Cours 10

Variable Aléatoire

Mathématiques, niveau baccalauréat

Mohamed Kadhem Karray

mohamedkadhem.com/baccalaureat

12 novembre 2026

Plan

Introduction

10.1 Variable Aléatoire et Loi

10.2 Espérance

10.3 Variance et Écart-Type

10.4 Indépendance

10.5 Lois de Bernoulli et Binomiale

10.6 Loi des Grands Nombres

À retenir

Introduction

Lancer un dé, compter les succès d'une série d'essais, mesurer un gain : on associe un nombre au résultat d'une expérience aléatoire. Cette leçon étudie ce nombre, sa loi, sa moyenne et sa dispersion, puis montre pourquoi la fréquence d'un événement se rapproche de sa probabilité.

Où se place cette leçon

- **Prérequis :** une probabilité sur un univers fini, l'indépendance d'événements et la modélisation d'épreuves répétées (cours 7 [2]) ; compter les parties d'un ensemble fini (cours 17 [3]) ; la limite d'une suite (cours 1 [1]).
- **Les objets :** une variable aléatoire et sa loi ; son espérance, sa variance et son écart-type ; l'indépendance de variables aléatoires ; les lois de Bernoulli et binomiale.
- **Ce qu'on saura faire :** déterminer la loi d'une variable aléatoire, calculer son espérance et sa variance, reconnaître une loi binomiale, majorer la probabilité d'un écart à la moyenne.
- **Et ensuite :** la loi des grands nombres justifie l'estimation d'une probabilité par une fréquence observée, base de la statistique.

Dans toute la leçon, sauf mention contraire, P est une probabilité sur un univers fini Ω (Définition 7.1 du cours 7 [2]).

10.1 Variable Aléatoire et Loi

Une variable aléatoire associe un réel à chaque issue ; sa loi donne la probabilité de chacune de ses valeurs.

Définition 10.1 (Variable Aléatoire). Une *variable aléatoire* est une fonction $X : \Omega \to \mathbb{R}$. Pour tout réel x, on note

$$\{X = x\} = \{\omega \in \Omega : X(\omega) = x\},$$

et de même $\{X \geq t\}$, $\{|X| \geq t\}\ldots$ Ce sont des événements, et l'on écrit $\mathrm{P}(X = x)$ pour $\mathrm{P}(\{X = x\})$.

Définition 10.2 (Loi d'une Variable Aléatoire). Soit X une variable aléatoire, et $x_1, \ldots, x_m$ les valeurs *distinctes* prises par X (il y en a un nombre fini, car Ω est fini). La *loi* de X est la donnée des couples $(x_i, \mathrm{P}(X = x_i))$, pour $i = 1, \ldots, m$. On la présente souvent dans un tableau.

Lire la définition. Si x n'est pas une valeur de X, l'événement $\{X = x\}$ est vide, et $\mathrm{P}(X = x) = 0$ (Théorème 7.1 (i) du cours 7 [2]).

La Somme des Probabilités

Théorème 10.1 (Somme des Probabilités). Si $x_1, \ldots, x_m$ sont les valeurs distinctes prises par X, alors

$$\sum_{i=1}^{m} \mathrm{P}\left(X = x_i\right) = 1.$$

Preuve. Chaque issue ω appartient à un et un seul des événements $\{X = x_i\}$, celui de la valeur $X(\omega)$: ces événements sont deux à deux incompatibles, et leur réunion est Ω. Par le Corollaire 7.1 et le Théorème 7.1 (i) du cours 7 [2], la somme de leurs probabilités est $\mathrm{P}(\Omega) = 1$.

Exemple : Lancer d'un Dé

Exemple 10.1 (Lancer d'un Dé). On lance un dé équilibré : $\Omega = \{1, 2, \ldots, 6\}$, muni de la probabilité uniforme (Définition 7.3 et Remarque 7.2 du cours 7 [2]). Soit $X(\omega) = 1$ si ω est pair, et $X(\omega) = -1$ si ω est impair. Les valeurs de X sont -1 et 1 ; $\{X = 1\} = \{2, 4, 6\}$ et $\{X = -1\} = \{1, 3, 5\}$, donc, par le Théorème 7.2 du cours 7 [2], la loi de X est

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

Exemple : le Carré d'une Issue

Exemple 10.2 (Le Carré). Soit $\Omega = \{-1, 0, 1\}$, muni de la probabilité uniforme, et $Y(\omega) = \omega^2$. Les valeurs de Y sont 0 et 1 ; $\{Y = 0\} = \{0\}$ et $\{Y = 1\} = \{-1, 1\}$, d'où $\mathrm{P}(Y = 0) = \frac{1}{3}$ et $\mathrm{P}(Y = 1) = \frac{2}{3}$ (Théorème 7.2 du cours 7 [2]). On vérifie que $\frac{1}{3} + \frac{2}{3} = 1$ (Théorème 10.1).

