

• Espaces probabilisés

- Ensembles finis ou dénombrables ; si E et F sont finis ou dénombrables, $E \cup F$ et $E \times F$ le sont aussi.
- Espace probabilisé (univers, tribu des événements, loi de probabilité), probabilité d'une réunion (au plus dénombrable) d'événements disjoints, théorèmes des limites croissante et décroissante.
- Variable aléatoire discrète $X : \Omega \rightarrow E$, tribu engendrée par une famille finie ou dénombrable d'atomes, notamment par la famille $(X^{-1}(\{e\}))_{e \in \text{Im } X}$, loi d'une variable aléatoire (et notations $X \sim Y$ et $X \hookrightarrow \mathcal{L}$).
- Cas des fonctions caractéristiques $1_A : \Omega \rightarrow \{0, 1\}$. Utilisation en combinatoire : $1_{A \cap B} = 1_A 1_B$, $1_{\Omega \setminus A} = 1_\Omega - 1_A$.
- Loi conjointe = loi de (X, Y) , marges d'une v.a. à valeurs dans $\mathbb{R}^2$.

La loi de (X, Y) est définie par les $p_{ij} = P(X = i, Y = j)$. Les v.a. X et Y sont indépendantes ssi $p_{ij} = a_i b_j$.

• Loi conditionnelles, formule de Bayes, formule de l'intersection et formule des probabilités totales

- Loi conditionnelle : $P_B(A) = P(A | B) = \frac{P(A \cap B)}{P(B)}$.
- Probabilités composées : $P(A \cap B) = P(A | B)P(B)$.
- Formule de l'intersection : $P(A_1 \cap A_2 \cap \dots \cap A_n) = P(A_1)P(A_2 | A_1)P(A_3 | A_1 \cap A_2) \dots P(A_n | A_1 \cap \dots \cap A_{n-1})$.

Autre formulation : avec $B_1 \supset B_2 \dots \supset B_n$, on a $P(B_n) = P(B_1)P(B_2 | B_1)P(B_3 | B_2) \dots P(B_n | B_{n-1})$.

Exemple : La probabilité que $(x_1, \dots, x_p) \in \{1, 2, \dots, n\}^p$ soit injectif est le produit des $1 - \frac{k}{n}$, avec $0 \leq k \leq p - 1$.

- **Formule des probabilités totales** : Si $(B_n)_{n \in I}$ système complet de Ω , $P(A) = \sum_{n \in I} P(A | B_n)P(B_n)$.

Formule de Bayes (s'interprète souvent comme une inversion cause/effet) : $P(A | B)P(B) = P(B | A)P(A)$.

• Espérance, variance des variables aléatoires réelles

- Espérance finie : $E(|X|) < +\infty$; dans ce cas, $E(X) = \sum_{x \in X(\Omega)} x P(X = x)$ est défini et on a $|E(X)| \leq E(|X|)$.
- **Théorème du transfert** : $E(f(X)) = \sum_{e \in \text{Im } X} P(X = e)f(e)$. Linéarité de l'espérance : $E(X + Y) = E(X) + E(Y)$.
- Moments d'ordre 2 : $E(X^2) < +\infty$. Dans ce cas, $E(|X|) < +\infty$ (espérance finie).

Sev $L^2(\Omega, \mathbb{R})$ des v.a. de moment d'ordre 2 fini, muni du (pseudo)-produit scalaire $\langle X, Y \rangle = E(XY)$.

On a en particulier l'inégalité de Cauchy-Schwarz : $E(|XY|)^2 \leq E(X^2)E(Y^2)$.

Remarque : Il s'agit d'un produit scalaire si on identifie les v.a.r. égales presque partout.

- Variance $V(X) = E((X - E(X))^2) = E(X^2) - E(X)^2$. On a $V(aX + b) = a^2 V(X)$.

Covariance $\text{Cov}(X, Y) = E((X - E(X))(Y - E(Y))) = E(XY) - E(X)E(Y)$. On a $\text{Cov}(XY)^2 \leq V(X)V(Y)$.

