Loi Binomiale

I) Répétition d'expériences identiques et indépendantes

La répétition d'expériences identiques consiste à répéter plusieurs fois de suite la même expérience aléatoire. Ces expériences aléatoires successives sont indépendantes si la réalisation de l'une ne modifie pas les probabilités des issues de l'autre.

Ex :

► On lance une pièce de monnaie trois fois de suite et on observe sur quelle face elle tombe. Le résultat "pile" ou "face" obtenu lors d'un lancer ne dépend pas du résultat des deux autres lancers.
Ces trois expériences aléatoires sont identiques et indépendantes.

► Un sac contient 4 boules numérotées de 1 à 4. On tire au hasard successivement deux boules en remettant à chaque fois la boule tirée dans le sac (**tirage avec remise**). On observe les numéros obtenus. Le résultat obtenu lors d'un tirage ne dépend pas du résultat de l'autre tirage.
On a là une répétition de deux expériences aléatoires identiques et indépendantes.

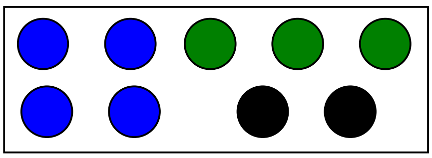
le tirage **avec remise** est indispensable pour que les expériences soient identiques !

a) modélisation d'une répétition :

Une répétition d'expériences aléatoires identiques et indépendantes peut être représentée par un arbre pondéré.
► une issue est une liste ordonnée de résultats représentée sur l'arbre par un chemin
► la probabilité d'une issue (de cette liste) est le produit des probabilités de chacun des résultats de la liste

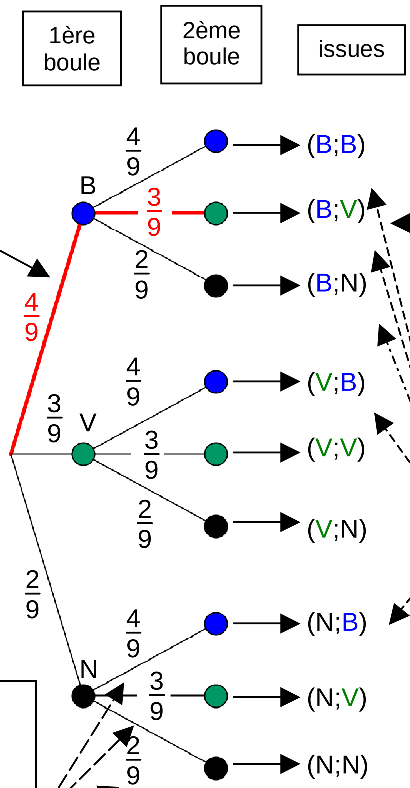
Ex : Une urne contient 4 boules bleues, 3 boules vertes et deux boules noires.
On tire **successivement** deux boules **avec remise**.

répétition d'expériences identiques et indépendantes !



Représentons cette répétition par un arbre <u>pondéré</u>.

on inscrit sur chaque branche la probabilité de l'événement correspondant !



le chemin tracé en rouge représente l'issue (B;V). La première boule obtenue est bleue et la seconde est verte.

la probabilité d'une issue est égale au produit des probabilités écrites sur les branches du chemin.

$$P(B;V) = \frac{4}{9} \times \frac{3}{9} = \frac{12}{81} = \frac{4}{27}$$

la probabilité d'un événement est **la somme des probabilités des issues** induisant la réalisation de l'événement.

Soit A l'événement «obtenir **au moins** une boule bleue dans le tirage»

$$P(A) = P(B;B) + P(B;V) + P(B;N) + P(V;B) + P(N;B)$$
$$= \frac{4}{9} \times \frac{4}{9} + \frac{4}{9} \times \frac{3}{9} + \frac{4}{9} \times \frac{2}{9} + \frac{3}{9} \times \frac{4}{9} + \frac{2}{9} \times \frac{4}{9}$$
$$= \frac{16}{81} + \frac{12}{81} + \frac{8}{81} + \frac{12}{81} + \frac{8}{81}$$
$$= \frac{56}{81}$$

la somme des probabilités sur les branches issues d'un même noeud est égale à 1.

$$\frac{4}{9} + \frac{3}{9} + \frac{2}{9} = \frac{9}{9} = 1$$

b) <u>exemple d'une variable aléatoire :</u>

Reprenons la répétition d'expériences précédente.
On considère la variable aléatoire X qui associe à chaque issue le nombre de boules vertes obtenues. La variable aléatoire prend donc pour valeurs 0 ; 1 et 2.

