Loi binomaile. Applications.

Clément BOULONNE

Session 2019

Préambule

Niveau de la leçon

Terminale S

Prérequis

Probabilités

Références

— M. LEZEN, *Leçon n° 3 : Coefficients binomiaux, dénombrement des combinaisons, formule du binôme. Applications*. URL : http://www.capes-de-maths.com/index.php?page=leconsNEW. 2011.
— G. CONNAN, *Une année de mathématiques en Terminale S*. Ch.14. URL : :http://tehessin.tuxfamily.org. 2009-2010.
— G. COSTANTINI, *Loi binomiale*. URL : http://bacmaths.net.
— C. SUQUET, *Intégration et Probabilité Élémentaires*. URL : http://math.univ-lille1.fr/~ipeis/. 2009-2010.
— C. GRAFFIGNE, *Démonstration de la formule du binôme de Newton*. Université Paris V, L1, S1. http://www.math-info.univ-paris5.fr/~avner/MC1/L1_S1/cours/eq/node11.html

Table des matières

1 Loi de Bernoulli

Définition 1.1 (Loi de Bernoulli). Soit $\mathcal{E}$ une épreuve comportant deux issues (succès et échec). On note p la probabilité de succès. Soit X la variable aléatoire qui est égale à 1 en cas de succès et 0 sinon. Alors, on dit que X suit une loi de Bernoulli de paramètres p. On note alors $X \sim \mathrm{Bern}(p)$.

Remarque 1.2. Si $X \sim \mathrm{Bern}(p)$, on notera :

$$P(X = 1) = p \quad \text{et} \quad P(X = 0) = 1 - p = q.$$

Exemple 1.3. On lance un dé non pipé. On note X la variable aléatoire qui prend comme valeur 1 si la face 6 apparaît lors du lancer et 0 sinon.

La variable aléatoire X est une variable aléatoire qui suit la loi de Bernoulli de paramètres $1/6$. Donc $X \sim \mathrm{Bern}(1/6)$.

Lemme 1.4. Si $X \sim \mathrm{Bern}(p)$ alors $X^2 \sim \mathrm{Bern}(p)$.

Démonstration. ◊ On a $X^2(\Omega) = \{0, 1\}$ et :

$$P(X^2 = 1) = P(X = 1) = p$$

donc $X^2 \sim \mathrm{Bern}(p)$. □

PROPOSITION 1.5. Si $X \sim \mathrm{Bern}(p)$ alors :

1. $\mathbf{E}(X) = p$
2. $\mathrm{Var}(X) = pq$.

Démonstration. ◊ On a :

$$\mathbf{E}(X) = P(X = 0) \times 0 + P(X = 1) \times 1 = q \times 0 + p \times 1 = p,$$

et :

$$\mathrm{Var}(X) = \mathbf{E}(X^2) - \mathbf{E}(X)^2 = \mathbf{E}(X^2) - p^2$$

or $X^2 \sim \mathrm{Bern}(p)$, donc on a : $\mathbf{E}(X^2) = \mathbf{E}(X) = p$.

Ainsi, $\mathrm{Var}(X) = p - p^2 = pq$. □

2 Loi binomiale

Définition 2.1 (Loi binomiale). Soit Ω l'univers associé à une expérience aléatoire. Soit X une variable aléatoire définie sur Ω. On dit que X suit une loi binomiale de paramètres $n \in \mathbb{N}^*$ et $p \in [0\,,1]$ lorsque :

1. $X(\Omega) = \{0, 1, \dots, n\}$;
2. pour tout $k \in \{0, 1, \dots, n\}$, $P(X = k) = \binom{n}{k} p^k (1-p)^{n-k} = \binom{n}{k} p^k q^{n-k}$.

Si X suit une loi binomiale de paramètres n et p alors on note $X \sim \mathrm{Bin}(n, p)$.

Remarque 2.2. Soit $X \sim \mathrm{Bin}(n, p)$. On a bien défini une variable aléatoire car :

$$\sum_{k=0}^{n} P(X = k) = \sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} = [p + (1-p)]^n = 1.$$

THÉORÈME 2.3. Soit $\mathcal{E}$ une épreuve comportant deux issues (succès et échec). On note p la probabilité de succès. On note n fois, de façons indépendantes, l'épreuve $\mathcal{E}$. Soit X la variable aléatoire correspondant au nombre de succès. Alors : X suit une loi binomiale de paramètres n et p.

Démonstration. $\Diamond$ La probabilité d'avoir k succès suivis de $n - k$ succès suivis de $n - k$ échecs est : $p^k(1-p)^{n-k}$. Mais les succès et les échecs n'apparaissent pas nécessairement dans cet ordre.

