

Fiche de cours Probabilités

Univers fini, conditionnement, indépendance, variables aléatoires, espérance, variance, inégalités de concentration

définitions · théorèmes et hypothèses · méthodes · démonstrations exigibles

1 UNIVERS, ÉVÉNEMENTS, VARIABLES ALÉATOIRES

Définition. Un **univers** est un ensemble *fini non vide* Ω , dont les éléments ω modélisent les issues d'une expérience aléatoire ; un **événement** est une partie A de Ω , un **événement élémentaire** un singleton. *Tout ce chapitre se limite au cas d'un univers fini.*

Notation. Le vocabulaire probabiliste relie le vocabulaire ensembliste : $\omega \in A$ « ω réalise A » ; $A \cap B$ « A et B » ; $A \cup B$ « A ou B » ; $\bar{A} = \Omega \setminus A$ événement **contraire** ; $A \subset B$ « A implique B » ; $A \cap B = \emptyset$ « A et B **incompatibles** (ou disjoints) ».

Définition. $(A_1, \dots, A_n)$ est un **système complet d'événements** lorsque les A_i sont deux à deux incompatibles et $\bigcup_{i=1}^n A_i = \Omega$. Exemple fondamental : $(A, \bar{A})$, et plus généralement $(\{X = x\})_{x \in X(\Omega)}$ pour une variable aléatoire X .

Définition. Une **variable aléatoire** sur Ω est une *application* $X : \Omega \rightarrow E$, où E est un ensemble *quelconque*. Pour $A \subset E$ et $x \in E$ on note $\{X \in A\} = (X \in A) = \{\omega \in \Omega, X(\omega) \in A\}$ et $\{X = x\} = \{X \in \{x\}\}$.

Remarque. Une variable aléatoire n'est ni un nombre ni une inconnue : c'est une fonction. On ne parlera d'espérance (resp. de variance) que lorsque $E \subset \mathbb{C}$ (resp. $E \subset \mathbb{R}$).

2 ESPACES PROBABILISÉS FINIS

Définition. Une **probabilité** sur Ω est une application $\mathbb{P} : \mathcal{P}(\Omega) \rightarrow [0; 1]$ telle que $\mathbb{P}(\Omega) = 1$ et $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$ pour tous événements *incompatibles* A, B . Le couple $(\Omega, \mathbb{P})$ est un **espace probabilisé fini**.

Définition. Une **distribution de probabilités** sur un ensemble E est une famille $(p_x)_{x \in E}$ d'éléments de $\mathbb{R}_+$ de somme 1.

Théorème. Une probabilité sur Ω est entièrement déterminée par la distribution $(\mathbb{P}(\{\omega\}))_{\omega \in \Omega}$, et réciproquement toute distribution $(p_\omega)_{\omega \in \Omega}$ provient d'une unique probabilité, donnée par $\mathbb{P}(A) = \sum_{\omega \in A} p_\omega$.

Définition. La **probabilité uniforme** sur Ω est celle de distribution constante : $\mathbb{P}(\{\omega\}) = \frac{1}{\text{Card } \Omega}$, et alors $\mathbb{P}(A) = \frac{\text{Card } A}{\text{Card } \Omega}$. C'est le modèle d'**équiprobabilité** : tout calcul s'y ramène à un dénombrement (chapitre 22).

Propriété — règles de calcul. Pour tous événements A, B : $\mathbb{P}(\emptyset) = 0$; $\mathbb{P}(\bar{A}) = 1 - \mathbb{P}(A)$; $\mathbb{P}(B \setminus A) = \mathbb{P}(B) - \mathbb{P}(A \cap B)$; **croissance** : $A \subset B \implies \mathbb{P}(A) \leq \mathbb{P}(B)$ et alors $\mathbb{P}(B \setminus A) = \mathbb{P}(B) - \mathbb{P}(A)$; enfin

$$\boxed{\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)} \quad \text{d'où} \quad \mathbb{P}(A \cup B) \leq \mathbb{P}(A) + \mathbb{P}(B).$$

Si $A_1, \dots, A_n$ sont deux à deux incompatibles, $\mathbb{P}(\bigcup_i A_i) = \sum_i \mathbb{P}(A_i)$; si c'est de plus un système complet, cette somme vaut 1.