► L' événement (X=0) est réalisé par les issues (B;B),(B;N),(N;B),(N;N)

Donc $P(X = 0) = \frac{4}{9} \times \frac{4}{9} + \frac{4}{9} \times \frac{2}{9} + \frac{2}{9} \times \frac{4}{9} + \frac{2}{9} \times \frac{2}{9} = \frac{36}{81} = \frac{4}{9}$

► L' événement (X=1) est réalisé par les issues (B;V),(V;B),(V;N),(N;V)

Donc $P(X = 1) = \frac{4}{9} \times \frac{3}{9} + \frac{3}{9} \times \frac{4}{9} + \frac{3}{9} \times \frac{2}{9} + \frac{2}{9} \times \frac{3}{9} = \frac{36}{81} = \frac{4}{9}$

► L' événement (X=2) est réalisé par l' issue (V;V) donc $P(X=2) = \frac{3}{9} \times \frac{3}{9} = \frac{9}{81} = \frac{1}{9}$

<u>La loi de probabilité X est résumée par ce tableau :</u>

x_i	0	1	2
$P(X = x_i)$	$\frac{4}{9}$	$\frac{4}{9}$	$\frac{1}{9}$

II) Épreuve et schéma de Bernoulli

a) épreuve de Bernoulli :

définition : Une épreuve de Bernoulli est une expérience aléatoire qui ne comprend que deux issues contraires.

▶ l'une est appelée «succès», on la note S et sa probabilité d'apparition est p,

▶ l'autre est appelée «échec», on la note $\overline{S}$ et sa probabilité d'apparition est $1 - p$,

p est le appelé paramètre de l'épreuve de Bernoulli !

Ex :

▶ On lance une pièce de monnaie. L'issue Pile sera par exemple appelée succès et l'issue Face échec. Le paramètre de l'épreuve de Bernoulli est $p = \frac{1}{2}$

▶ On lance un dé. On décide d'appeler succès la sortie du 1 et échec la sortie de 2,3,4,5 ou 6. Le paramètre de l'épreuve de Bernoulli est $p = \frac{1}{6}$

b) loi de Bernoulli :

définition : Dans une épreuve de Bernoulli de paramètre p, on définit la variable aléatoire X qui prend la valeur 1 quand l'issue est S et 0 en cas d'échec.

La loi de probabilité X est définie par le tableau suivant :

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

Son espérance est : $E(X) = 0(1 - p) + 1 \text{x}\, p = p$

Sa variance est : $V(X) = (1 - p)\text{x}0^2 + p \text{ x } 1^2 - E(X)^2 = p - p^2 = p(1 - p)$

On dit que X suit la loi de Bernoulli de paramètre p

Ex : Reprenons les deux exemples du paragraphe précédent

▶ Dans le premier cas, la probabilité du succès "Pile" est $\frac{1}{2}$

La loi de Bernoulli de paramètre $\frac{1}{2}$ est

x_i	0	1
$P(X = x_i)$	$\frac{1}{2}$	$\frac{1}{2}$

▶ Dans le deuxième cas, la probabilité du succès "obtenir 1" est $\frac{1}{6}$

La loi de Bernoulli de paramètre $\frac{1}{6}$ est

x_i	0	1
$P(X = x_i)$	$\frac{5}{6}$	$\frac{1}{6}$

c) schéma de Bernoulli :

définition : Un schéma de Bernoulli est une répétition d'épreuves de Bernoulli identiques et indépendantes

Ex : Reprenons un des exemples précédents.
On lance successivement 3 fois un dé non truqué.

