Colle semaine 11 et 12 (du 3 au 14 janvier 2023) - MP

Variables aléatoires discrètes

Cette section généralise aux variables aléatoires discrètes l'étude menée en première année des variables aléatoires à valeurs dans un ensemble fini. Cette généralisation nécessite d'introduire des notions générales de dénombrabilité et de théorie des probabilités, lesquelles font l'objet d'un exposé a minima. En particulier :

- *la notion de tribu, introduite pour donner un cadre rigoureux, n'appelle aucun développement théorique ;*
- *la construction d'espaces probabilisés n'est pas un objectif du programme.*

Ensembles dénombrables

- Un ensemble est dit dénombrable s'il est en bijection avec $\mathbb{N}$. Les parties infinies de $\mathbb{N}$ sont dénombrables.
- Un ensemble est fini ou dénombrable si et seulement s'il est en bijection avec une partie de $\mathbb{N}$. Un tel ensemble est dit au plus dénombrable.
- Un produit cartésien fini d'ensembles dénombrables est dénombrable. Une réunion finie ou dénombrable d'ensembles finis ou dénombrables est finie ou dénombrable. Les démonstrations ne sont pas exigibles.
- Les ensembles $\mathbb{N}^p$ $(p \in \mathbb{N}^*)$, $\mathbb{Z}$ et $\mathbb{Q}$ sont dénombrables. Le support d'une famille sommable de nombres complexes est dénombrable. L'ensemble $\mathbb{R}$ n'est pas dénombrable. La démonstration n'est pas exigible.

Espaces probabilisés

- Tribu sur un ensemble Ω. Espace probabilisable $(\Omega, \mathcal{A})$. La manipulation de tribus n'est pas un objectif du programme. Événements. Généralisation du vocabulaire relatif aux événements introduit en première année. Systèmes complets (ou quasi-complets) d'événements. Formule des probabilités totales (sans conditionnement).
- Probabilité sur un espace probabilisable, σ-additivité. Espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. Continuité monotone (croissante ou décroissante). Pour une suite $(A_n)_{n\in\mathbb{N}}$ d'événements (non nécessairement monotone), on a :

$$\mathbb{P}\Big(\bigcup_{k=0}^{+\infty} A_k\Big) = \lim_{n\to+\infty} \mathbb{P}\Big(\bigcup_{k=0}^{n} A_k\Big) \quad \text{et} \quad \mathbb{P}\Big(\bigcap_{k=0}^{+\infty} A_k\Big) = \lim_{n\to+\infty} \mathbb{P}\Big(\bigcap_{k=0}^{n} A_k\Big).$$

  Propriété de sous-additivité de $\mathbb{P}$ pour une réunion dénombrable d'événements (inégalité de Boole).
- Événements négligeables, événements presque sûrs. Une réunion (resp. intersection) finie ou dénombrable d'événements négligeables (resp. presque sûrs) est un événement négligeable (resp. presque sûr).
- Espaces probabilisés discrets. Si Ω est un ensemble, une distribution de probabilités discrètes sur Ω est une famille d'éléments de $\mathbb{R}^+$ indexée par Ω et de somme 1. Support d'une distribution de probabilités discrète ; le support est au plus dénombrable. Probabilité définie sur $\mathcal{A} = \mathcal{P}(\Omega)$ associée à une distribution de probabilités discrètes sur Ω. Si Ω est au plus dénombrable, on obtient ainsi toutes les probabilités sur $(\Omega, \mathcal{P}(\Omega))$.
- Une variable aléatoire discrète X définie sur l'espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ et à valeurs dans E est une application définie sur Ω, à valeurs dans l'ensemble E, telle que $X(\Omega)$ soit au plus dénombrable et que, pour tout $x \in X(\Omega)$, l'ensemble $X^{-1}(\{x\})$ appartienne à $\mathcal{A}$. Notations $(X = x)$, $(X \in A)$, $\{X = x\}$, $\{X \in A\}$. Lorsque $E = \mathbb{R}$ (ou $\mathbb{C}$), la variable aléatoire X est dite réelle (ou complexe). Notations $(X \leqslant x), (X \geqslant x), (X < x), (X > x)$ (et analogues avec accolades) pour une variable aléatoire réelle X. Variable aléatoire $f(X)$. À partir de maintenant, toutes les variables aléatoires considérées seront supposées discrètes.
- Loi $\mathbb{P}_X$ d'une variable aléatoire $X : \Omega \to E$. C'est une probabilité sur $(E, \mathcal{P}(E))$. La probabilité $\mathbb{P}_X$ est déterminée par la distribution de probabilités discrète $(\mathbb{P}(X = x))_{x\in E}$. Notation $X \sim Y$ (lorsque X et Y sont à valeurs dans un même ensemble E et ont même loi). Si $X \sim Y$ alors $f(X) \sim f(Y)$.
- Couple de variables aléatoires. Loi conjointe, lois marginales. Détermination des lois marginales à partir de la loi conjointe. Un couple est une variable aléatoire à valeurs dans un produit. Notation $\mathbb{P}(X = x, Y = y)$. Extension aux n-uplets de variables aléatoires.