La formule du crible est hors programme : la formule de la réunion n'est disponible que pour deux événements ; au-delà, on découpe en événements deux à deux incompatibles ou l'on passe au contraire.

3 PROBABILITÉS CONDITIONNELLES

Définition. Si $\mathbb{P}(B) > 0$, la **probabilité conditionnelle de A sachant B** est $\mathbb{P}(A \mid B) = \mathbb{P}_B(A) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$. *Convention officielle* : $\mathbb{P}(A \mid B)\mathbb{P}(B) = 0$ lorsque $\mathbb{P}(B) = 0$, de sorte que $\mathbb{P}(A \cap B) = \mathbb{P}(A \mid B)\mathbb{P}(B)$ soit vraie sans hypothèse.

Théorème. $\mathbb{P}_B$ est une *probabilité* sur Ω : toutes les règles ci-dessus lui s'appliquent ($\mathbb{P}_B(\bar{A}) = 1 - \mathbb{P}_B(A)$, etc.). En revanche $\mathbb{P}_B$ est une probabilité en la *première* variable seulement : $\mathbb{P}(A \mid B) + \mathbb{P}(A \mid \bar{B})$ n'a aucune raison de valoir 1.

Théorème — probabilités composées. Si $\mathbb{P}(A_1 \cap \dots \cap A_{n-1}) > 0$, alors

$$\mathbb{P}(A_1 \cap \dots \cap A_n) = \mathbb{P}(A_1) \mathbb{P}(A_2 \mid A_1) \mathbb{P}(A_3 \mid A_1 \cap A_2) \cdots \mathbb{P}(A_n \mid A_1 \cap \dots \cap A_{n-1}).$$

Idée de preuve. Récurrence sur n ; par croissance, tous les conditionnements écrits portent sur des événements de probabilité non nulle.

Théorème — probabilités totales. Si $(A_1, \dots, A_n)$ est un système complet d'événements, alors pour tout événement B

$$\mathbb{P}(B) = \sum_{i=1}^n \mathbb{P}(B \cap A_i) = \sum_{i=1}^n \mathbb{P}(B | A_i) \mathbb{P}(A_i)$$

(la seconde écriture avec la convention ci-dessus). *Idée de preuve (exigible).* Les $B \cap A_i$ sont deux à deux incompatibles et de réunion $B \cap \Omega = B$: on applique l'additivité puis la définition du conditionnement.

Théorème — formule de Bayes. Si $\mathbb{P}(A) > 0$ et $\mathbb{P}(B) > 0$, $\mathbb{P}(A | B) = \frac{\mathbb{P}(B | A) \mathbb{P}(A)}{\mathbb{P}(B)}$; et si $(A_1, \dots, A_n)$ est un système complet d'événements de probabilités non nulles,

$$\mathbb{P}(A_j | B) = \frac{\mathbb{P}(B | A_j) \mathbb{P}(A_j)}{\sum_{i=1}^n \mathbb{P}(B | A_i) \mathbb{P}(A_i)}.$$

Idée de preuve (exigible). Calculer $\mathbb{P}(A \cap B)$ de deux façons, puis développer le dénominateur par les probabilités totales.

Méthode — arbre de probabilités. un arbre dont le premier niveau est un système complet d'événements résume les trois formules : on *multiplie* le long d'une branche (composées), on *additionne* les branches aboutissant à un même événement (totales); Bayes remonte l'arbre.

4 LOIS DE VARIABLES ALÉATOIRES

Définition. La **loi** de $X : \Omega \rightarrow E$ est la probabilité $\mathbb{P}_X$ définie par $\mathbb{P}_X(A) = \mathbb{P}(X \in A)$; elle est déterminée par la distribution $(\mathbb{P}(X = x))_{x \in E}$, nulle hors de l'ensemble fini $X(\Omega)$. On note $X \sim Y$ lorsque $\mathbb{P}_X = \mathbb{P}_Y$. Si $f : E \rightarrow F$, alors $f(X) = f \circ X$ est une variable aléatoire et $X \sim Y \implies f(X) \sim f(Y)$. Attention, $X \sim Y$ n'entraîne pas $X = Y$: sur $\Omega = \{0, 1\}$ équiprobable, $X = \text{id}$ et $Y = 1 - X$ ont même loi et ne sont jamais égaux.

Définition — lois usuelles.