On décide d'appeler succès «obtenir 1». Sa probabilité d'apparition est $p = \frac{1}{6}$

On est en présence d'un schéma de Bernoulli avec la répétition trois fois, de façon indépendante, d'une même épreuve de Bernoulli de paramètre $\frac{1}{6}$

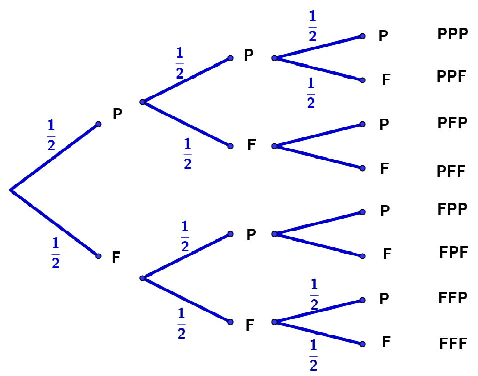
III) Épreuve et schéma de Bernoulli

a) loi binomiale :

définition : Considérons un schéma de Bernoulli, répétition de n épreuves de Bernoulli identiques et indépendantes de paramètre p. Appelons X la variable aléatoire qui associe à cette répétition de n épreuves le nombre de succès obtenus.
On dit que la loi de probabilité de X suit la loi binomiale de paramètres n et p.

Ex : Reprenons un des exemples précédents.
On lance successivement trois fois une pièce de monnaie. On décide d' appeler succès l'issue Pile.
Soit X la loi de probabilité correspondant au nombre de Pile obtenus.

X suit la loi binomiale de paramètres 3 et $\frac{1}{2}$



x_i	0	1	2	3
$P(X = x_i)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

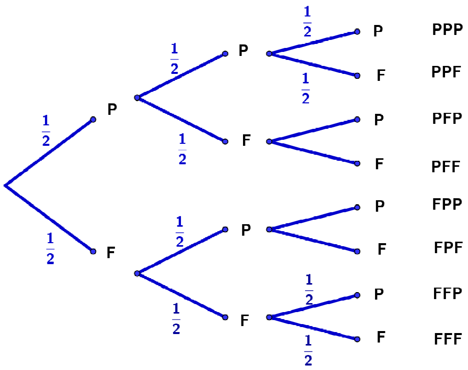
b) coefficients binomiaux :

définition : Considérons **un schéma de Bernoulli**, répétition de n épreuves de Bernoulli identiques et indépendantes représenté par un arbre. Soit un entier k compris entre 0 et n. Le nombre de chemins réalisant k succès est noté $\binom{n}{k}$.

Les nombres $\binom{n}{k}$ sont appelés coefficients binomiaux

k est un entier !

Ex : Reprenons l' exemple du schéma de Bernoulli précédent qui consistait à lancer successivement 3 fois une pièce de monnaie. L'issue Pile était le succès.



le nombre de chemins permettant de réaliser 2 succès est 3

$\binom{3}{2} = 3$

$\binom{3}{1} = 3$; $\binom{3}{0} = 1$

cas particuliers :
Soit un entier naturel n,
$\binom{n}{0} = 1$; $\binom{n}{1} = n$; $\binom{n}{n} = 1$

propriété : X est une variable aléatoire suivant la loi binomiale de paramètres n et p

Pour tout entier k compris entre 0 et n ($0 \leqslant k \leqslant n$), $P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$

Si un chemin réalise k succès, il réalise forcément n - k échecs.
Ce chemin conduira donc à une issue dont la probabilité est $p^k(1-p)^{n-k}$. Or, il y a $\binom{n}{k}$ chemins qui réalisent k succès. Par suite, $P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$

Ex : Prenons l' exemple d' un schéma de Bernoulli qui consiste à lancer successivement 5 fois un dé non truqué. On décide d'appeler succès «obtenir 1».
On note X la variable aléatoire indiquant le nombre de succès.
Calculons p(X = 3)

La variable aléatoire X suit une loi binomiale de paramètres n = 5 et $p = \frac{1}{6}$ donc

$$p(X = 3) = \binom{5}{3}\left(\frac{1}{6}\right)^3\left(1-\frac{1}{6}\right)^{5-3} = 10 \times \frac{1}{216} \times \left(\frac{5}{6}\right)^2 \simeq 0{,}03$$

pour calculer $\binom{5}{3}$, je peux utiliser la calculatrice ou le triangle de Pascal (voir plus bas) !!