- Lois usuelles. Revoir les lois de première année (Bernoulli, binomiale, uniforme).

  Pour p dans $]0,1[$, loi géométrique de paramètre p. Notations $\mathcal{G}(p)$, $X \sim \mathcal{G}(p)$. Variable géométrique de paramètre p.

  Pour λ dans $\mathbb{R}_+^*$, loi de Poisson de paramètre λ. Notations $\mathcal{P}(\lambda)$, $X \sim \mathcal{P}(\lambda)$. Variable de Poisson de paramètre λ. Interprétation en termes d'événements rares ; si X_n suit la loi binomiale $\mathcal{B}(n, p_n)$ et $p_n \underset{n\to+\infty}{\sim} \dfrac{\lambda}{n}$ avec $\lambda > 0$, alors :

$$\forall k \in \mathbb{N} \quad \mathbb{P}(X_n = k) \underset{n\to+\infty}{\longrightarrow} e^{-\lambda}\frac{\lambda^k}{k!}.$$

Conditionnement et indépendance

- Extension des résultats vus en première année : probabilité conditionnelle, formule des probabilités composées, formule des probabilités totales (avec conditionnement), formule de Bayes. Notations $\mathbb{P}_B(A), \mathbb{P}(A|B)$. Loi conditionnelle d'une variable aléatoire X sachant un événement A.
- Par définition, les événements A et B sont indépendants si $\mathbb{P}(A \cap B) = \mathbb{P}(A)\,\mathbb{P}(B)$. Lorsque $\mathbb{P}(B) > 0$, l'indépendance de A et B s'écrit $\mathbb{P}(A \mid B) = \mathbb{P}(A)$. Famille d'événements indépendants. L'indépendance deux à deux n'implique pas l'indépendance. Si A et B sont indépendants, A et $\overline{B}$ le sont aussi. Extension à une famille d'événements mutuellement indépendants (l'indépendance est conservée par passage au complémentaire).
- Couple de variables aléatoires indépendantes, famille finie de variables aléatoires indépendantes. Notation $X \perp\!\!\!\perp Y$. Les variables aléatoires X et Y sont indépendantes si et seulement si la distribution de probabilités de (X, Y) est le produit des distributions de probabilités de X et Y. Extension aux n-uplets de variables aléatoires. Famille quelconque de variables aléatoires indépendantes.
- Fonctions de variables aléatoires indépendantes : si $X \perp\!\!\!\perp Y$, alors $f(X) \perp\!\!\!\perp g(Y)$ Extension au cas de plus de deux variables.
- Lemme des coalitions : si les variables aléatoires $X_1, \ldots, X_n$ sont indépendantes, les variables aléatoires $f(X_1, \ldots, X_m)$ et $g(X_{m+1}, \ldots, X_n)$ le sont aussi. Extension au cas de plus de deux coalitions.
- Existence d'espaces probabilisés portant une suite de variables indépendantes de lois discrètes données. La démonstration est hors programme. Suites i.i.d. Modélisation du jeu de pile ou face infini : suite i.i.d. de variables de Bernoulli. Interprétation de la loi géométrique comme rang du premier succès dans le jeu de pile ou face infini.