■ FORMULAIRE

- $X \sim \mathcal{U}(E)$, E fini non vide : X est à valeurs dans E et $\mathbb{P}(X = x) = \frac{1}{\text{Card } E}$ pour tout $x \in E$.
- $X \sim \mathcal{B}(p)$, $p \in [0; 1]$: $X(\Omega) \subset \{0, 1\}$, $\mathbb{P}(X = 1) = p$, $\mathbb{P}(X = 0) = 1 - p$ (X code le succès d'une expérience). Pour $A \subset \Omega$, l'indicatrice $\mathbf{1}_A$ suit $\mathcal{B}(\mathbb{P}(A))$.
- $X \sim \mathcal{B}(n, p)$ (loi **binomiale**), $n \in \mathbb{N}^*$: X est à valeurs dans $\llbracket 0; n \rrbracket$ et $\mathbb{P}(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$. Interprétation : nombre de succès lors de n répétitions *indépendantes* d'une expérience de probabilité de succès p . Contrôle : $\sum_{k=0}^n \binom{n}{k} p^k (1 - p)^{n-k} = (p + (1 - p))^n = 1$ (binôme).

Définition. Pour $\mathbb{P}(A) > 0$, la **loi conditionnelle de X sachant A** est la loi de X pour $\mathbb{P}_A$, c'est-à-dire la distribution $(\mathbb{P}(X = x | A))_{x \in E}$.

Définition. Le **couple** $(X, Y) : \Omega \rightarrow E \times F$ est une variable aléatoire; sa loi, dite **loi conjointe**, est donnée par les $\mathbb{P}(X = x, Y = y)$. Les lois de X et Y sont les **lois marginales**, et on les retrouve par

$$\mathbb{P}(X = x) = \sum_{y \in F} \mathbb{P}(X = x, Y = y), \quad \mathbb{P}(Y = y) = \sum_{x \in E} \mathbb{P}(X = x, Y = y),$$

probabilités totales appliquées au système complet $(\{Y = y\})_{y \in F}$ (resp. $(\{X = x\})_{x \in E}$). Tout s'étend aux n -uplets.

Remarque. Les marges ne déterminent pas la loi conjointe : les tableaux $\begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix}$ et $\begin{pmatrix} 1/4 & 1/4 \\ 1/4 & 1/4 \end{pmatrix}$ ont les mêmes marges, toutes deux uniformes sur $\{0, 1\}$.

5 INDÉPENDANCE

Définition. A et B sont **indépendants** lorsque $\mathbb{P}(A \cap B) = \mathbb{P}(A) \mathbb{P}(B)$; si $\mathbb{P}(B) > 0$, cela s'écrit $\mathbb{P}(A | B) = \mathbb{P}(A)$. La famille $(A_1, \dots, A_n)$ est **indépendante** lorsque $\mathbb{P}(\bigcap_{i \in I} A_i) = \prod_{i \in I} \mathbb{P}(A_i)$ pour toute partie I de $\llbracket 1; n \rrbracket$.

Propriété. Si A et B sont indépendants, A et $\bar{B}$ le sont aussi (donc aussi $\bar{A}$ et $\bar{B}$); dans une famille indépendante, on peut remplacer certains A_i par leurs contraires sans perdre l'indépendance.

Remarque — les deux confusions à ne jamais faire. **Incompatibilité et indépendance sont sans rapport** : deux événements incompatibles de probabilités non nulles ne sont *jamais* indépendants, puisque $\mathbb{P}(A \cap B) = 0 < \mathbb{P}(A) \mathbb{P}(B)$. Et **l'indépendance deux à deux n'implique pas l'indépendance** de la famille : il faut vérifier toutes les parties I , y compris $I = \llbracket 1; n \rrbracket$.

Définition. $X : \Omega \rightarrow E$ et $Y : \Omega \rightarrow F$ sont **indépendantes**, ce qu'on note $X \perp\!\!\!\perp Y$, lorsque pour toutes parties $A \subset X(\Omega)$ et $B \subset Y(\Omega)$ les événements $(X \in A)$ et $(Y \in B)$ sont indépendants.