10.2 Espérance

L'espérance est la moyenne des valeurs d'une variable aléatoire, pondérées par leurs probabilités ; elle est linéaire.

Définition 10.3 (Espérance). Soit X une variable aléatoire, de valeurs distinctes $x_1, \ldots, x_m$. L'*espérance* de X est le réel

$$\mathrm{E}[X] = \sum_{i=1}^{m} x_i \, \mathrm{P}(X = x_i).$$

Lire la définition. C'est la moyenne des valeurs de X, chacune comptée avec sa probabilité. Elle ne dépend que de la loi de X.

Exemple 10.3 (Lancer d'un Dé). Avec la variable X de l'Exemple 10.1,

$$\mathrm{E}[X] = (-1) \times \frac{1}{2} + 1 \times \frac{1}{2} = 0.$$

L'Espérance par les Issues

Théorème 10.2 (L'Espérance par les Issues). Pour toute variable aléatoire X,

$$\mathrm{E}[X] = \sum_{\omega \in \Omega} X(\omega)\, \mathrm{P}(\{\omega\}).$$

Preuve. Soient $x_1, \ldots, x_m$ les valeurs distinctes de X. Chaque issue appartient à un et un seul des événements $\{X = x_i\}$ (preuve du Théorème 10.1) : on regroupe les termes de la somme selon la valeur de $X(\omega)$. Sur $\{X = x_i\}$, on a $X(\omega) = x_i$, donc

$$\sum_{\omega \in \Omega} X(\omega)\, \mathrm{P}(\{\omega\}) = \sum_{i=1}^{m} x_i \sum_{\omega \in \{X = x_i\}} \mathrm{P}(\{\omega\}) = \sum_{i=1}^{m} x_i\, \mathrm{P}(X = x_i),$$

par la Définition 7.1 du cours 7 [2], qui définit $\mathrm{P}(A)$ comme la somme des poids des issues de A.

Linéarité de l'Espérance

Si X et Y sont deux variables aléatoires et a, b deux réels, $aX + bY$, XY, $X^2 \ldots$ sont encore des fonctions de Ω dans $\mathbb{R}$, donc des variables aléatoires ; un réel c est une variable aléatoire constante.

Théorème 10.3 (Linéarité). Soient X et Y deux variables aléatoires, et a, b deux réels.

(i) $\mathrm{E}[aX + bY] = a\,\mathrm{E}[X] + b\,\mathrm{E}[Y]$;

(ii) $\mathrm{E}[b] = b$, et $\mathrm{E}[aX + b] = a\,\mathrm{E}[X] + b$;

(iii) si $X \geq 0$ (c'est-à-dire $X(\omega) \geq 0$ pour tout ω), alors $\mathrm{E}[X] \geq 0$.

Preuve. (i) Par le Théorème 10.2, $\mathrm{E}[aX + bY] = \sum_{\omega \in \Omega} (aX(\omega) + bY(\omega))\,\mathrm{P}(\{\omega\})$, et l'on sépare la somme en deux. (ii) La variable constante b a une seule valeur, b, et $\mathrm{P}(\Omega) = 1$; on applique ensuite (i) avec la variable constante égale à 1 à la place de Y. (iii) Dans la somme du Théorème 10.2, tous les termes sont positifs.

Espérance d'une Fonction de X

Théorème 10.4 (Transfert). Soient X une variable aléatoire, de valeurs distinctes $x_1, \ldots, x_m$, et f une fonction définie sur $\{x_1, \ldots, x_m\}$, à valeurs réelles. Alors

$$\mathrm{E}\left[f\left(X\right)\right] = \sum_{i=1}^{m} f\left(x_i\right) \mathrm{P}\left(X = x_i\right).$$

En particulier, pour tout entier $k \geq 1$, $\mathrm{E}\left[X^k\right] = \sum_{i=1}^{m} x_i^k \, \mathrm{P}\left(X = x_i\right)$.

