Mise à niveau - Probabilités

1 Cadre général des probabilités

Definition. Un **espace de probabilité** est un triplet $(\Omega, \mathcal{F}, \mathbb{P})$ où
- Ω est un ensemble regroupant toutes les issues possibles du hasard, appelées **réalisations** $\omega \in \Omega$;
- $\mathcal{F}$ est un ensemble d'**événements** $A \subset \Omega$ qu'on appelle aussi **mesurables**, car ce sont des ensembles A de réalisations $\omega \in \Omega$ auquel on peut assigner une probabilité ;
- $\mathbb{P}$ est une **mesure de probabilité**, c'est-à-dire une fonction $\mathbb{P} : \mathcal{F} \to [0, 1]/$

Ce triplet doit vérifier les conditions suivantes :
- Ω est mesurable, avec $\mathbb{P}[\Omega] = 1$;
- si A est mesurable, alors son complémentaire A^c l'est aussi, avec $\mathbb{P}[A^c] = 1 - \mathbb{P}[A]$;
- si $(A_n)_{n\geq 1}$ est une suite (finie ou infinie) d'ensembles mesurables, l'union $\cup_{n\geq 1} A_n$ l'est aussi. De plus dès que les événements A_n sont **disjoints deux à deux** on a

$$\mathbb{P}\left[\bigcup_{n\geq 1} A_n\right] = \sum_{n\geq 1} \mathbb{P}\left[A_n\right]$$

En particulier, pour deux événements A et B disjoints, $\mathbb{P}[A \cup B] = \mathbb{P}[A] + \mathbb{P}[B]$.

Exercice 1. Montrer que $\emptyset$ est mesurable et que $\mathbb{P}[\emptyset] = 0$. Soit A et B deux événements, montrer que $A \cap B$ est mesurable (on pourra faire un dessin).

Exercice 2. On considère l'espace de probabilité $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), Leb)$, où $\mathcal{B}(\mathbb{R}^d)$ est l'ensemble des **boréliens** $B \subset \mathbb{R}^d$ sur lesquels on peut définir la mesure de Lebesgue Leb. On rappelle que tout ouvert de $\mathbb{R}^d$ est un borélien et que

$$Leb\Big(]a_1, b_1[\times \cdots \times]a_d, b_d[\Big) = |b_1 - a_1| \cdots |b_d - a_d|.$$

On définit la distance $\|x - y\|_\infty$ de deux points $x, y \in \mathbb{R}^d$ par $\sup_i |x_i - y_i|$.

Montrer que $A = \{x \in \mathbb{R}^d : \exists n \in \mathbb{N}^*, \|x - (n, \ldots, n)\|_\infty < 1/(n+1)\}$ est un borélien, et montrer que $Leb(A)$ est fini dès que $d \geq 2$.

Il faut s'habituer à ce que les événements ne soient pas toujours notés par des ensembles, mais par des phrases. Ces phrases renvoient (sans le dire à chaque fois), à l'ensemble A des réalisations $\omega \in \Omega$ qui correspondent à la situation décrite dans la phrase.

Exercice 3. Supposons que A et B soient deux événements disjoints tels que $\mathbb{P}[A] = 0.3$ et $\mathbb{P}[B] = 0.5$. Quelles sont les probabilités de "A ou B", "A et B, "A mais pas B" ?

Exercice 4. Soient B_1, B_2 et B_3 trois événements disjoints deux à deux, mais tels qu'au moins un des évenements soit à chaque fois réalisé (c'est-à-dire $B_1 \cup B_2 \cup B_3 = \Omega$). Montrer que pour tout événement A,

$$\mathbb{P}[A] = \mathbb{P}[A \text{ et } B_1] + \mathbb{P}[A \text{ et } B_2] + \mathbb{P}[A \text{ et } B_3].$$

2 Variables aléatoires

En probabilités, on s'intéresse à des **variables aléatoires**, notées $X, Y, ...$, dont la valeur n'est pas déterminée *a priori*. La valeur d'une variable aléatoire X dépend d'un aléa (qui peut être dû au hasard, ou à une incertitude, un manque d'information...). Nous représentons donc naturellement X comme une certaine fonction de Ω. Chaque réalisation ω fixe une certaine valeur $X(\omega)$ pour la variable aléatoire X.