Théorème — caractérisation. $X \perp\!\!\!\perp Y$ si et seulement si $\mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x)\mathbb{P}(Y = y)$ pour tous $(x, y) \in E \times F$: la loi conjointe est le produit des lois marginales. De même $X_1, \dots, X_n$ sont indépendantes lorsque $\mathbb{P}(X_1 = x_1, \dots, X_n = x_n) = \prod_{i=1}^n \mathbb{P}(X_i = x_i)$ pour tous $x_1, \dots, x_n$.

Notation. On modélise n expériences aléatoires indépendantes par une famille finie $(X_i)_{1 \leq i \leq n}$ de variables aléatoires indépendantes ; la construction explicite d'un univers portant ce n -uplet n'est pas un objectif du programme.

Théorème. Si $X_1, \dots, X_n$ sont indépendantes de même loi $\mathcal{B}(p)$, alors $X_1 + \dots + X_n \sim \mathcal{B}(n, p)$. *Idée de preuve (exigible).* $\{\sum X_i = k\}$ est la réunion disjointe, sur les $\binom{n}{k}$ parties I de cardinal k , des $\bigcap_{i \in I} (X_i = 1) \cap \bigcap_{i \notin I} (X_i = 0)$, chacune de probabilité $p^k(1-p)^{n-k}$ par indépendance.

Théorème — stabilité, lemme des coalitions. Si $X \perp\!\!\!\perp Y$, alors $f(X) \perp\!\!\!\perp g(Y)$ pour toutes applications f, g . Plus généralement (**lemme des coalitions**, admis) : si $X_1, \dots, X_n$ sont indépendantes et $1 \leq m < n$, alors $f(X_1, \dots, X_m) \perp\!\!\!\perp g(X_{m+1}, \dots, X_n)$; l'énoncé s'étend à plus de deux coalitions, formées de blocs d'indices *disjoints*.

6 ESPÉRANCE

Définition. Pour $X : \Omega \rightarrow \mathbb{C}$, l'**espérance** de X est $E(X) = \sum_{x \in X(\Omega)} x \mathbb{P}(X = x) = \sum_{\omega \in \Omega} X(\omega) \mathbb{P}(\{\omega\})$: c'est un **indicateur de**

position. X est dite **centrée** si $E(X) = 0$; $X - E(X)$ l'est toujours. L'espérance ne dépend que de la loi, celle d'une constante c vaut c , et $E(\mathbf{1}_A) = \mathbb{P}(A)$.

Théorème — propriétés. Pour X, Y sur Ω à valeurs complexes et $\lambda, \mu \in \mathbb{C}$: **linéarité** $E(\lambda X + \mu Y) = \lambda E(X) + \mu E(Y)$ — sans aucune hypothèse d'indépendance ; **positivité** : $X \geq 0 \implies E(X) \geq 0$; **croissance** : $X \leq Y \implies E(X) \leq E(Y)$; **inégalité triangulaire** : $|E(X)| \leq E(|X|)$. En particulier $m \leq X \leq M \implies m \leq E(X) \leq M$.

Idée de preuve (exigible). Tout se lit sur $E(X) = \sum_{\omega} X(\omega) \mathbb{P}(\{\omega\})$: linéarité de la somme, positivité des $\mathbb{P}(\{\omega\})$, inégalité triangulaire du module.

Théorème — formule de transfert. Pour $X : \Omega \rightarrow E$ et $f : E \rightarrow \mathbb{C}$, $E(f(X)) = \sum_{x \in X(\Omega)} f(x) \mathbb{P}(X = x)$: on calcule $E(f(X))$

sans déterminer la loi de $f(X)$. Elle vaut aussi pour les couples et n -uplets : $E(f(X, Y)) = \sum_{x, y} f(x, y) \mathbb{P}(X = x, Y = y)$.

Idée de preuve (exigible). Regrouper $\sum_{\omega} f(X(\omega)) \mathbb{P}(\{\omega\})$ selon la valeur de $X(\omega)$, ce qui est licite car $(\{X = x\})_{x \in X(\Omega)}$ est un système complet d'événements.

Théorème. Si $X \perp\!\!\!\perp Y$, alors $E(XY) = E(X)E(Y)$; plus généralement $E(\prod_{i=1}^n X_i) = \prod_{i=1}^n E(X_i)$ si $X_1, \dots, X_n$ sont indépendantes. La réciproque est fautive (voir « décorrélées »).

