

La loi hypergéométrique

Marc Lorenzi

6 juin 2025

1 Introduction

1.1 Les paramètres

On se donne $N \in \mathbb{N}$, $a \in \llbracket 0, N \rrbracket$ et $n \in \llbracket 0, N \rrbracket$. On pose $b = N - a$.

Soit E un ensemble fini de cardinal N . Soit $F \subseteq E$ de cardinal a . Nous nous plaçons dans l'espace probabilisé $(\Omega, \mathfrak{T}, \mathbb{P})$ où

- $\Omega = \mathcal{P}_n(E)$, l'ensemble des parties de E de cardinal n .
- $\mathfrak{T} = \mathcal{P}(\Omega)$, l'ensemble des parties de Ω .
- $\mathbb{P}$ est la probabilité uniforme.

1.2 La variable aléatoire

Soit $X : \Omega \longrightarrow \mathbb{N}$ la variable aléatoire définie, pour tout $A \in \Omega$, par

$$X(A) = |A \cap F|$$

Pour tout $A \in \Omega$, $A \cap F$ est inclus dans F , donc $X(A) \leq |F| = a$. On a donc $X(\Omega) \subseteq \llbracket 0, a \rrbracket$. Nous préciserons un peu plus loin quelles sont exactement les valeurs prises par X .

1.3 La loi de X

Soit $k \in \llbracket 0, a \rrbracket$. Pour tout $A \in \{X = k\}$, le couple $(A \cap F, A \cap (E \setminus F))$ appartient à $\mathcal{P}_k(F) \times \mathcal{P}_{n-k}(E \setminus F)$. Considérons l'application

$$\Phi : \{X = k\} \longrightarrow \mathcal{P}_k(F) \times \mathcal{P}_{n-k}(E \setminus F)$$

définie pour tout $A \in \{X = k\}$ par

$$\Phi(A) = (A \cap F, A \cap (E \setminus F))$$

Cette application est bijective, et $\Phi^{-1} = \Psi$ où $\Psi(U, V) = U \cup V$. Ainsi,

$$|\{X = k\}| = |\mathcal{P}_k(F)| \times |\mathcal{P}_{n-k}(E \setminus F)| = \binom{a}{k} \binom{b}{n-k}$$

En divisant par le cardinal de Ω , qui est $\binom{N}{n}$,

$$\mathbb{P}(X = k) = \frac{\binom{a}{k} \binom{b}{n-k}}{\binom{N}{n}}$$

1.4 Les valeurs prises par X

Soit $k \in \llbracket 0, n \rrbracket$. On a $\{X = k\} \neq \emptyset$ si et seulement si le cardinal de cet événement est différent de 0, c'est à dire $k \leq a$ et $n - k \leq b = N - a$, ou encore $a - (N - n) \leq k \leq a$. Ainsi,

$$X(\Omega) = \llbracket \max(0, a - (N - n)), \min(a, n) \rrbracket$$

1.5 Deux exemples

Une urne contient N boules. Parmi celles-ci, a sont blanches et $N - a = b$ sont noires. On tire n boules sans remise. Pour modéliser cette expérience, on prend par exemple $E = \llbracket 1, N \rrbracket$ et $F = \llbracket 1, a \rrbracket$. Un élément A de Ω est un *tirage*. L'entier $X(A)$ est le nombre de boules blanches du tirage. La figure ci-dessous donne, pour $k \in \llbracket 0, n \rrbracket$, la probabilité $\mathbb{P}(X = k)$.