Definition. Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité. Une **variable aléatoire** à valeurs dans $\mathbb{R}^d$ (en abrégé **v.a.**) est une fonction $X : \Omega \to \mathbb{R}^d$ qui est **mesurable**, c'est-à-dire que, pour tout borélien $B \in \mathcal{B}(\mathbb{R}^d)$, l'ensemble $\{w \in \Omega : X(\omega) \in B\}$ est mesurable. Cet événement est aussi noté $X \in B$.

En particulier, pour une v.a. X et un borélien B, l'événement $X \in B$ a une probabilité donnée par $\mathbb{P}[X \in B] = \mathbb{P}[\{w \in \Omega : X(\omega) \in B\}]$. Plus généralement, toute phrase faisant intervenir X fait référence à l'ensemble des réalisations $\omega \in \Omega$ pour lesquels $X(\omega)$ est dans la situation décrite par la phrase (par exemple $X \geq 0$, "X est un entier pair", ou $X = Y$).

Exercice 5. Soit X une v.a. réelle telle que $\mathbb{P}\Big[X \in]-1;1[\Big] = 0.3$. Calculer $\mathbb{P}\Big[X^2 \geq 1\Big]$.

Grâce à la notion de loi, on peut calculer les probabilités $\mathbb{P}[X \in B]$ sans revenir à l'ensemble $\{w \in \Omega : X(\omega) \in B\}$, qui est trop abstrait.

Definition. Soit X une v.a. à valeurs dans $\mathbb{R}^d$. La **loi** (ou **distribution**) de X est la mesure de probabilité P_X sur $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ définie par $P_X(B) = \mathbb{P}[X \in B]$ pour tout borélien $B \in \mathcal{B}(\mathbb{R}^d)$.

1. Lorsque X prend un nombre fini ou dénombrable de valeurs, on dit que la v.a. est **discrète**, et le support de la loi est alors fini ou dénombrable (ce qui simplifie la description de P_X).
2. Lorsqu'il existe une fonction mesurable $f : \mathbb{R}^d \to \mathbb{R}_+$ telle que

$$P_X(B) = \int_B f(x)dx,$$

on dit que la v.a. est **à densité** (la description de P_X est alors donnée par la densité f).

Connaître les lois usuelles discrètes : Bernoulli, binomiales, géométriques, Poisson ; et continues : uniforme sur un intervalle, exponentielle, gaussienne.

Exercice 6. Soit X une v.a. réelle telle que $\mathbb{P}[X < 0] = 0.3$. Identifier la loi de la v.a. $\mathbb{1}_{X \geq 0}$.

Exercice 7. Soit X une v.a. suivant la loi uniforme sur $[-1; 1]$. Calculer $\mathbb{P}[|X| > 1/2]$.

Exercice 8. Soit X une v.a. réelle à densité, mais dont la densité f est paire ($f(x) = f(-x)$). Soit $x > 0$. Montrer que

1. $\mathbb{P}[X > x] = \mathbb{P}[X < -x]$;
2. $\mathbb{P}[|X| > x] = 2\mathbb{P}[X > x]$;
3. $\mathbb{P}[|X| < x] = 2\mathbb{P}[X < x] - 1$.

3 Fonction de répartition

Pour des v.a. réelles, il n'y a pas forcément besoin de calculer $P_X(B)$ pour tous les boréliens B. En effet, une mesure de probabilité sur $\mathbb{R}$ est déterminée de manière unique par sa valeur sur les boréliens $]-\infty; t]$, ce qui motive la définition suivante.