b) espérance et variance de la loi binomiale :

propriétés (admise) :

X est une variable aléatoire suivant la loi binomiale de paramètres n et p

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

c) propriétés des coefficients binomiaux :

► Pour tous entiers n et k, si $0 \leqslant k \leqslant n$ alors $\binom{n}{k} = \binom{n}{n-k}$

► **démonstration**

Supposons $0 \leqslant k \leqslant n$

Sur l'arbre, il y a autant de chemins qui réalisent k succès que de chemins qui réalisent k échecs, or, réaliser k échecs équivaut à réaliser n – k succès

donc $\binom{n}{k} = \binom{n}{n-k}$

► Pour tous entiers n et k, si $0 \leqslant k \leqslant n-1$ alors $\binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1}$ ← formule de Pascal

► **démonstration**

Supposons $0 \leqslant k \leqslant n-1$

Considérons l'arbre à n+ 1 répétitions d'une épreuve de Bernoulli.

Le nombre de chemins réalisant k + 1 succès est $\binom{n+1}{k+1}$

Parmi ces chemins, il y en a deux sortes :

- ceux qui commencent par un succès, ils indiquent k succès lors des n premières répétitions et leur nombre est donc $\binom{n}{k}$
- ceux qui commencent par un échec, ils indiquent k + 1 succès lors des n premières répétitions et leur nombre est donc $\binom{n}{k+1}$

donc $\binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1}$

une application

Le triangle de Pascal permet de calculer de proche en proche les coefficients binomiaux en utilisant les propriétés précédentes. On admet que $\binom{0}{0} = 1$

n \ k	0	1	2	3	4	5	6
0	1						
1	1	1					
2	1	2	1				
3	1	3	3	1			
4	1	4	6	4	1		
5	1	5	10 +	10	5	1	
6	1	6	15	= 20	15	6	1

$\binom{3}{2}$

$$\binom{5}{3} = \binom{4+1}{2+1} = \binom{4}{2} + \binom{4}{3} = 6 + 4 = 10$$

IV) Échantillonnage - intervalle de fluctuation

a) échantillonnage :

définition : On obtient un échantillon de taille n en prélevant au hasard successivement et avec remise n individus d'une population

Après avoir prélevé un échantillon, on s'intéressera à la **valeur d'un caractère** des individus de la population. Dans le cours, les caractères étudiés n'auront que **deux valeurs possibles**.

intérêt de l'échantillonnage :
Il permet de vérifier une information donnée sur une population d'effectif important à partir de l'étude d'un échantillon de quelques dizaines ou centaines d'unités.

On rencontre souvent cette situation dans l'industrie quand on désire vérifier la qualité d'une production par exemple !

Soit une population d'individus dans laquelle une proportion (une fréquence) d' individus présente un certain caractère.
La démarche consiste à observer sur un échantillon de taille n issu de cette population la proportion f d'individus présentant le caractère donné.
On considère que l'échantillon prélevé est représentatif de la population.

L'échantillon n'est évidemment pas exactement représentatif de la population où il a été prélevée ! Il est donc nécessaire d'évaluer la marge d'erreur (fluctuation) afin de pouvoir utiliser l'information obtenue en connaissance de cause.

Ex :
On prélève au hasard (de manière identique et indépendante) des pièces dans une production et on observe si la pièce présente un défaut ou non.
► La population est l'ensemble des pièces de la production.
► Le caractère étudié est "avoir un défaut". Ses valeurs sont Vrai, Faux.

b) intervalle de fluctuation :

propriété (admise): Soit une population dont une proportion p des individus admet un caractère donné.
Dans un échantillon de taille n prélevé dans cette population, l'effectif des individus présentant ce caractère est une variable aléatoire suivant la loi binomiale de paramètres n et p.