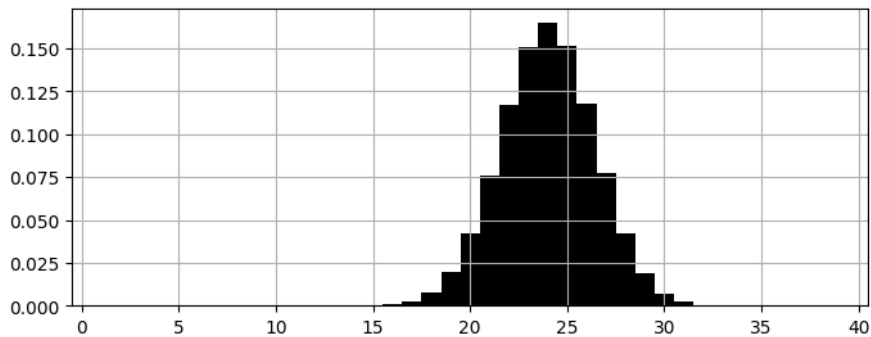


FIGURE 1 – Un tirage sans remise avec $N = 100$, $a = 40$, $n = 60$

Un exemple concret de tirage sans remise est celui du loto. Pour modéliser un tirage du loto, prenons $E = \llbracket 1, 49 \rrbracket$, de cardinal $N = 49$. L'ensemble F , de cardinal $a = 6$, est l'ensemble des numéros gagnants. Une *grille* est une partie A de E de cardinal $n = 6$. Pour toute grille A , $X(A)$ est le nombre de numéros gagnants de la grille. La figure ci-dessous donne, pour $k \in \llbracket 0, 6 \rrbracket$, la probabilité $\mathbb{P}(X = k)$.

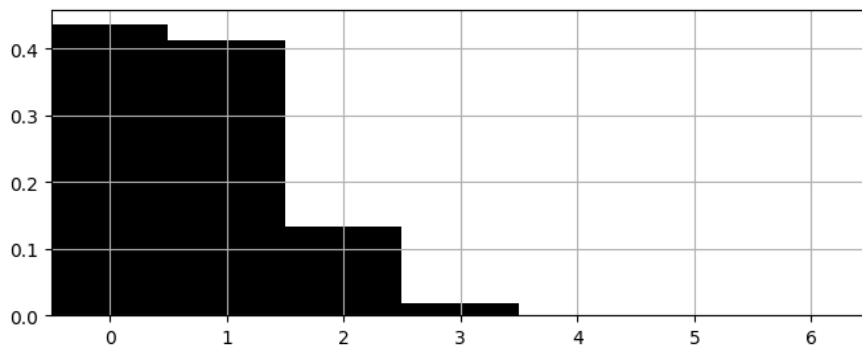


FIGURE 2 – Probabilité, pour $k \in \llbracket 0, 6 \rrbracket$, d'obtenir k numéros gagnants sur une grille de loto

2 L'espérance de X

2.1 Variables de Bernoulli

Nous supposons dorénavant que $N \geq 1$ et $n \geq 1$.

La variable aléatoire X peut s'écrire comme une somme de variables aléatoires suivant une même loi de Bernoulli.

Soit $A \in \Omega$. On a

$$X(A) = |A \cap F| = \sum_{x \in A \cap F} 1 = \sum_{f \in F} \mathbf{1}_A(f)$$

Pour tout $f \in F$, soit $X_f : \Omega \longrightarrow \{0, 1\}$ définie par $X_f(A) = \mathbf{1}_A(f)$. On a alors

$$X(A) = \sum_{f \in F} X_f(A)$$

et donc

$$X = \sum_{f \in F} X_f$$

La variable aléatoire X_f , à valeurs dans $\{0, 1\}$, suit une loi de Bernoulli.

Proposition 1. *Pour tout $f \in F$,*

$$X_f \sim \mathcal{B}\left(\frac{n}{N}\right)$$

Démonstration. Soit $f \in F$. Pour tout $A \in \Omega$, $A \in \{X_f = 1\}$ si et seulement si $f \in A$, c'est à dire $A = A' \cup \{f\}$ où $A' \in \mathcal{P}_{n-1}(E \setminus \{f\})$. On a donc

$$\{X_f = 1\} = \{A' \cup \{f\} : A' \in \mathcal{P}_{n-1}(E \setminus \{f\})\} \simeq \mathcal{P}_{n-1}(E \setminus \{f\})$$