On considère l'ensemble des « mots » de n lettres qui ne contiennent que des S (Succès) et des E (Échecs). On sait qu'il y en a exactement $\binom{n}{p}$ qui contiennent exactement k fois la lettre S (et donc $n - k$ fois la lettre E).

On en déduit m

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

et ceci pour tout $k \in \{0, 1, \dots, n\}$. $\square$

Remarques 2.4.

1. La probabilité d'avoir n succès : $P(X = n) = p^n$ et d'avoir aucun succès $P(X = 0) = q^n$. Par conséquent, la probabilité d'avoir au moins un succès est :

$$P(X \geq 1) = 1 - P(X = 0) = 1 - q^n.$$

2. La loi de Bernoulli est un cas particulier de la loi binomiale où l'épreuve $\mathcal{E}$ n'est réalisée qu'une seule fois.
3. Toute variable aléatoire X suivant une loi binomiale de paramètres $n \in \mathbb{N}^*$ et $p \in [0\,,1]$ peut s'écrire comme somme $X = X_1 + \cdots + X_n$ où, pour tout $k \in \{0, 1, \dots, n\}$, X_k est une variable aléatoire suivant une loi de Bernoulli de paramètre p (X_k vaut 1 en cas de succès à la k^{e} réalisation de $\mathcal{E}$ et 0 sinon).

Exemples 2.5. La probabilité qu'un tireur atteigne sa cible est $p = \frac{3}{4}$. On suppose qu'il fait deux tirs et on note X la variable aléatoire associant à cette épreuve le nombre de succès obtenus ($X = 0$, 1 ou 2).

1. Calculer la probabilité des événements $\{X = 0\}$, $\{X = 1\}$ et $\{X = 2\}$.
2. Calculer $\sum_{k=0}^{2} P(X = k)$.
3. On suppose qu'il fait sept tirs et on note Y la variable aléaoire associant à cette épreuve le nombre de succès obtenus. Calculer $P(X = 1)$ et $P(X = 2)$.

THÉORÈME 2.6 (ESPÉRANCE ET VARIANCE D'UNE LOI BINOMIALE). Si $X \sim \mathrm{Bin}(n, p)$ avec $n \in \mathbb{N}^*$ et $p \in [0\,,1]$ alors :

$$\mathbf{E}(X) = np \quad \text{et} \quad \mathrm{Var}(X) = npq.$$

Démonstration. $\Diamond$ Puisque $X \sim \mathrm{Bin}(n, p)$, il existe des variables aléatoires (réelles) $X_1, X_2, \dots, X_n$ définies sur Ω indépendantes, de loi de Bernoulli de même paramètre p telles que $X = \sum_{i=1}^{n} X_i$.

Par linéarité de l'espérance :

$$\mathbf{E}(X) = \mathbf{E}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathbf{E}(X_i)$$

et d'après ce qui précède :

$$\mathbf{E}(X) = \sum_{i=1}^{n} p = np.$$

De même pour la variance :

$$\mathrm{Var}(X) = \mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathrm{Var}(X_i) = \sum_{i=1}^{n} pq = npq.$$

$\square$

Exemple 2.7. La probabilité qu'un tireur atteigne sa cible est $p = \frac{3}{4}$. On suppose qu'il tire $n = 7$ fois. On note X la variable aléatoire associant à cette expérience aléatoire le nombre de succès obtenus. Calculer son espérance et sa variance.

3 Propriétés sur les coefficients binomiaux

3.1 Définitions et propriétés

Définition 3.1 (Combinaisons). Soient n et p deux entiers naturels et E un ensemble contenant n éléments. Un sous-ensemble de E contenant p éléments est appelé une *combinaison* de p éléments de E.

Le nombre de p-combinaisons d'un ensemble contenant n éléments est noté $\binom{n}{p}$ ou $\binom{n}{p}$.

Exemple 3.2. Pour gagner au Loto, il faut trouver 3 numéros parmi 5. On veut savoir combien il y a de grilles possibles. Considérons une grille quelconque (c'est-à-dire une 3-combinaison de l'ensemble des 5 numéros) : par exemple $\{1, 3, 4\}$. Il y a 3! façons possibles d'ordonner ces nombres. Or, il y a $\binom{5}{3} \times 3!$ suites de 3 nombres ordonnées. Mais, on compte $5 \times 4 \times 3$ de ces dernières suites. Donc :

$$\binom{5}{3} = \frac{5 \times 4 \times 3}{3!}.$$

On peut maintenant généraliser la formule :