Propriété. Si $X \sim \mathcal{B}(p)$, $E(X) = p$; si $X \sim \mathcal{B}(n, p)$, $E(X) = np$; si $X \sim \mathcal{U}(\llbracket 1 ; n \rrbracket)$, $E(X) = \frac{n+1}{2}$. *Idée de preuve (exigible)* pour la binomiale. Écrire $X = X_1 + \dots + X_n$ avec $X_i \sim \mathcal{B}(p)$ indépendantes et appliquer la linéarité. (Le calcul direct par $k \binom{n}{k} = n \binom{n-1}{k-1}$ marche aussi, mais la décomposition en indicatrices est la bonne méthode.)

Méthode — exemple type : décomposer en indicatrices. pour $\sigma \sim \mathcal{U}(\mathfrak{S}_n)$, le nombre de points fixes s'écrit $X = \sum_{i=1}^n \mathbf{1}_{A_i}$ avec $A_i = \{\sigma(i) = i\}$; comme $\text{Card}\{\sigma, \sigma(i) = i\} = (n-1)!$, $\mathbb{P}(A_i) = \frac{1}{n}$ et $E(X) = 1$: un point fixe en moyenne, quel que soit n . Les A_i n'étant pas indépendants, la variance réclamera les covariances.

7 VARIANCE, ÉCART TYPE, COVARIANCE

Définition. Pour $X : \Omega \rightarrow \mathbb{R}$: **variance** $V(X) = E((X - E(X))^2) \geq 0$ et **écart type** $\sigma(X) = \sqrt{V(X)}$ — des **indicateurs de dispersion**. X est **réduite** si $V(X) = 1$.

Théorème — formule de König-Huygens. $V(X) = E(X^2) - E(X)^2$, d'où $E(X)^2 \leq E(X^2)$. *Idée de preuve (exigible).* Avec $m = E(X)$ constante, développer $(X - m)^2 = X^2 - 2mX + m^2$ et appliquer la linéarité : $V(X) = E(X^2) - 2m^2 + m^2$.

Propriété. $V(X) = 0$ si et seulement s'il existe c tel que $\mathbb{P}(X = c) = 1$. Pour $a, b \in \mathbb{R}$: $V(aX + b) = a^2 V(X)$ et $\sigma(aX + b) = |a| \sigma(X)$. Si $\sigma(X) > 0$, la variable $X^* = \frac{X - E(X)}{\sigma(X)}$ est **centrée réduite**.

Définition. La **covariance** est $\text{Cov}(X, Y) = E((X - E(X))(Y - E(Y)))$; si $\text{Cov}(X, Y) = 0$, X et Y sont dites **décorrélées**.

Théorème. $\boxed{\text{Cov}(X, Y) = E(XY) - E(X)E(Y)}$; Cov est symétrique, bilinéaire, et $\text{Cov}(X, X) = V(X)$. Si $X \perp\!\!\!\perp Y$, alors X et Y sont **décorrélées**; la réciproque est fautive. *Idée de preuve (exigible).* Développer le produit et appliquer la linéarité (comme pour König-Huygens); l'implication est le théorème $E(XY) = E(X)E(Y)$.

Théorème — variance d'une somme. $V\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n V(X_i) + 2 \sum_{1 \leq i < j \leq n} \text{Cov}(X_i, X_j)$. Si les X_i sont **deux à deux décorré-**
lées (en particulier si elles sont indépendantes), $V(\sum X_i) = \sum V(X_i)$. En particulier, pour X et Y décorréées, $V(X + Y) = V(X - Y) = V(X) + V(Y)$; prendre $Y = X$ montre que l'hypothèse est indispensable, car $V(2X) = 4V(X)$.

Propriété. Si $X \sim \mathcal{B}(p)$, $V(X) = p(1 - p)$; si $X \sim \mathcal{B}(n, p)$, $V(X) = np(1 - p)$; si $X \sim \mathcal{U}([1; n])$, $V(X) = \frac{n^2 - 1}{12}$. *Idée de preuve (exigible).* Pour $\mathcal{B}(p)$: $X^2 = X$, donc $V(X) = p - p^2$. Pour la binomiale : $X = \sum X_i$ avec les X_i **indépendantes** de loi $\mathcal{B}(p)$, donc les variances s'ajoutent.