De là,

$$|\{X_f = 1\}| = |\mathcal{P}_{n-1}(E \setminus \{f\})| = \binom{N-1}{n-1}$$

d'où

$$\mathbb{P}(X_f = 1) = \frac{\binom{N-1}{n-1}}{\binom{N}{n}} = \frac{n}{N}$$

d'où le résultat. $\square$

Conséquence. Pour tout $f \in F$,

$$\mathbb{E}(X_f) = \frac{n}{N} \quad \text{et} \quad \mathbb{V}(X_f) = \frac{n(N-n)}{N^2}$$

2.2 L'espérance de X

Notation. Notons p la proportion d'éléments de E qui appartiennent à F :

$$p = \frac{|F|}{|E|} = \frac{a}{N}$$

Proposition 2. *L'espérance de X est*

$$\mathbb{E}(X) = np$$

Démonstration. Par linéarité de l'espérance,

$$\mathbb{E}(X) = \sum_{f \in F} \mathbb{E}(X_f) = \sum_{f \in F} \frac{n}{N} = \frac{n}{N} |F| = \frac{na}{N} = np$$

□

3 La variance de X

Nous supposons dans la suite que $N \geq 2$ et $n \geq 2$.

3.1 Covariances

Les variables de Bernoulli X_f ne sont pas indépendantes. Déterminons leurs covariances.

Proposition 3. *Soient $f, g \in F$ distincts. On a*

$$\text{Cov}(X_f, X_g) = -\frac{n(N-n)}{N^2(N-1)}$$

Démonstration. La variable aléatoire $X_f X_g : \Omega \rightarrow \{0, 1\}$ suit une loi de Bernoulli. Pour tout $A \in \Omega$, $A \in \{X_f X_g = 1\}$ si et seulement si $f \in A$ et $g \in A$. On a donc

$$\begin{aligned} \{X_f X_g = 1\} &= \{A \in \Omega : f, g \in A\} \\ &= \{A' \cup \{f, g\} : A' \in \mathcal{P}_{n-2}(E \setminus \{f, g\})\} \\ &\simeq \mathcal{P}_{n-2}(E \setminus \{f, g\}) \end{aligned}$$

De là,

$$|\{X_f X_g = 1\}| = |\mathcal{P}_{n-2}(E \setminus \{f, g\})| = \binom{N-2}{n-2}$$

d'où

$$\mathbb{P}(X_f X_g = 1) = \frac{\binom{N-2}{n-2}}{\binom{N}{n}} = \frac{n(n-1)}{N(N-1)}$$

Ainsi,

$$X_f X_g \sim \mathcal{B}\left(\frac{n(n-1)}{N(N-1)}\right)$$

On en déduit

$$\text{Cov}(X_f, X_g) = \mathbb{E}(X_f X_g) - \mathbb{E}(X_f)\mathbb{E}(X_g) = \frac{n(n-1)}{N(N-1)} - \frac{nn}{NN} = -\frac{n(N-n)}{N^2(N-1)}$$

□

3.2 La variance de X

Proposition 4. *La variance de X est*

$$\mathbb{V}(X) = np(1-p) \frac{N-n}{N-1}$$

Démonstration. On a

$$\begin{aligned}\mathbb{V}(X) &= \mathbb{V}\left(\sum_{f \in F} X_f\right) \\ &= \sum_{f \in F} \mathbb{V}(X_f) + \sum_{f, g \in F, f \neq g} \text{Cov}(X_f, X_g)\end{aligned}$$

Comme nous l'avons déjà vu, on a pour tout $f \in F$,

$$\mathbb{V}(X_f) = \frac{n(N-n)}{N^2}$$

Par ailleurs,

$$|\{(f, g) \in F^2 : f \neq g\}| = a^2 - a = Np(Np-1)$$

De là,

$$\begin{aligned}\mathbb{V}(X) &= \frac{na(N-n)}{N^2} - Np(Np-1) \frac{n(N-n)}{N^2(N-1)} \\ &= \frac{np(N-n)}{N} - p(Np-1) \frac{n(N-n)}{N(N-1)}\end{aligned}$$

d'où, par une simplification facile, le résultat. $\square$