Espérance d'une variable aléatoire réelle ou complexe. Variance et covariance dans le cas réel.

- Si X est une variable aléatoire à valeurs dans $\mathbb{R}^+ \cup \{+\infty\}$, l'espérance de X est la somme, dans $[0, +\infty]$, de la famille $(x\ \mathbb{P}(X = x))_{x \in X(\Omega)}$. Notation $\mathbb{E}(X)$. Pour une variable aléatoire à valeurs dans $\mathbb{N} \cup \{+\infty\}$, égalité $\mathbb{E}(X) = \displaystyle\sum_{n=0}^{+\infty} \mathbb{P}(X > n)$.
- Une variable aléatoire complexe X est dite d'espérance finie si la famille $(x\ \mathbb{P}(X = x))_{x \in X(\Omega)}$ est sommable ; dans ce cas, la somme de cette famille est l'espérance de X. Notation $\mathbb{E}(X)$. Variables centrées. La notation $X \in L^1$ signifie que X est d'espérance finie. On ne soulèvera aucune difficulté quant à la définition précise de L^1. Espérance d'une variable géométrique, d'une variable de Poisson.

  Formule de transfert : soit X une variable aléatoire discrète, f une fonction définie sur $X(\Omega)$ à valeurs complexes ; alors $f(X)$ est d'espérance finie si et seulement si la famille $(f(x)\ \mathbb{P}(X = x))_{x \in X(\Omega)}$ est sommable ; si tel est le cas :

$$\mathbb{E}\left(f(X)\right) = \sum_{x \in X(\Omega)} f(x)\,\mathbb{P}(X = x).$$

  Linéarité, positivité, croissance, inégalité triangulaire. Caractérisation des variables aléatoires à valeurs dans $\mathbb{R}^+$ d'espérance nulle. Si $|X| \leqslant Y$ et si $Y \in L^1$, alors $X \in L^1$. Si X et Y sont dans L^1 et indépendantes, alors XY est dans L^1 et :

$$\mathbb{E}(XY) = \mathbb{E}(X)\ \mathbb{E}(Y).$$

  Extension au cas de n variables aléatoires.
- À partir de maintenant, les variables aléatoires sont réelles. Si $\mathbb{E}\left(X^2\right) < +\infty$, X est d'espérance finie. La notation $X \in L^2$ signifie que X^2 est d'espérance finie. On ne soulèvera aucune difficulté quant à la définition précise de L^2.

- Inégalité de Cauchy-Schwarz : si X et Y sont dans L^2, XY est dans L^1 et $\mathbb{E}(XY)^2 \leqslant \mathbb{E}\left(X^2\right)\ \mathbb{E}\left(Y^2\right)$. Cas d'égalité.
- Pour $X \in L^2$, variance et écart type de X. Notations $\mathbb{V}(X), \sigma(X)$. Variables réduites. Caractérisation des variables aléatoires de variance nulle. Relation $\mathbb{V}(X) = \mathbb{E}\left(X^2\right) - \mathbb{E}(X)^2$. Relation $\mathbb{V}(aX+b) = a^2\ \mathbb{V}(X)$. Si $\sigma(X) > 0$, la variable aléatoire $\dfrac{X - \mathbb{E}(X)}{\sigma(X)}$ est centrée réduite. Variance d'une variable géométrique, d'une variable de Poisson. Covariance de deux variables aléatoires de L^2. Relation $\mathrm{Cov}(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X)\ \mathbb{E}(Y)$. Cas de variables indépendantes. Variance d'une somme de n variables aléatoires, cas de variables décorrélées.
- Fonction génératrice de la variable aléatoire X à valeurs dans $\mathbb{N}$: $G_X(t) = \mathbb{E}\left(t^X\right) = \displaystyle\sum_{k=0}^{+\infty} \mathbb{P}(X=k)\ t^k$. La série entière définissant G_X est de rayon supérieur ou égal à 1 et converge normalement sur le disque fermé de centre 0 et de rayon 1. Continuité de G_X. Détermination de la loi de X par G_X.
- La variable aléatoire X est d'espérance finie si et seulement si G_X est dérivable en 1 ; dans ce cas $\mathbb{E}(X) = {G_X}'(1)$. La démonstration de la réciproque n'est pas exigible. Utilisation de G_X pour le calcul de $\mathbb{E}(X)$ et $\mathbb{V}(X)$.