8 INÉGALITÉS DE CONCENTRATION

Théorème — inégalité de Markov. Si X est positive, alors pour tout $a > 0$: $\boxed{\mathbb{P}(X \geq a) \leq \frac{E(X)}{a}}$. *Idée de preuve (exigible).* On a $a \mathbf{1}_{(X \geq a)} \leq X$ sur tout Ω (disjonction des cas selon que ω réalise ou non $(X \geq a)$, en utilisant $X \geq 0$); on prend l'espérance par croissance et $E(\mathbf{1}_{(X \geq a)}) = \mathbb{P}(X \geq a)$.

Théorème — inégalité de Bienaymé-Tchebychev. Pour X réelle et $\varepsilon > 0$: $\boxed{\mathbb{P}(|X - E(X)| \geq \varepsilon) \leq \frac{V(X)}{\varepsilon^2}}$. *Idée de preuve (exigible).* Appliquer Markov à la variable positive $Y = (X - E(X))^2$ et au seuil $a = \varepsilon^2 > 0$, en remarquant que $(Y \geq \varepsilon^2) = (|X - E(X)| \geq \varepsilon)$ et $E(Y) = V(X)$.

Théorème — concentration d'une moyenne. Soient $X_1, \dots, X_n$ **indépendantes de même loi**, d'espérance m et de variance σ^2 , et $\overline{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$. Alors $E(\overline{X}_n) = m$, $V(\overline{X}_n) = \frac{\sigma^2}{n}$, et

$$\forall \varepsilon > 0, \quad \mathbb{P}(|\overline{X}_n - m| \geq \varepsilon) \leq \frac{\sigma^2}{n\varepsilon^2}.$$

Idée de preuve (exigible). Linéarité pour l'espérance; $V(\overline{X}_n) = \frac{1}{n^2} \sum V(X_i) = \frac{\sigma^2}{n}$ par indépendance et $V(aX) = a^2 V(X)$; puis Bienaymé-Tchebychev appliqué à $\overline{X}_n$.

Remarque — interprétation fréquentiste. Pour $X_i \sim \mathcal{B}(p)$ indépendantes, $\overline{X}_n$ est la **fréquence** de succès et $\sigma^2 = p(1 - p) \leq \frac{1}{4}$, d'où $\mathbb{P}(|\overline{X}_n - p| \geq \varepsilon) \leq \frac{1}{4n\varepsilon^2}$: à ε fixé, le majorant est aussi petit qu'on veut pour n grand, ce qui justifie la lecture d'une probabilité comme une fréquence. La majoration est **non asymptotique** — vraie pour **chaque** n , elle fournit un n explicite. On la démontre, on ne l'invoque pas sous un nom : « la loi faible des grands nombres » n'est pas un résultat du cours de première année, et la **convergence en probabilité** est hors programme.

► SAVOIR-FAIRE

- Décrire un univers fini et une probabilité dessus; reconnaître l'équiprobabilité, se ramener à un dénombrement.
- Construire un système complet d'événements, puis appliquer composées, totales, Bayes (arbre); prouver ou réfuter l'indépendance de deux événements, d'une famille, de deux variables aléatoires.
- Déterminer une loi, une loi conjointe, des lois marginales, une loi conditionnelle sachant un événement.
- Calculer E et V par la loi, par transfert, ou par décomposition en somme d'indicatrices — presque toujours la bonne idée; calculer une covariance, en déduire la variance d'une somme; majorer une probabilité par Markov ou Bienaymé-Tchebychev et trouver un n assurant une précision donnée.

✖ PIÈGES

- Confondre $\mathbb{P}(A | B)$ et $\mathbb{P}(B | A)$; écrire $\mathbb{P}(A | B)$ sans $\mathbb{P}(B) > 0$; appliquer les probabilités totales à une famille qui n'est pas un système complet. Confondre incompatibles et indépendants.
- Écrire $E(XY) = E(X)E(Y)$ ou $V(X + Y) = V(X) + V(Y)$ sans justifier l'indépendance (resp. la décorrélation); croire que décorréées $\implies$ indépendantes.
- Confondre X et le nombre $E(X)$; écrire $E(X^2) = E(X)^2$ ou $E(1/X) = 1/E(X)$.
- Modéliser un tirage **sans** remise par une binomiale; appliquer Markov à une variable non positive, ou Bienaymé-Tchebychev sans centrer. Formule du crible, loi géométrique ou de Poisson, fonction génératrice, densité, univers infini : hors programme.