Intégration & probabilités
MACS 1
Introduction de la partie probabilités

Laurent Tournier

Novembre 2018

Notion d'expérience aléatoire

(Ce qui suit n'est pas une définition mathématique)

Dans la vie courante, une expérience est dite **aléatoire** si son déroulement est imprévisible, à chaque fois qu'on la répète dans les "mêmes conditions".

Remarques :

- Les "conditions" changent selon l'**information** dont on dispose sur l'expérience. Le "hasard" vient d'ailleurs souvent d'un manque d'information. Il est donc essentiel, pour décrire une expérience aléatoire, de préciser exactement ce qui est connu ou non.
- "Répéter l'expérience" signifie que l'on "régénère" la situation, de telle sorte que le résultat de la première expérience ne donne *aucune information* sur celui de la deuxième : les deux expériences sont **indépendantes**, mais se déroulent selon le même protocole.

Bien qu'imprévisible, une suite de répétitions indépendantes d'une même expérience présente (le plus souvent) des "régularités". Notamment, on peut observer la **loi des grands nombres** :

Considérons une propriété A qui dépend de l'expérience. Si on répète l'expérience n fois, où n est "grand", alors la proportion des fois où la propriété A est réalisée s'approche d'un certain nombre réel $P(A) \in [0, 1]$:

$$\frac{\text{nombre de fois où } A \text{ est réalisé parmi } n \text{ expériences}}{n} \simeq P(A).$$

Ce nombre $P(A)$ correspond donc à la "proportion de chance" que A soit réalisé lors d'une expérience donnée.
Dans les cas simples, $P(A)$ peut être deviné à partir d'arguments de symétrie.

Fondement mathématique : notion d'espace de probabilité

Essayons de formaliser les propriétés de $P(A) =$ "proportion limite de fois où A est réalisé" :

- si A est une propriété toujours fausse, $P(A) = 0$;
- si A est une propriété toujours vraie, $P(A) = 1$;
- si A et B sont des propriétés incompatibles (jamais simultanément vraies), $P(A \text{ ou } B) = P(A) + P(B)$.

Choisissons une représentation mathématique Ω de l'ensemble des résultats de l'expérience. Une propriété A s'identifie à un sous-ensemble de Ω (l'ensemble des résultats qui vérifient la propriété).
Ainsi P est une fonction $P : \mathcal{P}(\Omega) \to [0, 1]$ telle que $P(\emptyset) = 0$, $P(\Omega) = 1$ et si A et B sont des parties disjointes de Ω, $P(A \uplus B) = P(A) + P(B)$.

Pour obtenir une théorie mathématique satisfaisante, calquée sur la théorie de l'intégration,

- On fait l'hypothèse plus forte que cette propriété s'étend à une infinité dénombrable de parties disjointes, si bien que P est une **mesure** et
- On suppose que l'on veut seulement définir la probabilité de certains ensembles (qui forment nécessairement une tribu). On les appellera des **événements**.

Un **espace de probabilité** est un espace mesuré $(\Omega, \mathcal{A}, P)$ tel que $P(\Omega) = 1$. Autrement dit, $\mathcal{A}$ est une tribu sur Ω, appelée tribu des événements, et $P : \mathcal{A} \to [0, 1]$ vérifie :

$$\text{pour toute suite } (A_n)_n \text{ d'événements disjoints,} \qquad P\Big(\biguplus_{n\in\mathbb{N}} A_n\Big) = \sum_{n\in\mathbb{N}} P(A_n).$$

Soit $(\Omega, \mathcal{A}, P)$ un espace de probabilité.

- un élément $\omega \in \Omega$ est une **réalisation** de l'expérience : chaque élément de Ω correspond à une façon dont l'expérience peut se dérouler
- Ω est l'**espace des éventualités** ou l'**univers**
- un ensemble $A \in \mathcal{A}$ (donc $A \subset \Omega$) est un **événement**. Si $\omega \in A$, on dit que ω **réalise** A.
- $P(A)$ est la **probabilité** de l'événement A
- Si $P(A) = 1$ (donc $P(A^c) = 0$) on dit que A est un événement **presque sûr**, ou que A est réalisé **presque sûrement**, en abrégé **p.s.**.
- Une **variable aléatoire** est une fonction mesurable $X : \Omega \to E$ (où $(E, \mathcal{E})$ est un espace mesurable). Elle associe une valeur $X(\omega)$ à chaque réalisation $\omega \in \Omega$. On parle de **variable aléatoire réelle** si $X : \Omega \to \mathbb{R}$.