- Variance d'une somme : $V(X_1 + \dots + X_n) = \sum_{i=1}^n V(X_i) + 2 \sum_{i < j} \text{Cov}(X_i X_j)$.

- Variables aléatoires non corrélées : $\text{Cov}(XY) = 0$: C'est notamment le cas si X et Y sont indépendantes.

Variance d'une somme de v.a. deux à deux non corrélées : $V(X_1 + \dots + X_n) = V(X_1) + \dots + V(X_n)$.

• Variables aléatoires indépendantes

- Variables indépendantes : $P(X = x, Y = y) = P(X = x)P(Y = y)$, et plus généralement avec $P(X \in E', Y \in F')$.

Variables mutuellement indépendantes : $P(X_1 = x_1, \dots, X_n = x_n) = P(X_1 = x_1)P(X_2 = x_2) \dots P(X_n = x_n)$.

- Événements indépendants : 1_A et 1_B indépendants, ce qui équivaut à $P(A \cap B) = P(A)P(B)$, i.e. $P(A | B) = P(A)$.

Événements mutuellement indépendants (à distinguer de deux à deux indépendants).

- Variables aléatoires indépendantes (mutuellement). Si X et Y sont indépendantes, alors $f(X)$ et $g(Y)$ le sont.

Si X, Y, Z sont (mutuellement) indépendantes, alors X et (Y, Z) le sont, donc $f(X)$ et $g(Y, Z)$ sont indépendants.

- Si X et Y indépendantes, alors $E(XY) = E(X)E(Y)$, donc $\text{Cov}(XY) = E(XY) - E(X)E(Y) = 0$.

- Les variables aléatoires constantes (presque partout) sont les seules variables indépendantes de toute variable.

• Inégalités de Markov et de Bienaymé-Tchebychev

- Inégalité de Markov : Si X est à valeurs positives, alors $P(X \geq \lambda) \leq \frac{E(X)}{\lambda}$.

Inégalité Bienaymé-Tchebychev : $P(|X - \mu| \geq \alpha) \leq \frac{V(X)}{\alpha^2}$. Exemple : $P(|X - \mu| \geq k\sigma(X)) \leq \frac{1}{k^2}$.

- Loi faible des grands nombres :

Si $(X_n)_{n \in \mathbb{N}}$ suite de v.a. i.i.d. de moment d'ordre 2 fini, on a $\forall \varepsilon > 0, \lim_{n \rightarrow +\infty} P(|\frac{1}{n}S_n - \mu| \geq \varepsilon) = 0$.

On dit que $(X_n)_{n \in \mathbb{N}}$ converge en probabilité vers la v.a. constante μ , où $\mu = E(X_n)$.

• Valeurs aléatoires à valeurs entières

- $E(X) = \sum_{n=0}^{+\infty} nP(X = n)$ et la série converge ssi X est d'espérance finie (par définition).

On a aussi : X est d'espérance finie ssi $\sum P(X > n)$ converge, et on a $E(X) = \sum_{n=0}^{+\infty} P(X > n)$.

En effet, avec $R_n = \sum_{k=n+1}^{+\infty} a_k$, on a : $\sum_{n=0}^N na_n = \sum_{n=1}^N n(R_{n-1} - R_n) = \sum_{n=0}^{N-1} R_n - N R_N$.

Or, $N R_N \leq \sum_{k=N+1}^{+\infty} ka_k$, donc tend vers 0 lorsque n tend vers $+\infty$. Sens réciproque aisé.

- Série génératrice $G_X(z) = E(z^X) = \sum_{n=0}^{+\infty} a_n z^n$, avec $a_n = P(X = n)$, définie (et continue) pour $|z| \leq 1$.

On a en particulier $G_X(1) = 1$, et la série entière $\sum_{n=0}^{+\infty} a_n z^n$ converge normalement sur le disque unité.

- Si X et Y sont indépendantes, $G_{X+Y}(z) = G_X(z)G_Y(z)$. Extension à la somme de n v.a. indépendantes.