Preuve. Par le Théorème 10.2 appliqué à la variable $f\left(X\right)$, puis en regroupant les issues selon la valeur de X comme dans sa preuve,

$$\mathrm{E}\left[f\left(X\right)\right] = \sum_{\omega \in \Omega} f\left(X\left(\omega\right)\right) \mathrm{P}\left(\{\omega\}\right) = \sum_{i=1}^{m} f\left(x_i\right) \sum_{\omega \in \{X = x_i\}} \mathrm{P}\left(\{\omega\}\right) = \sum_{i=1}^{m} f\left(x_i\right) \mathrm{P}\left(X = x_i\right).$$

Lire le théorème. Les $f\left(x_i\right)$ peuvent ne pas être distincts : la formule ne découle donc pas directement de la Définition 10.3.

10.3 Variance et Écart-Type

La variance mesure la dispersion d'une variable aléatoire autour de son espérance ; elle se calcule par la formule de König-Huygens.

Variance et Écart-Type

Définition 10.4 (Variance et Écart-Type). Soit X une variable aléatoire. La *variance* de X est le réel

$$\mathrm{Var}(X) = \mathrm{E}\left[(X - \mathrm{E}[X])^2\right],$$

qui est positif par le Théorème 10.3 (iii). L'*écart-type* de X est $\sigma(X) = \sqrt{\mathrm{Var}(X)}$.

Lire la définition. Par le Théorème 10.4, $\mathrm{Var}(X) = \sum_{i=1}^{m} (x_i - \mathrm{E}[X])^2 \, \mathrm{P}(X = x_i)$: c'est la moyenne des carrés des écarts à l'espérance. L'écart-type, exprimé dans la même unité que X, donne l'ordre de grandeur des écarts de X à son espérance.

Formule de König-Huygens

Théorème 10.5 (Formule de König-Huygens). Pour toute variable aléatoire X,

$$\mathrm{Var}(X) = \mathrm{E}\left[X^2\right] - \mathrm{E}[X]^2.$$

Preuve. Notons $\mu = \mathrm{E}[X]$, qui est un réel. On développe $(X-\mu)^2 = X^2 - 2\mu X + \mu^2$, puis, par le Théorème 10.3 (i) et (ii),

$$\mathrm{Var}(X) = \mathrm{E}\left[X^2\right] - 2\mu\,\mathrm{E}[X] + \mu^2 = \mathrm{E}\left[X^2\right] - 2\mu^2 + \mu^2 = \mathrm{E}\left[X^2\right] - \mu^2.$$

Exemple 10.4 (Lancer d'un Dé). Avec la variable X de l'Exemple 10.1, $X^2 = 1$, donc $\mathrm{E}\left[X^2\right] = 1$ (Théorème 10.3 (ii)) ; comme $\mathrm{E}[X] = 0$ (Exemple 10.3), le Théorème 10.5 donne $\mathrm{Var}(X) = 1$, et $\sigma(X) = 1$.

Variance d'une Fonction Affine de X

Théorème 10.6 (Variance de $aX + b$). Pour toute variable aléatoire X et tous réels a et b,

$$\operatorname{Var}(aX + b) = a^2 \operatorname{Var}(X) \qquad \text{et} \qquad \sigma(aX + b) = |a|\, \sigma(X).$$

Preuve. Par le Théorème 10.3 (ii), $\mathrm{E}[aX + b] = a\,\mathrm{E}[X] + b$, donc $aX + b - \mathrm{E}[aX + b] = a(X - \mathrm{E}[X])$. Par la Définition 10.4 et le Théorème 10.3 (i),

$$\operatorname{Var}(aX + b) = \mathrm{E}\left[a^2 (X - \mathrm{E}[X])^2\right] = a^2 \operatorname{Var}(X),$$

et $\sqrt{a^2} = |a|$.

Lire le théorème. Ajouter une constante déplace les valeurs sans changer leur dispersion ; multiplier par a multiplie les écarts par $|a|$.

10.4 Indépendance

Deux variables aléatoires sont indépendantes quand les événements qu'elles définissent le sont ; l'espérance du produit et la variance de la somme s'en déduisent.

Variables Aléatoires Indépendantes

Définition 10.5 (Variables Indépendantes).

(i) Deux variables aléatoires X et Y sont *indépendantes* si, pour tous réels x et y, les événements $\{X = x\}$ et $\{Y = y\}$ sont indépendants (Définition 7.6 du cours 7 [2]) :

$$\mathrm{P}\left(\{X = x\} \cap \{Y = y\}\right) = \mathrm{P}\left(X = x\right)\mathrm{P}\left(Y = y\right).$$

(ii) Des variables aléatoires $X_1, \ldots, X_n$ sont *indépendantes* si, pour tous réels $a_1, \ldots, a_n$, les événements $\{X_1 = a_1\}, \ldots, \{X_n = a_n\}$ sont mutuellement indépendants (Définition 7.7 du cours 7 [2]).