PROPOSITION 3.3. Le nombre de p-combinaisons d'un ensemble contenant n éléments est noté

$$\binom{n}{p} = \frac{n(n-1)(n-2)\cdot(n-(p-1))}{p!} \tag{1}$$

$$= \frac{n!}{p!(n-p)!} \tag{2}$$

$\Diamond$ *Démonstration de la proposition 3.3.* On part de la formule (1) pour arriver à la formule (2) :

$$\begin{aligned}\binom{n}{p} &= \frac{n(n-1)(n-2)\cdots(n-p+1)}{p!} \\ &= \frac{n(n-1)(n-2)\cdots(n-p+1)}{p!}\frac{(n-p)(n-p-1)\cdots 2\times 1}{(n-p)(n-p-1)\cdots 2\times 1} \\ &= \frac{n!}{p!(n-p)!}\end{aligned}$$

Une autre façon de voir la formule (2). Il y a A_n^p manières de tirer p objets parmi n en les ordonnant soit

$$\mathrm{A}_n^p = \frac{n!}{(n-p)!}.$$

Une fois les p objets tirés, il y a $p!$ manières de les ordonner. Il y a donc $\frac{\mathrm{A}_n^p}{p!}$ manières de tirer p objets parmi sans les ordonner. D'où

$$\binom{n}{p} = \frac{\mathrm{A}_n^p}{p!} = \frac{1}{p!}\frac{n!}{(n-p)!}.$$

□

Définition 3.4 (Coefficients binomiaux). Soit p un entier naturel non nul. Les nombres $\binom{n}{p}$ sont appelés les *coefficients binomiaux*.

PROPOSITION 3.5 (FORMULE DE PASCAL). Soit $n, p \in \mathbb{N}$ tel que $p < n$. On a :

$$\binom{n}{p} = \binom{n-1}{p} + \binom{n-1}{p-1}.$$

$\Diamond$ *Démonstration de la formule de Pascal.* Soit un ensemble E à n éléments. On suppose que l'on a « extrait » une partie à p éléments. Si l'on retire un élément $\{a\}$ à E, c'est soit un élément de la combinaison, soit non. Dans le premier cas, les $p-1$ restants forment une partie de l'ensemble $E\setminus\{a\}$ de cardinal $n-1$, et dans le second, ce sont les p éléments qui forment une partie de $E \setminus \{a\}$. Cette union étant disjointe, les cardinaux s'ajoutent pour aboutir à l'égalité demandée. □

$n\backslash p$	0	1	2	3	$\cdots$
0	1				
1	1	1			
2	1	2	1		
3	1	3	3	1	
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$

FIGURE 1 – Triangle de Pascal

PROPOSITION 3.6 (FORMULE ITÉRÉE DE PASCAL). Soit $p \leq n$ deux entiers naturels. Alors

$$\sum_{k=p}^{n}\binom{k}{p} = \binom{n+1}{p+1}.$$

$\Diamond$ *Démonstration de la formule itérée de Pascal.* On effectue une récurrence sur l'entier n.

Initialisation Lorsque $n = p$, les deux membres valent 1.

Hérédité On suppose que la formule est vraie au rang n et on montre qu'elle est encore vraie au rang $n+1$:

$$\sum_{k=p}^{n+1} \binom{k}{p} = \sum_{k=p}^{n} \binom{k}{p} + \binom{n+1}{p}$$

et d'après l'hypothèse de récurrence,

$$\sum_{k=p}^{n+1} \binom{k}{p} = \binom{p+1}{n+1} + \binom{n+1}{p} = \binom{n+2}{p+1}.$$

La dernière égalité est justifiée par l'emploi de la formule de Pascal.

□

On note $A = \mathbb{C}$ (ou $\mathbb{R}$ ou $\mathbb{Q}$ ou $\mathbb{Z}$).

THÉORÈME 3.7 (FORMULE DU BINÔME). Soient deux éléments a, b de A qui commutent. Alors :

$$\forall n \in \mathbb{N}, \quad (a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}.$$

◊ *Démonstration de la formule du binôme de Newton.* Pour $n = 1$, nous avons :

$$\sum_{k=0}^{1} \binom{1}{k} a^k b^{1-k} = \binom{1}{0} b + \binom{a}{=} a + b.$$

La formule du binôme est vraie pour $n = 1$.