Definition. La **fonction de répartition** F_X d'une v.a. X réelle est l'application réelle

$$F_X : t \mapsto \mathbb{P}[X \leq t].$$

La loi P_X est en fait complétement déterminé par F_X, car $P_X(]-\infty; t])$ est donné par $F_X(t)$ pour tout $t \in \mathbb{R}$.

Exercice 9. Montrer que si F_X est une fonction dérivable, alors X est à densité. Inversement, montrer que si la v.a. a une densité f continue sauf en un nombre fini de points $t_1, \ldots, t_n$, alors F_X est continue, et dérivable en tout $t \neq t_1, \ldots, t_n$ avec $F'_X(t) = f(t)$.

Exercice 10. Montrer que si F_X et F_Y sont les fonctions de répartition de X et Y, alors $\inf(X, Y)$ a pour fonction de répartition $F_{\inf(X,Y)} = 1 - (1 - F_X)(1 - F_Y)$.

4 Espérance de variables aléatoires

Definition. Soit X une v.a. à valeurs dans $\mathbb{R}^d$ et de loi P_X. Pour toute fonction mesurable bornée $g : \mathbb{R}^d \to \mathbb{R}$, l'espérance de $g(X)$ est donnée par

$$\mathbb{E}[g(X)] = \int_{\mathbb{R}^d} g(x) dP_X(x).$$

En particulier, en intégrant la variable aléatoire $\mathbb{1}_B$, on obtient

$$\mathbb{E}[\mathbb{1}_B] = \mathbb{P}[B].$$

Plus généralement, on peut définir $\mathbb{E}[g(X)]$ dès que $\int_{\mathbb{R}^d} |g| dP_X < +\infty$.

Exercice 11. Calculer $\mathbb{E}[X^2]$ pour une v.a. réelle de loi $P_X = \frac{1}{3}\delta_0 + \frac{1}{3}\delta_1 + \frac{1}{3}\delta_2$

Exercice 12. Soit X une v.a. réelle et $x > 0$. Montrer que $\mathbb{P}[|X| > x] = \mathbb{E}[\mathbb{1}_{[-x;x]^c}(X)]$. En remarquant que $\mathbb{1}_{[-x;x]^c}(X) \leq |X|/x$, en déduire l'inégalité de Markov suivante :

$$\mathbb{P}[|X| > x] \leq \frac{\mathbb{E}[|X|]}{x}.$$

L'espérance est une intégrale comme une autre, et on a le théorème de Fubini(-Tonelli).

Exercice 13. Résoudre les questions suivantes en appliquant le théorème de Fubini.

1. Soit N une v.a. à valeurs dans $\mathbb{N}$. Montrer que
$$\mathbb{E}[N] = \sum_{n \geq 1} \mathbb{P}[N \geq 1].$$
2. Soit X une v.a. à valeurs dans $\mathbb{R}_+$ et $\alpha > 0$. Montrer que
$$\mathbb{E}[X^\alpha] = \alpha \int_0^{+\infty} t^{\alpha-1} \mathbb{P}[X > t] dt,$$
et écrire une formule plus générale pour $\mathbb{E}[g(X)]$ avec g monotone et C^1 sur $]0; +\infty[$.
3. Soit $(A_n)_{n \geq 1}$ une suite d'événements. On note N le nombre (aléatoire) d'événements parmi ceux-ci qui se produisent. Calculer $\mathbb{E}[N]$ en fonction des probabilités $\mathbb{P}[A_n]$.

Pour trouver la loi de X, il suffit de trouver une mesure de probabilité P telle que pour toute fonction mesurable bornée $g : \mathbb{R}^d \to \mathbb{R}$,

$$\mathbb{E}[g(X)] = \int_{\mathbb{R}^d} g(x) dP(x).$$

En effet, la loi de X est alors donnée par P car le cas particulier $g = \mathbb{1}_B(X)$ nous donne $P_X(B) = P(B)$.