Lire la définition. Il suffit de vérifier l'égalité de (i) quand x et y sont des valeurs de X et de Y : sinon, ses deux membres sont nuls. Dans (ii), deux des variables X_i et X_j $(i \neq j)$ sont indépendantes au sens de (i), par la Définition 7.7 appliquée à une sous-famille de deux événements.

Remarque 10.1 (Épreuves Répétées). Quand une expérience est répétée dans les mêmes conditions et que X_i ne dépend que du résultat de la i-ème répétition, on suppose $X_1, \ldots, X_n$ indépendantes : c'est l'hypothèse de modèle de la Remarque 7.4 du cours 7 [2].

Théorème 10.7 (Produit et Somme).

(i) Si X et Y sont indépendantes, alors $\mathrm{E}[XY] = \mathrm{E}[X]\,\mathrm{E}[Y]$.

(ii) Si $X_1, \ldots, X_n$ sont indépendantes, alors $\mathrm{Var}(X_1 + \cdots + X_n) = \mathrm{Var}(X_1) + \cdots + \mathrm{Var}(X_n)$. En particulier, si X et Y sont indépendantes, $\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$.

Preuve. (i) Soient $x_1, \ldots, x_m$ et $y_1, \ldots, y_r$ les valeurs distinctes de X et de Y. Chaque issue appartient à un et un seul des événements $\{X = x_i\} \cap \{Y = y_j\}$, sur lequel XY vaut $x_i y_j$. En regroupant ainsi les termes de la somme du Théorème 10.2, puis par la Définition 7.1 du cours 7 [2] et la Définition 10.5 (i),

$$\mathrm{E}[XY] = \sum_{i=1}^{m} \sum_{j=1}^{r} x_i y_j \,\mathrm{P}(\{X = x_i\} \cap \{Y = y_j\}) = \left(\sum_{i=1}^{m} x_i \,\mathrm{P}(X = x_i) \right) \left(\sum_{j=1}^{r} y_j \,\mathrm{P}(Y = y_j) \right).$$

Preuve du Théorème 10.7 (ii)

Notons $\mu_i = \mathrm{E}[X_i]$ et $S = X_1 + \cdots + X_n$; alors $\mathrm{E}[S] = \mu_1 + \cdots + \mu_n$ (Théorème 10.3 (i)), et $(S - \mathrm{E}[S])^2$ est la somme des $(X_i - \mu_i)^2$ et des produits $(X_i - \mu_i)(X_j - \mu_j)$, pour $i \neq j$. Pour $i \neq j$, X_i et X_j sont indépendantes (lecture de la Définition 10.5), donc, par le Théorème 10.3 (i) et (ii) puis par le Théorème 10.7 (i),

$$\mathrm{E}[(X_i - \mu_i)(X_j - \mu_j)] = \mathrm{E}[X_i X_j] - \mu_i \mu_j = 0.$$

Il reste, par le Théorème 10.3 (i), $\mathrm{Var}(S) = \sum_{i=1}^{n} \mathrm{E}\left[(X_i - \mu_i)^2\right] = \sum_{i=1}^{n} \mathrm{Var}(X_i)$.

Exemple : Deux Dés

Exemple 10.5 (Deux Dés). On lance deux dés équilibrés, avec l'univers $\Omega = \{1, \ldots, 6\}^2$ muni de la probabilité uniforme (Exemple 7.8 du cours 7 [2]). Soient X et Y les résultats du premier et du second dé.

Étape 1 : indépendance. Pour i, j dans $\{1, \ldots, 6\}$, $\{X = i\} \cap \{Y = j\} = \{(i, j)\}$, de probabilité $\frac{1}{36} = \frac{1}{6} \times \frac{1}{6}$, et $\mathrm{P}(X = i) = \mathrm{P}(Y = j) = \frac{6}{36} = \frac{1}{6}$ (Théorème 7.2 du cours 7 [2]) : X et Y sont indépendantes (Définition 10.5 (i) et sa lecture).