Supposons que la formule du binôme soit vraie au rang $n \geq 1$. Alors,

$$(a+b)^{n+1} = (a+b) \cdot (a+b)^n = (a+b) \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}.$$

En distribuant le produit, nous obtenons

$$(a+b)^{n+1} = \sum_{k=0}^{n} \binom{n}{k} a^{k+1} b^{n-k} + \sum_{k=0}^{n} \binom{n}{k} a^k b^{n+1-k}.$$

Nous effectuons alors la translation d'indices $l = k+1$ dans la première somme :

$$(a+b)^{n+1} = \sum_{l=1}^{n+1} \binom{n}{l-1} a^l b^{n+1-l} + \sum_{k=0}^{n} \binom{n}{k} a^k b^{n+1-k}.$$

L'indice de sommation étant muet, nous pouvons regrouper les deux sommes :

$$(a+b)^{n+1} = \binom{n}{n} a^{n+1} + \sum_{k=1}^{n} \left[\binom{n}{k-1} + \binom{n}{k} \right] a^k b^{n+1-k} + \binom{n}{0} b^{n+1}.$$

On utilise ensuite la formule du triangle de Pascal :

$$(a+b)^{n+1} = \binom{n}{n} a^{n+1} + \sum_{k=1}^{n} \binom{n+1}{k} a^k b^{n+1-k} + \binom{n}{0} b^{n+1}.$$

On remarque que : $\binom{n}{0} = 1 = \binom{n+1}{0}$ et que $\binom{n}{n} = 1 = \binom{n+1}{n+1}$ pour faire entrer les deux termes isolés dans la somme.

$$(a+b)^{n+1} = \sum_{k=0}^{n+1} \binom{n+1}{k} a^k b^{n+1-k}.$$

□

Corollaire 3.8. On a les égalités suivantes :

1. $\sum_{k=0}^{n} \binom{n}{k} = 2^n$,
2. $\sum_{k=0}^{n} (-1)^k \binom{n}{k} = 0$.

◊ *Démonstration du corollaire 3.8.*

1. On utilise le binôme de Newton avec $a = 1$ et $b = 1$.
2. On utilise le binôme de Newton avec $a = -1$ et $b = 1$.

□

Remarque 3.9. On remarque que l'égalité 1 du corollaire 3.8 traduit le fait que le nombre de parties d'un ensemble à n éléments est 2^n. En effet, ce nombre est la somme des nombres de parties ayant respectivement 0, 1, ... éléments (le cardinal d'une union disjointe est la somme des cardinaux), ce qui correspond bien à la somme indiquée.

PROPOSITION 3.10 (FORMULE DE VAN DER MONDE). Pour tous entiers m, n et p tels que $p \leq m + n$, on a l'égalité :

$$\binom{m+n}{p} = \sum_{k=0}^{p} \binom{m}{k} \binom{n}{p-k}.$$

Remarque 3.11. On remarque que l'égalité 1 du corollaire 3.8 traduit le fait que le nombre de parties d'un ensemble à n éléments est 2^n. En effet, ce nombre est la somme des nombres de parties ayant respectivement 0, 1, ... éléments (le cardinal d'une union disjointe est la somme des cardinaux), ce qui correspond bien à la somme indiquée.

◊ *Démonstration de la formule de Van der Monde.* Soit x un réel. Alors :

$$(1+x)^m (1+x)^n = (1+x)^{m+n} = \sum_{p=0}^{m+n} \binom{m+n}{p} x^p.$$

Or

$$\begin{aligned}
(1+x)^m (1+x)^n &= \left(\sum_{i=0}^{m} \binom{m}{i} x^i \right) \left(\sum_{j=0}^{n} \binom{n}{j} x^j \right) = \sum_{i=0}^{m} \sum_{j=0}^{n} \binom{m}{i} \binom{n}{j} x^{i+j} \\
&= \left(\binom{m}{0} \binom{n}{0} \right) + \left((\binom{m}{0} \binom{n}{1} + \binom{m}{1} \binom{n}{0}) x \right) \\
&\qquad + \left((\binom{m}{0} \binom{n}{2} + \binom{m}{1} \binom{n}{1} + \binom{m}{2} \binom{n}{0}) x^2 \right) + \cdots \\
&= \sum_{p=0}^{m+n} \Big((\sum_{\substack{i,j>0 \\ i+j=p}} \binom{m}{i} \binom{n}{j}) x^p \Big).
\end{aligned}$$

Par identification des coefficients de ce polynôme de degré p, on obtient finalement que, pour tout entier $0 \leq p \leq m + n$,

$$\binom{m+n}{p} = \sum_{\substack{i,j>0 \\ i+j=p}} \binom{m}{i} \binom{n}{j} = \sum_{i=0}^{p} \binom{m}{i} \binom{n}{p-i}.$$

□

4 Stabilité additive de la loi binomiale

Théorème 4.1 (Stabilité additive de la loi binomiale). Si $X \sim \mathrm{Bin}(m,p)$ et $Y \sim \mathrm{Bin}(n,p)$ avec X et Y indépendantes, alors $X+Y = \mathrm{Bin}(m+n,p)$.