Exercice 14. Soit X une v.a. réelle de densité $f : \mathbb{R} \to \mathbb{R}_+$. Calculer la loi de $2X + 2$, puis la loi de X^2.

Exercice 15.
1. Soit X une v.a. réelle de loi exponentielle $\mathcal{E}(1)$. Déterminer la loi de $\lfloor X \rfloor + 1$.
2. Soit U une v.a. de loi uniforme sur $[-1; 1]$. Déterminer la loi de $\arcsin(U)$.
3. Soit X une v.a. de loi gaussienne centrée réduite. Déterminer la loi de $|X|$.

5 Indépendance de variables aléatoires

Definition. Les v.a. $X_1, \dots, X_n$ définies sur le même espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$ sont indépendantes si et seulement si l'une des affirmations équivalentes suivantes est vérifiée :

1. Pour tous boréliens $B_1, \dots, B_n$, on a
$$\mathbb{P}[X_1 \in B_1 \text{ et } X_2 \in B_2 \dots \text{ et } X_n \in B_n] = \mathbb{P}[X_1 \in B_1] \times \cdots \times \mathbb{P}[X_n \in B_n];$$
2. pour toutes fonctions mesurables bornées $f_1, \dots, f_n$, on a
$$\mathbb{E}[f_1(X_1) \cdots f_n(X_n)] = \mathbb{E}[f(X_1)] \times \cdots \times \mathbb{E}[f(X_n)];$$
3. la loi du n-tuple $(X_1, \dots, X_n)$ est donnée par le produit des lois $P_{(X_1,\dots,X_n)} = P_{X_1} \otimes \cdots \otimes P_{X_n}$, ce qui implique que
$$\mathbb{E}[f(X_1, \dots, X_n)] = \int f(x_1, \dots, x_n) dP_{X_1}(x_1) \cdots dP_{X_n}(x_n)$$
pour toute fonction mesurable bornée.

Exercice 16. En utilisant le théorème de Fubini, montrer que si X et Y sont indépendantes, alors toute fonction mesurable bornée F de ces variables vérifie
$$\mathbb{E}[F(X,Y)] = \mathbb{E}[F_1(X)] = \mathbb{E}[F_2(Y)]$$
pour $F_1(x) := \mathbb{E}[F(x,Y)]$ et $F_2(y) := \mathbb{E}[F(X,y)]$.

Exercice 17. On s'intéresse à la somme de deux variables indépendantes X et Y à valeurs dans $\mathbb{R}$.

1. Montrer que si X et Y sont variables de densité f_X et f_Y, alors la v.a. $X+Y$ a pour densité la fonction $f_X \star f_Y$ suivante
$$f_X \star f_Y(z) = \int_{\mathbb{R}} f_X(z-y) f_Y(y) dy.$$
2. On suppose que X et Y sont deux variables exponentielles de paramètre $\mathcal{E}(1)$. Calculer la loi de $X+Y$.

Exercice 18. Soit (X,Y) une v.a. à valeurs dans $\mathbb{R}^2$ de densité $p(x,y) = e^{-y} \mathbb{1}_{0<x<y}$.
— Déterminer la loi de X et celle de $Y - X$.
— Montrer que X et $Y - X$ sont indépendantes.

Exercice 19. Soit N, X_1, X_2 trois v.a. réelles indépendantes. On suppose que N est à valeurs dans $\{1,2\}$, et que X_1 et X_2 sont à densité. Montrer que la v.a. X_N est à densité également (X_N est égale à la v.a. X_1 dès que $N = 1$ et est égale à la v.a. X_2 dès que $N = 2$).

Exercice 20. On suppose que X et Y sont deux binomiales indépendantes de paramètres respectifs $\mathcal{B}(n,p)$ et $\mathcal{B}(m,p)$. Calculer la loi de $X+Y$. On pourra utiliser l'identité de Vandermonde
$$\sum_{i=0}^{n+m} \binom{n}{i} \binom{m}{k-i} = \binom{n+m}{k}.$$