<u>définition</u> : L'intervalle de fluctuation à 95% d'une fréquence, sur un échantillon de taille n prélevé au hasard, selon la loi binomiale de paramètres n et p est $\left[\frac{a}{n} , \frac{b}{n}\right]$

▶ a est le plus petit entier tel que P(X⩽ a) > 0,025

▶ b est le plus petit entier tel que P(X⩽ b) ⩾ 0,975

L'intervalle au seuil de 95% est le plus utilisé (on peut cependant rencontrer des seuils différents comme 90% ou 99%)

c) <u>prise de décision</u>

Supposons que la fréquence d'un caractère dans une population est p.
Pour vérifier cette hypothèse, on prélève un échantillon de taille n.
On observe la fréquence f du caractère dans l'échantillon.

Il s'agit maintenant de prendre la décision d'accepter ou de refuser l'information donnée par l'échantillon.
Peut-on se fier à la fréquence f trouvée et en déduire que l'hypothèse est vraie ?

<u>Règle de prise de décision :</u>

▶ Si f appartient à $\left[\frac{a}{n} , \frac{b}{n}\right]$, l' hypothèse est acceptée au seuil de 95%

▶ Si f n'appartient pas à $\left[\frac{a}{n} , \frac{b}{n}\right]$, l'hypothèse est rejetée au seuil de 5%

La probabilité de rejeter l' hypothèse alors qu'elle est vraie est **inférieure** à 5% !

<u>Ex :</u> Dans une société de production d'assiettes, 75% des assiettes produites n'ont aucun défaut et seront vendues dans les boutiques de la société. Les autres, ayant des imperfections, seront distribuées dans des magasins discount.
Déterminons pour un échantillon de taille 150 un intervalle de fluctuation au seuil de 95% du nombre d'assiettes sans aucun défaut.

La variable aléatoire X donnant pour chaque échantillon de 150 assiettes le nombre d'assiettes sans défaut suit une loi binomiale de paramètres n=150 et p = 0,75

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

<u>A l'aide d'un tableur ou d'une calculatrice, on détermine :</u>

- le plus petit entier a tel que P(X⩽ a) > 0,025
- le plus petit entier b tel que P(X⩽ b) ⩾ 0,975

=LOI.BINOMIALE(A2;150;0,75;1)

=LOI.BINOMIALE(A2;150;0,75;0)

=A2/150

	A	B	C	D
1	k	p(X=k)	p(X<=k)	k/n
2	100	0,0050926	0,0136186	0,6666667
3	101	0,0075633	0,0211819	0,6733333
4	102	0,0109001	0,032082	0,68
5	103	0,0152389	0,0473209	0,6866667
6	104	0,0206604	0,0679813	0,6933333
7	105	0,0271537	0,0951351	0,7
8	106	0,0345826	0,1297176	0,7066667
9	107	0,0426626	0,1723803	0,7133333
10	108	0,0509581	0,2233384	0,72
11	109	0,0589057	0,2822441	0,7266667
12	110	0,0658673	0,3481115	0,7333333
13	111	0,0712079	0,4193194	0,74
14	112	0,0743868	0,4937062	0,7466667
15	113	0,0750451	0,5687514	0,7533333
16	114	0,0730703	0,6418216	0,76
17	115	0,0686225	0,7104441	0,7666667
18	116	0,0621152	0,7725593	0,7733333
19	117	0,0541517	0,826711	0,78
20	118	0,0454324	0,8721434	0,7866667
21	119	0,0366513	0,9087947	0,7933333
22	120	0,0284048	0,9371995	0,8
23	121	0,0211275	0,958327	0,8066667
24	122	0,0150663	0,9733934	0,8133333
25	123	0,0102892	0,9836826	0,82
26	124	0,0067212	0,9904038	0,8266667

On obtient P(X⩽101) = 0,02 et P(X⩽102) ≃ 0,03 donc a = 102

On a aussi P(X⩽122) = 0,973 et P(X⩽123)≃ 0,983 donc b = 123

Donc, $\frac{a}{n} = 0,68$ et $\frac{b}{n} = 0,82$

L'intervalle de fluctuation à 95% de la fréquence des assiettes "sans aucun défaut" est [0,68 ; 0,82] et sur un échantillon de 150 assiettes, l'intervalle de fluctuation à 95% du nombre d'assiettes est [102,123]

Si je trouve dans échantillon de 150 assiettes un nombre d'assiettes sans défaut compris entre 102 et 123, je peux considérer que la société de production produit pratiquement 75% d' assiettes sans défaut !