Étape 2 : variance de X. $\mathrm{E}[X] = \frac{1 + 2 + \cdots + 6}{6} = \frac{7}{2}$ et $\mathrm{E}\left[X^2\right] = \frac{1 + 4 + \cdots + 36}{6} = \frac{91}{6}$ (Théorème 10.4), donc $\mathrm{Var}(X) = \frac{91}{6} - \frac{49}{4} = \frac{35}{12}$ (Théorème 10.5) ; de même pour Y.

Étape 3 : la somme. Par le Théorème 10.7 (ii), $\mathrm{Var}(X + Y) = \frac{35}{12} + \frac{35}{12} = \frac{35}{6}$.

10.5 Lois de Bernoulli et Binomiale

Une variable de Bernoulli code le succès d'un essai ; le nombre de succès dans n essais indépendants suit une loi binomiale.

Loi de Bernoulli

Définition 10.6 (Loi de Bernoulli). Soit $p \in [0,1]$. Une variable aléatoire X suit la *loi de Bernoulli de paramètre p* si elle ne prend que les valeurs 0 et 1, avec

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

On dit aussi que X est une variable de Bernoulli de paramètre p ; l'événement $\{X=1\}$ est appelé *succès*.

Théorème 10.8 (Espérance et Variance). Si X suit la loi de Bernoulli de paramètre p, alors

$$\mathrm{E}[X] = p \qquad \text{et} \qquad \mathrm{Var}(X) = p(1-p).$$

Preuve. Par la Définition 10.3 et le Théorème 10.4 (les termes de valeur 0 sont nuls), $\mathrm{E}[X] = 1 \times p + 0 \times (1-p) = p$ et $\mathrm{E}\left[X^2\right] = 1^2 \times p + 0^2 \times (1-p) = p$. Par le Théorème 10.5, $\mathrm{Var}(X) = p - p^2 = p(1-p)$.

Coefficients Binomiaux

Définition 10.7 (Factorielle et Coefficient Binomial). Pour tout entier $n \geq 1$, on pose $n! = 1 \times 2 \times \cdots \times n$, et l'on convient que $0! = 1$. Pour tous entiers $n \geq 0$ et $k \in \{0, 1, \ldots, n\}$, le *coefficient binomial* est

$$\binom{n}{k} = \frac{n!}{k!\,(n-k)!}.$$

On le note aussi C_n^k.

Théorème 10.9 (Nombre de Parties). Pour $0 \leq k \leq n$, un ensemble à n éléments a exactement $\binom{n}{k}$ parties à k éléments. Ce résultat de dénombrement est admis ici (cours 17 [3]).

Exemple 10.6 (Valeurs Simples). $\binom{n}{0} = \binom{n}{n} = 1$, $\binom{n}{1} = n$ pour $n \geq 1$, et $\binom{4}{2} = \frac{24}{2 \times 2} = 6$: les parties à deux éléments de $\{1, 2, 3, 4\}$ sont $\{1, 2\}$, $\{1, 3\}$, $\{1, 4\}$, $\{2, 3\}$, $\{2, 4\}$, $\{3, 4\}$.

Loi Binomiale

Définition 10.8 (Loi Binomiale). Soient $n \geq 1$ un entier et $p \in [0,1]$. Une variable aléatoire X suit la *loi binomiale de paramètres n et p*, notée $\mathcal{B}(n,p)$, si elle prend ses valeurs dans $\{0,1,\ldots,n\}$ et si

$$\mathrm{P}(X=k) = \binom{n}{k} p^k (1-p)^{n-k} \qquad \text{pour tout } k \in \{0,1,\ldots,n\},$$

avec la convention $0^0 = 1$ (utile si $p=0$ ou $p=1$).

Théorème 10.10 (Nombre de Succès). Soient $X_1,\ldots,X_n$ des variables aléatoires indépendantes, qui suivent toutes la loi de Bernoulli de *même* paramètre p. Alors $S = X_1 + \cdots + X_n$, le nombre de succès, suit la loi $\mathcal{B}(n,p)$.

Lire le théorème. C'est la situation de n répétitions indépendantes d'une épreuve à deux issues, succès de probabilité p ou échec (Remarque 10.1) : on parle de *schéma de Bernoulli*. Si les paramètres diffèrent, S ne suit pas, en général, une loi binomiale.

Preuve du Théorème 10.10

Comme les X_i valent 0 ou 1, S prend ses valeurs dans $\{0, 1, \ldots, n\}$. Soit $k \in \{0, \ldots, n\}$.