Soit $(A_i)_{1\leq i\leq n}$ une suite d'événements. On note : $\coprod_{i=0}^{n} A_i$ si les événements sont disjoints.

Démonstration. $\Diamond$ On pose $S = X+Y$. On a clairement $S(\Omega) = \{0,\ldots,m+n\}$.
Calculons $P(S^{-1}(k))$ pour tout $1 \leq k \leq m+n$:

$$S^{-1}(k) = \coprod_{i=0}^{k} X^{-1}(i) \cap Y^{-1}(k-i).$$

D'où :

$$P(S^{-1}(k)) = \sum_{i=0}^{k} P(X^{-1}(i) \cap Y^{-1}(k-i)).$$

Et comme X et Y sont indépendantes :

$$P(S^{-1}(k)) = \sum_{i=0}^{k} P(X^{-1}(i))P(Y^{-1}(k-i)).$$

Comme $X \sim \mathrm{Bin}(m,p)$ et $Y \sim \mathrm{Bin}(n,p)$:

$$\begin{aligned} P(S^{-1}(k)) &= \sum_{i=0}^{k} \binom{m}{i} p^i (1-p)^{m-i} \binom{n}{k-i} p^{k-i}(1-p)^{n-(k-i)} \\ &= \left(\sum_{i=0}^{k} \binom{m}{i}\binom{n}{k-i}\right) p^k (1-p)^{m+n-k}. \end{aligned}$$

Et comme $\sum_{i=0}^{k} \binom{m}{i}\binom{n}{k-i} = \binom{m+n}{k}$.

$$P(S^{-1}(k)) = \binom{m+n}{k} p^k (1-p)^{m+n-k}.$$

Donc $S \sim \mathrm{Bin}(m+n,p)$. □

5 Applications et compléments

5.1 Loi binomiale et suites

1. On lance n fois un dé équilibré. Déterminer la probabilité p_n d'obtenir au moins un 6.
2. Déterminer le nombre minimal de lancers pour qu'on ait $p_n \geq 0,99$.

$\Diamond$ *Solutions.* 1. Soit X_n le nombre de fois que l'on obtienne un 6 lors de n lancers d'un dé équilibré ($n \in \mathbb{N}^*$). Les expériences sont répétés identiquement et sont indépendants les unes aux autres. Ainsi, X_n suit la loi binomiale de paramètres n et $p = \frac{1}{6}$ (proba d'obtenir un 6). On veut calculer la probabilité de l'événement « $X_n \geq 1$ ». Pour cela, on utilise la probabilité de l'événement contraire « $X_n < 1$ ».

$$p_n = P(X_n \geq 1) = 1 - P(X_n < 1).$$

Or, les valeurs possibles de X_n appartiennent à l'ensemble $\{0, 1, \ldots, n\}$. Ainsi, $X_n < 1$ équivaut à l'événement $X_n = 0$ d'où :

$$\begin{aligned} p_n &= P(X_n \geq 1) = 1 - P(X_n < 1) = 1 - P(X_n = 0) \\ &= 1 - \left(\binom{n}{0} \times \left(\frac{1}{6}\right)^0 \times \left(\frac{5}{6}\right)^n \right) = 1 - 1 \times 1 \times \left(\frac{5}{6}\right)^n = 1 - \left(\frac{5}{6}\right)^n. \end{aligned}$$

2. On veut déterminer le nombre minimal de lancers pour qu'on ait $p_n \geq 0,99$. Pour cela, on résout l'inéquation $p_n \geq 0,99$.

$$\begin{aligned} p_n \geq 0,99 &\Leftrightarrow 1 - \left(\frac{5}{6}\right)^n \geq 0,99 \Leftrightarrow -\left(\frac{5}{6}\right)^n \geq -0,01 \\ &\Leftrightarrow \left(\frac{5}{6}\right)^n \leq 0,01 \Leftrightarrow \ln\left(\left(\frac{5}{6}\right)^n\right) \leq \ln(0,01) \Leftrightarrow n \ln\left(\frac{5}{6}\right) \leq \ln(0,01). \end{aligned}$$

Ici, on doit changer le sens du signe de l'inéquation car $\frac{5}{6} < 1$ et $\ln(\frac{5}{6}) < 0$.

$$p_n \geq 0,99 \Leftrightarrow n \ln\left(\frac{5}{6}\right) \leq \ln(0,01) \Leftrightarrow n \geq \frac{\ln(0,01)}{\ln(\frac{5}{6}} \Leftrightarrow n \geq 25,26.$$

Conclusion : À partir du 26e lancer de dés, la probabilité d'obtenir au moins un 6 est supérieure à $0,99$.