- X est d'espérance finie ssi G_X (définie sur $[0, 1]$) est dérivable en 1, et $E(X) = \sum_{n=0}^{+\infty} na_n = G'_X(1)$.

- X admet un moment d'ordre 2 fini ssi G_X est (définie sur $[0, 1]$) est dérivable deux fois en 1.

Et alors $E(X^2) = \sum_{n=0}^{+\infty} n(n-1)a_n + \sum_{n=0}^{+\infty} na_n = G''_X(1) + G'_X(1)$. Donc $V(X) = G''_X(1) + G'_X(1) - G'_X(1)^2$.

Remarque : Le moment $E(X^p)$ existe ssi G_X est de classe C^p sur $[0, 1]$.

- Fonction caractéristique $\phi_X(t) = E(e^{itX}) = G_X(e^{it}) = \sum_{n=0}^{+\infty} a_n e^{int}$. On a notamment $\phi'_X(0) = iE(X)$.

Formules de Cauchy : $a_n = \frac{1}{2\pi i} \int_0^{2\pi} \phi_X(t) e^{-int} dt$. En effet, $\int_0^{2\pi} e^{ikt} e^{-int} dt = 0$ si $k \neq n$, et 2π si $k = n$.

Remarque : Théorème de Lévy : $(X_n)_{n \in \mathbb{N}}$ converge en loi vers X ssi $(\phi_{X_n})_{n \in \mathbb{N}}$ converge vers ϕ_X sur $[0, 2\pi]$.

• Lois usuelles

- Loi de Rademacher (pile ou face) $\rightarrow$ marche aléatoire (somme de n v.a. i.i.d. de loi de Rademacher).

- Loi de Bernoulli $\mathcal{B}(p)$. On a $E(X) = p$, $V(X) = p - p^2$ et $G_X(z) = q + pz$, où $q = 1 - p$.

- **Loi binomiale** $\mathcal{B}(n, p)$: somme S de n v.a. i.i.d. de loi de Bernoulli).

On a $E(S) = np$, $V(S) = n(p - p^2)$ et $G_S(z) = (q + pz)^n$, d'où on retrouve $P(S = k) = \binom{n}{k} q^{n-k} p^k$.

- **Loi géométrique** $\mathcal{G}(p)$ = loi du premier succès (à valeurs dans $\mathbb{N}^*$) ou du nombre N d'échecs avant succès ($\mathbb{N}$).

On a $P(M = k) = q^{k-1}p$ et $P(M > k) = q^k$, d'où $E(M) = \sum_{k=0}^{+\infty} q^k = \frac{1}{p}$. On a $N = M - 1$, donc $E(N) = \frac{1}{p} - 1 = \frac{q}{p}$.

On a $G_N(z) = \frac{p}{1 - qz}$, d'où on peut calculer $E(N^2)$ et $V(N)$. Remarque ; On a $M = 1 + N$, où M est le numéro du premier tirage avec succès, on a donc $P(M > k) = q^k$ et $E(M) = \frac{1}{p}$ et $G_M(z) = zG_N(z)$.

- *complément* : **Loi binomiale négative** = loi du m -ième premier succès : somme S de m v.a. géométrique.

On a $\forall k \geq m$, $P(S = k) = \binom{k-1}{m-1} q^{k-m} p^m$. Remarque : On a $G_S(z) = \left(\frac{pz}{1-qz} \right)^m$.

- **Loi de Poisson** $\mathcal{P}(\lambda)$: nombre de tops par unité de temps pour un flux moyen de λ tops par unité de temps.

On a $P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}$ et $G_X(z) = e^{\lambda(z-1)}$. Donc $E(X) = \lambda$ et $V(X) = \lambda^2 + \lambda - \lambda^2 = \lambda$.

Théorème des événements rares : Si $p_n \sim \frac{\lambda}{n}$, la loi $\mathcal{B}(n, p_n)$ converge vers $\mathcal{P}(\lambda)$. On parle de convergence en loi.