Étape 1 : découper $\{S = k\}$. Pour toute partie J de $\{1, \ldots, n\}$ à k éléments, soit A_J l'événement « $X_i = 1$ pour $i \in J$, et $X_i = 0$ pour $i \notin J$ ». Une issue ω est dans $\{S = k\}$ si et seulement si exactement k des nombres $X_i(\omega)$ valent 1, c'est-à-dire si ω est dans l'un des A_J ; et elle est dans un seul d'entre eux, celui de la partie $J = \{i : X_i(\omega) = 1\}$. Les A_J sont donc deux à deux incompatibles, de réunion $\{S = k\}$.

Étape 2 : probabilité de A_J. A_J est l'intersection de n événements $\{X_i = a_i\}$, avec $a_i = 1$ pour $i \in J$ et $a_i = 0$ sinon ; ils sont mutuellement indépendants (Définition 10.5 (ii)), donc, par la Définition 7.7 du cours 7 [2] (si $n = 1$, directement par la Définition 10.6), $\mathrm{P}(A_J)$ est le produit de k facteurs p et de $n - k$ facteurs $1 - p$: $\mathrm{P}(A_J) = p^k (1 - p)^{n-k}$.

Étape 3 : conclure. Il y a $\binom{n}{k}$ parties J (Théorème 10.9). Par le Corollaire 7.1 du cours 7 [2], $\mathrm{P}(S = k) = \binom{n}{k} p^k (1 - p)^{n-k}$.

Théorème 10.11 (Espérance et Variance). Si X suit la loi $\mathcal{B}(n, p)$, alors

$$\mathrm{E}[X] = np \qquad \text{et} \qquad \mathrm{Var}(X) = np(1-p).$$

Preuve. L'espérance et la variance d'une variable ne dépendent que de sa loi (Définition 10.3, Théorèmes 10.4 et 10.5). Il suffit donc de les calculer pour une variable $S = X_1 + \cdots + X_n$ comme au Théorème 10.10, qui suit la même loi. On admet qu'un tel modèle existe pour tous n et p (par exemple n lancers indépendants d'une pièce qui tombe sur pile avec la probabilité p, Remarque 10.1). Par le Théorème 10.3 (i), le Théorème 10.7 (ii) et le Théorème 10.8,

$$\mathrm{E}[S] = \mathrm{E}[X_1] + \cdots + \mathrm{E}[X_n] = np, \qquad \mathrm{Var}(S) = \mathrm{Var}(X_1) + \cdots + \mathrm{Var}(X_n) = np(1-p).$$

Exemple : Quatre Lancers d'un Dé

Exemple 10.7 (Quatre Lancers d'un Dé). On lance quatre fois un dé équilibré ; soit X le nombre de 6 obtenus. Pour $i = 1, \ldots, 4$, soit $X_i = 1$ si le i-ème lancer donne 6, et $X_i = 0$ sinon. Le dé étant équilibré, on munit l'univers des résultats de chaque lancer, $\{1, \ldots, 6\}$, de la probabilité uniforme, donc X_i suit la loi de Bernoulli de paramètre $\frac{1}{6}$ (Remarque 7.2 et Théorème 7.2 du cours 7 [2], Définition 10.6). Les lancers étant indépendants, les X_i le sont (Remarque 10.1), et $X = X_1 + \cdots + X_4$ suit la loi $\mathcal{B}\left(4, \frac{1}{6}\right)$ (Théorème 10.10), donc

$$\mathrm{P}(X = 2) = \binom{4}{2} \left(\frac{1}{6}\right)^2 \left(\frac{5}{6}\right)^2 = 6 \times \frac{25}{1296} = \frac{25}{216}, \qquad \mathrm{E}[X] = \frac{4}{6} = \frac{2}{3} \quad \text{(Théorème 10.11).}$$

10.6 Loi des Grands Nombres

Les inégalités de Markov et de Tchebychev majorent la probabilité d'un grand écart ; on en déduit que la moyenne de nombreuses répétitions s'écarte rarement de l'espérance.

Inégalité de Markov

Théorème 10.12 (Inégalité de Markov). Soit X une variable aléatoire à valeurs positives ($X \geq 0$). Pour tout réel $t > 0$,

$$\mathrm{P}(X \geq t) \leq \frac{\mathrm{E}[X]}{t}.$$

Preuve. Soient $x_1, \ldots, x_m$ les valeurs distinctes de X, toutes positives. Dans la somme de la Définition 10.3, on garde les seuls termes où $x_i \geq t$, ce qui ne peut que diminuer la somme, puis on minore chacun d'eux en remplaçant x_i par t :

$$\mathrm{E}[X] \geq \sum_{i:\, x_i \geq t} x_i \,\mathrm{P}(X = x_i) \geq t \sum_{i:\, x_i \geq t} \mathrm{P}(X = x_i) = t\,\mathrm{P}(X \geq t),$$

la dernière égalité par le Corollaire 7.1 du cours 7 [2], car $\{X \geq t\}$ est la réunion des événements deux à deux incompatibles $\{X = x_i\}$, $x_i \geq t$ (si aucun x_i n'est $\geq t$, les deux membres sont nuls). On divise par $t > 0$.