□

5.2 La planche de Galton

On dispose d'une planche longue en bois, d'un récipient de billes en plomb à son extrémité, des clous disposaient sur plusieurs niveaux et des receptacles.

Les clous sont disposés sur plusieurs niveaux de la manière suivante :
— 1 clou centré au premier niveau ;
— 2 clous centrés au deuxième niveau ;
— 3 clous centrés et alignés horizontalement au troisième niveau ;
— ...
— n clous centrés et alignés horizontalement au n^{e} niveau ;

Les clous sont espacés régulièrement pour que les billes puissent passser entre eux.

On lève la planche pour que la gravité puisse faire tomber les billes du récipient vers les clous puis vers les réceptacles.

Cette expérience mathématique est dûe à Sir Franck GALTON (1822-1911), mathématicien et scientifique britanique. La question que se posait Galton quand il construit cette planche était : « Dans quel receptacle il y aura le plus de billes ? ».

Au fur et à mesure de ces expériences, il constate que quand il fait tomber un très grand nombre de billes de plomb, la répartition des billes dans les bacs se font en cloche de Gauss.

On peut aussi calculer la probabilité que la bille touche un certain clou dans la planche. Pour cela, on aura besoin du triangle de Pascal que l'on a introduit plus haut dans la leçon.

Ainsi, la probabilité que la bille arrive dans le k^{e} receptacle (s'il y a n rangées de clous et n receptacles) est de :

$$P_n = \frac{\binom{n}{k}}{2^n}.$$

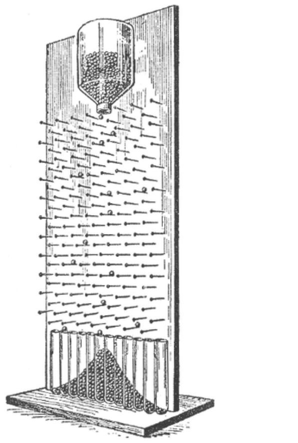

FIGURE 2 – Planche de Galton

$1/1$

$1/2 \quad 1/2$

$1/4 \quad 2/4 \quad 1/4$

$1/8 \quad 3/8 \quad 3/8 \quad 1/8$

$1/16 \quad 4/16 \quad 6/16 \quad 4/16 \quad 1/16$

$1/32 \quad 5/32 \quad 10/32 \quad 10/32 \quad 5/32 \quad 1/32$

$1/64 \quad 6/64 \quad 15/64 \quad 20/64 \quad 15/64 \quad 6/64 \quad 1/64$

$1/128 \quad 7/128 \quad 21/128 \quad 35/128 \quad 35/128 \quad 21/128 \quad 7/128 \quad 1/128$

FIGURE 3 – Probabilité de toucher un clou dans la planche

5.3 Convergence

5.3.1 Vers la loi de Poisson

THÉORÈME 5.1. Lorsque n tend vers l'infini et que simultanément $p_n \to 0$ de sorte que $\lim_n np_n = a > 0$, la loi binomiale de paramètres n et p_n converge vers la loi de Poisson de paramètre a. En pratique, on remplace la loi binomiale par une loi de Poisson dès que $n > 30$ et $np < 5$ ou dès que $n > 50$ et $p < 0.1$.

Démonstration. $\diamondsuit$ On décompose $P(X = k)$:

$$\binom{n}{k} p_n^k (1-p_n)^{n-k} = \frac{n(n-1)\cdots(n-k+1)}{k!} p_n^k (1-p_n)^{n-k}$$
$$= \frac{(np_n)^k}{k!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{k-1}{n}\right)(1-p_n)^{n-k}.$$

On se place dans la situation où p_n est équivalent à $\frac{a}{n}$ en l'infini.
- Lorsque n tend vers l'infini, les facteurs $\left(1-\frac{1}{n}\right)$, $\left(1-\frac{2}{n}\right)$, …, $\left(1-\frac{k-1}{n}\right)$ tendent vers 1. Le produit de ces termes tend également vers 1 puisqu'ils sont en nombre fini fixé k.
- On a :
$$(1-p_n)^{n-k} = (1-p_n)^n (1-p_n)^{-k},$$
or, $\lim_{p\to 0}(1-p)^{-k} = 1$ et de plus, $(1-p_n)^n \simeq (1-\frac{a}{n})^n$ et ce dernier terme tend vers e^{-a} quand n tend vers l'infini.

On trouve donc :
$$\lim_{n\to+\infty} \binom{n}{k} p_n^k (1-p_n)^{n-k} = \frac{a^k}{k!}\mathrm{e}^{-a},$$
qui est la probabilité de k pour la loi de Poisson de paramètre a. $\square$

5.3.2 Vers la loi normale

THÉORÈME 5.2. Soit $(X_n)_n$ une suite de variable aléatoires indépendnates de même loi de Bernoulli $\mathrm{Bern}(p)$ et $S_n = X_1 + \cdots + X_n$ suit la loi binomiale $\mathrm{Bin}(n, p)$.

