0{,}4$

Exemple :

Dans une urne contenant 40 % de boules rouges, on prélève un échantillon de taille $n=15$.

Soit X la variable aléatoire qui compte le nombre de boules rouges dans l'échantillon prélevé.

Les valeurs possibles de X sont les entiers compris entre 0 et 15.

La loi de probabilité de X est donnée par :

			Intervalle de fluctuation													
k	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$p(X=k)$	0,000	0,005	0,022	0,063	0,127	0,186	0,207	0,177	0,118	0,061	0,025	0,007	0,002	0,000	0,000	0,000
	0,027															
	0,991															

- $a=2$ car $p(X \leqslant 1) \simeq 0{,}005$ et $p(X \leqslant 2) \simeq 0{,}027$.
- $b=10$ car $p(X \leqslant 9) \simeq 0{,}966$ et $p(X \leqslant 10) \simeq 0{,}991$.

L'intervalle $[0{,}13\,;0{,}67]$ est un intervalle de fluctuation au seuil 0,95 associé à X.

Calculatrice :