Lire le théorème. L'hypothèse $X \geq 0$ est indispensable : si X vaut -100 ou 1, avec la probabilité $\frac{1}{2}$ chacun, alors $\mathrm{P}(X \geq 1) = \frac{1}{2}$, alors que $\mathrm{E}[X] < 0$.

Inégalité de Tchebychev

Théorème 10.13 (Inégalité de Tchebychev). Pour toute variable aléatoire X et tout réel $t > 0$,

$$\mathrm{P}\left(|X - \mathrm{E}[X]| \geq t\right) \leq \frac{\mathrm{Var}(X)}{t^2}.$$

Preuve. Soit $Y = (X - \mathrm{E}[X])^2$: c'est une variable aléatoire positive, et $\mathrm{E}[Y] = \mathrm{Var}(X)$ (Définition 10.4). Comme $t > 0$ et que la fonction carré est strictement croissante sur $[0, +\infty[$, pour toute issue, $|X - \mathrm{E}[X]| \geq t$ équivaut à $Y \geq t^2$: les deux événements sont égaux. On applique le Théorème 10.12 à Y, avec $t^2 > 0$.

Lire le théorème. Avec $t = k\,\sigma(X)$, si $\sigma(X) > 0$ et $k > 0$: la probabilité que X s'écarte de son espérance d'au moins k écarts-types est au plus $\frac{1}{k^2}$. Par exemple, au plus $\frac{1}{4}$ pour deux écarts-types. Pour un écart *strictement* supérieur à t, on utilise l'inclusion $\{|X - \mathrm{E}[X]| > t\} \subset \{|X - \mathrm{E}[X]| \geq t\}$ et le Théorème 7.1 (ii) du cours 7 [2].

Inégalité de Concentration

Théorème 10.14 (Inégalité de Concentration). Soient $X_1, \ldots, X_n$ des variables aléatoires indépendantes, de même loi, d'espérance μ et de variance V, et

$$M_n = \frac{X_1 + \cdots + X_n}{n}$$

leur moyenne. Pour tout réel $t > 0$,

$$\mathrm{P}\left(|M_n - \mu| \geq t\right) \leq \frac{V}{nt^2}.$$

Preuve. Par le Théorème 10.3 (i), $\mathrm{E}\left[M_n\right] = \frac{n\mu}{n} = \mu$. Par le Théorème 10.6 puis le Théorème 10.7 (ii), $\mathrm{Var}\left(M_n\right) = \frac{1}{n^2}\mathrm{Var}\left(X_1 + \cdots + X_n\right) = \frac{nV}{n^2} = \frac{V}{n}$. On applique le Théorème 10.13 à M_n.

Loi des Grands Nombres

Pour faire tendre n vers $+\infty$, on change de modèle à chaque n : sur un univers fini fixé, il n'existe pas, en général, de suite infinie de variables indépendantes de même loi.

Théorème 10.15 (Loi des Grands Nombres). Soient μ et $V \geq 0$ deux réels, et $t > 0$. Pour tout entier $n \geq 1$, soient $X_1, \ldots, X_n$ des variables aléatoires indépendantes, de même loi, d'espérance μ et de variance V, définies sur un univers fini qui peut dépendre de n, et $p_n = \mathrm{P}(|M_n - \mu| \geq t)$. Alors

$$\lim_{n \to +\infty} p_n = 0.$$

Preuve. Par le Théorème 7.1 (ii) du cours 7 [2] et le Théorème 10.14, $0 \leq p_n \leq \dfrac{V}{t^2} \times \dfrac{1}{n}$ pour tout $n \geq 1$. Comme $\lim\limits_{n \to +\infty} \dfrac{1}{n} = 0$, le majorant tend vers 0, et le théorème des gendarmes donne la limite (exemples de la Définition 1.6, Théorèmes 1.6 et 1.8 (i) du cours 1 [1]).