D'après le théorème central limite, la loi de S_n peut re approximée par la loi normale $\mathfrak{N}(\mathbf{E}(S_n), \mathrm{Var}(S_n))$, c'est-à-dire par la loi $\mathfrak{N}(np, npq)$.

Remarque 5.3. En pratique, lorsque $n \geq 30$, $np \geq 15$ et $npq > 5$, la loi binomiale $\mathrm{Bin}(n, p)$ peut être approximée par la loi normale $\mathfrak{N}(np, npq)$.

5.4 Échantillonnage

5.4.1 Premier problème : proportion de boules dans une urne

Dans une urne contenant une dizaine de boules, il y a 2 boules noires et 8 boules blanches. La proportion de boules noires est donc de $1/5$.

On pioche dans l'urne avec ordre et remise une vingtaine de boules et on s'intéresse à la proportion de boules noires obtenues.

Cette expérience a été recommencée 100 fois à l'aide d'un tableur et voici les proportions obtenues.

Proportion	0	$0,05$	$0,1$	$0,15$	$0,2$	$0,25$	$0,3$	$0,35$	$0,4$	$0,45$	$0,5$	Total
Nb d'échantillons	0	9	13	20	27	16	9	5	0	1	0	100

1. Quel est le nombre d'échantillons qui ont une proportion de boules noires de $0,3$?

2. Quel est le nomb re d'échantillons qui ont une proportion de boules noires de $0,6$?
3. Quel est le nombre d'échantillons qui ont une proportion de boules noires entre $0,1$ et $0,4$?
4. Le but de cette partie est de retrouver par le calcul ce dernier nombre. On considère la variable aléatoire X qui lors de l'expérience compte le nombre boules noires obtenues.
   (a) Justifier que X suit une loi binomiale dont on précisera les paramètres.
   (b) Calculer $P(2 \leq X \leq 8)$.
   (c) En déduire la probabilité que la proportion de boules noires soit comprise entre 0 et $0,4$.

$\Diamond$ *Solution.*

Proportion	0	$0,05$	$0,1$	$0,15$	$0,2$	$0,25$	$0,3$	$0,35$	$0,4$	$0,45$	$0,5$	Total
Nb d'échantillons	0	9	13	20	27	16	9	5	0	1	0	100

1. Le nombre d'échantillons qui ont une proportion de boules noires de $0,3$ est 9.
2. Le nombre d'échantillons qui ont une proportion de boules noires de $0,6$ est 0. En effet, tous les échantillons sont déjà dans le tableau.
3. Le nombre d'échantillons qui ont une proportion de boules noires comprise entre $0,1$ et $0,4$ est $13 + 20 + 27 + 16 + 9 + 5 = 90$. Soit 90%.
4. (a) On recommence 20 fois de manière indépendante une expérience ayant deux issues possibles, succès ou échec. La variable aléatoire qui compte le nombre de succès suit une loi binomiale de paramètres 20 et $1/5$.
   (b) $P(2 \leq X \leq 8) = 0,92$.
   (c) On cherche la probabilité que la proportion de boules noires dans un échantillon soit comprise entre $0,1$ et $0,4$; c'est-à-dire la probabilité qu'il y ait entre 10% et 40% de boules noires. Or chaque échantillonnage contient 20 boules. Ainsi 10% de boules noires pari ces 20 boules représente exactement 2 boules noires. De même 40% représente 8 boules noires. Finalement, chercher la probabilité que la proportion de boules noires dans les échantillonnages soit comprise entre $0,1$ et $0,4$ revient à chercher la probabilité de piocher entre 2 et 8 boules noires parmi les 20 boules. C'est exactement la probabilité que l'on a calculé à la question 4b, soit $0,92$. Ce qui correspond à peu près au 90% trouvé grâce au tableau.

$\square$

5.4.2 Second problème : proportion de camions sur une autoroute

Sur une autoroute, la proportion des camions par rapport à l'ensemble des véhicules est $0,07$.

1. Soit X le nombre de camions parmi 100 véhicules choisis au hasard. Calculer $P(X \geq 5)$.
2. Soit Y le nombre de camions parmi 1000 véhicules choisis au hasard. Calculer $P(65 \leq Y \leq 75)$.
3. On choisit n véhicules au hasard. Pour quelles valeurs de n peut-on affirmer que la proportion de camions est entre $0,06$ et $0,08$ avec un risque d'erreur inférieur à 5% ?