  Les étudiants doivent savoir calculer rapidement la fonction génératrice d'une variable aléatoire de Bernoulli, binomiale, géométrique, de Poisson. Fonction génératrice d'une somme finie de variables aléatoires indépendantes à valeurs dans $\mathbb{N}$.

Inégalités probabilistes et loi faible des grands nombres

- Inégalité de Markov. Inégalité de Bienaymé-Tchebychev.
- Loi faible des grands nombres : si $(X_n)_{n\geqslant 1}$ est une suite i.i.d. de variables aléatoires de variance finie, alors, pour tout $\varepsilon > 0$,

$$\mathbb{P}\left(\left|\frac{S_n}{n} - m\right| \geqslant \varepsilon\right) \underset{n\to\infty}{\longrightarrow} 0 \qquad \text{où} \qquad S_n = \sum_{k=1}^{n} X_k \quad \text{et} \quad m = \mathbb{E}(X_1).$$

Remarques

- Voici une liste de question de cours avec démonstration pour les colleurs qui souhaitent démarrer ainsi. Pour la semaine 11, on se limitera aux questions de rang impair et pour la semaine 12 aux questions de rang pair.
  - i) Énoncer et démontrer la formule de Bayes pour un système complet d'événements.
  - ii) Soit A et B deux événements indépendants. Montrer que $\overline{A}$ et B sont indépendants.
  - iii) Soit X une variable aléatoire à valeurs dans $\mathbb{N} \cup \{+\infty\}$. Montrer que $\mathbb{E}(X) = \displaystyle\sum_{n=0}^{+\infty} \mathbb{P}(X > n)$.
  - iv) Soit $(X,Y) \in (L^1)^2$. Montrer que $X + Y \in L^1$ et $\mathbb{E}(X+Y) = \mathbb{E}(X) + \mathbb{E}(Y)$.
  - v) Soit $(X,Y) \in (L^1)^2$ *indépendantes*. Montrer que $XY \in L^1$ et $\mathbb{E}(XY) = \mathbb{E}(X)\,\mathbb{E}(Y)$.
  - vi) Inégalité de Cauchy-Schwarz.
  - vii) Soit X, Y deux variables aléatoires indépendantes à valeurs dans $\mathbb{N}$. En utilisant un produit de Cauchy, montrer que :

    $$\forall t \in [0,1] \quad G_{X+Y}(t) = G_X(t) G_Y(t).$$
  - viii) Calculer sur $[0,1]$ la fonction génératrice d'une variable aléatoire X qui suit une loi de Poisson de paramètre $\lambda > 0$.
  - ix) Calculer sur $[0,1]$ la fonction génératrice d'une variable aléatoire X qui suit une loi géométrique de paramètre $p \in \,]0,1[$.
  - x) Énoncer et démontrer l'inégalité de Markov.
  - xi) Énoncer et démontrer l'inégalité de Bienaymé-Tchebychev.
  - xii) Énoncer et démontrer la loi faible des grands nombres.
- On ne posera pas d'exercice de dénombrabilité ni portant spécifiquement sur la notion de tribu.