Corollaire 10.1 (Fréquence d'un Succès). Soient $p \in [0,1]$ et $t > 0$. Pour tout $n \geq 1$, soit S_n le nombre de succès dans n répétitions indépendantes d'une épreuve dont le succès a la probabilité p (sur un univers fini qui dépend de n, comme au Théorème 10.15), de sorte que S_n suit la loi $\mathcal{B}(n,p)$ (Théorème 10.10). Alors

$$\mathrm{P}\left(\left|\frac{S_n}{n} - p\right| \geq t\right) \leq \frac{p(1-p)}{nt^2} \qquad \text{et} \qquad \lim_{n\to+\infty} \mathrm{P}\left(\left|\frac{S_n}{n} - p\right| \geq t\right) = 0.$$

Preuve. Soit $X_i = 1$ si la i-ème répétition est un succès, et $X_i = 0$ sinon : $S_n = X_1 + \cdots + X_n$, où les X_i sont indépendantes (Remarque 10.1) et suivent la loi de Bernoulli de paramètre p (Définition 10.6). Ainsi S_n/n est la moyenne M_n de n variables de Bernoulli indépendantes de paramètre p, d'espérance p et de variance $p(1-p)$ (Théorème 10.8) : on applique les Théorèmes 10.14 et 10.15.

Remarque 10.2 (Ce que Dit la Loi). La fréquence S_n/n n'est pas assurée d'être proche de p ; c'est la *probabilité* qu'elle s'en écarte d'au moins t qui tend vers 0. Pour n donné, l'inégalité du Corollaire 10.1 en donne un majorant explicite.

Exemple 10.8 (Mille Lancers d'une Pièce). Pour 1000 lancers indépendants d'une pièce équilibrée, avec $p = \frac{1}{2}$ (Remarque 7.2 et Théorème 7.2 du cours 7 [2]) et $t = 0{,}05$, le Corollaire 10.1 donne $\mathrm{P}\left(\left|\frac{S_{1000}}{1000} - \frac{1}{2}\right| \geq 0{,}05\right) \leq \frac{1/4}{1000 \times 0{,}0025} = 0{,}1.$

À retenir

On retient la loi d'une variable aléatoire, l'espérance et sa linéarité, la variance et la formule de König-Huygens, l'indépendance, les lois de Bernoulli et binomiale, et les inégalités qui mènent à la loi des grands nombres.

À retenir

Ce que cette leçon a établi.

- Une variable aléatoire est une fonction $X : \Omega \to \mathbb{R}$; sa loi donne $\mathrm{P}(X = x_i)$ pour chaque valeur, et ces probabilités ont pour somme 1 (Définitions 10.1 et 10.2, Théorème 10.1).
- $\mathrm{E}[X] = \sum x_i \mathrm{P}(X = x_i)$ est linéaire, et $\mathrm{E}[f(X)] = \sum f(x_i)\,\mathrm{P}(X = x_i)$ (Définition 10.3, Théorèmes 10.2 à 10.4).
- $\mathrm{Var}(X) = \mathrm{E}\left[X^2\right] - \mathrm{E}[X]^2$ et $\mathrm{Var}(aX + b) = a^2 \mathrm{Var}(X)$ (Définition 10.4, Théorèmes 10.5 et 10.6).
- Pour des variables indépendantes : $\mathrm{E}[XY] = \mathrm{E}[X]\,\mathrm{E}[Y]$, et la variance d'une somme est la somme des variances (Définition 10.5, Théorème 10.7).
- Bernoulli : $\mathrm{E} = p$, $\mathrm{Var} = p(1-p)$; le nombre de succès dans n épreuves indépendantes, de même probabilité de succès p, suit $\mathcal{B}(n, p)$, d'espérance np et de variance $np(1-p)$ (Théorèmes 10.8, 10.10 et 10.11).
- Markov, Tchebychev, concentration ; la probabilité que la moyenne s'écarte de μ d'au moins t tend vers 0 (Théorèmes 10.12 à 10.15, Corollaire 10.1).

Merci pour votre attention.

Références

[1] Mohamed Kadhem Karray.
Cours 1. Suites numériques, niveau baccalauréat.
https://mohamedkadhem.com/baccalaureat/.

[2] Mohamed Kadhem Karray.
Cours 7. Probabilité, niveau baccalauréat.
https://mohamedkadhem.com/baccalaureat/.

[3] Mohamed Kadhem Karray.
Cours 17. Dénombrement, combinatoire, niveau baccalauréat.
https://mohamedkadhem.com/baccalaureat/.