$\Diamond$ *Solution.* 1. Soit X une variable aléatoire de loi binomiale $\mathrm{Bin}(100, 0.07)$. $100 \geq 30$, $100 \times 0,07 = 7 < 15$, $0,07 \leq 0,1$ donc l'approximation à utiliser est celle par la loi de Poisson $\mathrm{Pois}(7)$ et :

$$P(X \geq 5) \approx 1 - \mathrm{e}^{-7} \sum_{k=0}^{4} \frac{7^k}{k!} \approx 0,827.$$

2. Y suit la loi binomiale $\mathrm{Bin}(1000, 0.07)$. $1000 \geq 30$, $1000 \times 0,07 = 70 \geq 15$, $70 \times 0,93 = 64,1 > 4$ donc l'approximation à utiliser est celle par la loi normale $\mathfrak{N}(70, 65.1)$ et si F désigne la fonction de répartition de la loi $\mathfrak{N}(70, 65.1)$,

$$\begin{aligned} P(65 \leq Y \leq 75) &\approx F(75.5) - F(64.5) = \Phi\left(\frac{5.5}{\sqrt{65.1}}\right) - \Phi\left(-\frac{5.5}{\sqrt{65.1}}\right) \\ &= 2\Phi\left(\frac{5.5}{\sqrt{65.1}}\right) - 1 \approx 2\Phi(0.68) \approx 0.5 \end{aligned}$$

3. On choisit n véhicules au hasard. Le nombre S_n des camions parmi ces n véhicules suit la loi binomiale $\mathrm{Bin}(n, 0.07)$ et la proportion des camions est $\frac{S_n}{n}$.
   On cherche n tel que
   $$P\left(\left|\tfrac{S_n}{n} - 0.07\right| \geq 0.01\right) = 0.05.$$
   Si $n \geq 30$, $0.07n \geq 15$ et $0.07 \times 0.93 \times n > 5$, c'est-à-dire $n \geq 215$, on peut approximer la loi de $\frac{S_n}{n}$ par la loi normale $\mathfrak{N}(0.07, \frac{0.0651}{n})$ et la loi de $\frac{S_n}{n} - 0.07$ par la loi normale $\mathfrak{N}(0, \frac{0.065}{n})$.
   On a alors :
   $$\begin{aligned} P\left(\left|\frac{S_n}{n} - 0.07\right| \geq 0.01\right) &= P\left(\left|\frac{\sqrt{n}}{\sqrt{0.0651}}\left(\frac{S_n}{n} - 0.07\right)\right| \geq \frac{\sqrt{n}}{\sqrt{0.0651}}\frac{1}{100}\right) \\ &\approx 2\left(1 - \Phi\left(\frac{\sqrt{n}}{\sqrt{651}}\right)\right) \approx 0.05 \end{aligned}$$
   On a donc $\Phi\left(\frac{\sqrt{n}}{\sqrt{651}}\right) \approx 0.975 \approx \Phi(1.96)$ et $n \approx 1.96^2 \times 651 \approx 2501$. $2501 \geq 90$, ce qui légitime l'approximation.

□

5.5 Loi multinomiale

Définition 5.4 (Loi multinomiale). Le vecteur aléatoire N suit la loi multinomiale de paramètres n et $(p_1, \ldots, p_d)$ où $n \in \mathbb{N}^*$ et les p_i sont strictement positifs et de somme 1 si pour tout d-uple $(j_1, j_2, \ldots, j_d)$ d'entiers tels que $j_1 + j_2 + \cdots + j_d = n$,

$$P[N = (j_1, j_2, \ldots, j_d)] = \frac{n!}{j_1! j_2! \cdots j_d!} p_1^{j_1} p_2^{j_2} \cdots p_d^{j_d}.$$

Exemple 5.5. On considère 20 tirages d'une boule avec remise dans une urne contenant 1 boule bleue, 3 jaunes, 4 rouges et 2 vertes. Notons $N = (N_1, N_2, N_3, N_4)$ où N_i est le nombre de boules de la couleur i en numérotant les couleurs par ordre alphabétique (`b, j, r, v`). On a $(p_1, p_2, p_3, p_4) = (\frac{1}{10}, \frac{3}{10}, \frac{4}{10}, \frac{2}{10})$. La probabilité d'obtenir en 20 tirages 3 bleues, 5 jaunes, 10 rouges et 2 vertes est :

$$P(N = (3, 5, 10, 2)) = \frac{20!}{3!5!10!2!}\left(\frac{1}{10}\right)^3\left(\frac{3}{10}\right)^5\left(\frac{4}{10}\right)^{10}\left(\frac{2}{10}\right)^2 \simeq 0,004745.$$