Les événements relatifs à une variable aléatoire seront généralement notés de façon fonctionnelle, c'est-à-dire : pour $B \in \mathcal{E}$, on notera

$$\{X \in B\} = \{\omega \in \Omega \,|\, X(\omega) \in B\} = X^{-1}(B) \qquad \in \mathcal{A}.$$

- La **loi** de la variable aléatoire X est la mesure image de P par X. C'est donc la probabilité P_X sur $(E, \mathcal{E})$ donnée par :

  $$\text{pour tout } B \in \mathcal{E}, \qquad P_X(B) = P(X \in B) = P(\{\omega \in \Omega \,|\, X(\omega) \in B\}).$$

  On dira que X **suit la loi** P_X, et on notera parfois $X \sim P_X$.
- L'**espérance** de $X : \Omega \to \mathbb{R}$ est son intégrale par rapport à P : on note $E[X] = \int_\Omega X \, dP$ (si existe : $X \geq 0$ ou $E[|X|] < \infty$).

Théorème de transfert

Rappel :

- Une **variable aléatoire** est une fonction mesurable $X : \Omega \to E$
- La **loi** de la variable aléatoire X est la mesure image de P par X. C'est donc la probabilité P_X sur $(E, \mathcal{E})$ donnée par :

$$\text{pour tout } B \in \mathcal{E}, \qquad P_X(B) = P(X \in B) = P(\{\omega \in \Omega \,|\, X(\omega) \in B\}).$$

- L'**espérance** de $X : \Omega \to \mathbb{R}$ est son intégrale par rapport à P : on note $E[X] = \int_\Omega X \, dP$ (si existe : $X \geq 0$ ou $E[|X|] < \infty$).

Par le théorème de transfert, pour toute variable aléatoire $X : \Omega \to E$, et toute fonction mesurable $\varphi : E \to \mathbb{R}$,

$$E[\varphi(X)] = \int_{\mathbb{R}} \varphi(x) \, dP_X(x),$$

dès lors que $\varphi \geq 0$ ou $\varphi(X)$ est intégrable.

$\rightsquigarrow$ Pour calculer $E[\varphi(X)]$, il suffit de connaître la loi P_X de X

Remarques sur les opérations ensemblistes

Un événement est un ensemble (une partie de Ω). Néanmoins, il est généralement décrit par une condition vérifiée ou non par l'expérience. Les opérations de logique entre conditions correspondent à des opérations d'ensembles entre événements, à garder à l'esprit :

- $A^c = \Omega \setminus A$ est l'événement "A n'est pas réalisé", c'est-à-dire le contraire de A
- $A \cup B$ est l'événement "A ou B est réalisé"
- $A \cap B$ est l'événément "A et B sont réalisés"
- $A \cap B = \emptyset$ signifie que "A et B sont incompatibles"
- $A \subset B$ signifie que "A implique B"
- $\bigcup\limits_{1 \leq k \leq n} A_k$ est l'événement "il existe $1 \leq k \leq n$ tel que A_k est réalisé"
- $\bigcap\limits_{1 \leq k \leq n} A_k$ est l'événement "pour tous $1 \leq k \leq n$, A_k est réalisé"
- $B \cap A^c$ est l'événement "B est réalisé, mais pas A"

Usage de la fonction indicatrice

Si A est un événement, $\mathbf{1}_A : \Omega \to \{0, 1\}$ est une variable aléatoire (très pratique).

Elle vaut 1 si A est réalisé, et 0 sinon : $\mathbf{1}_A(\omega) = \begin{cases} 1 & \text{si } \omega \in A, \\ 0 & \text{sinon.} \end{cases}$

Probabilité conditionnelle, indépendance

On a signalé que disposer d'une information modifie l'expérience. Plus précisément, cela modifie la probabilité :

Soit $B \in \mathcal{A}$ tel que $P(B) > 0$. La **probabilité conditionnelle sachant** B est la probabilité $P(\cdot \mid B)$ sur $(\Omega, \mathcal{A})$ donnée par

$$\text{pour tout } A \in \mathcal{A}, \qquad P(A \mid B) = \frac{P(A \cap B)}{P(B)}.$$

$P(A \mid B)$ correspond à la probabilité de A lorsque l'on modifie l'expérience par l'information que B est réalisé.

On a évoqué l'indépendance dans la façon de répéter une expérience. C'est une notion clé en probabilité, qui signifie qu'une certaine information n'a pas d'incidence.

- Deux événements A et B sont **indépendants** si $P(A \cap B) = P(A)P(B)$;
- Deux variables aléatoires $X : \Omega \to E$ et $Y : \Omega \to F$ sont **indépendantes** si

$$\text{pour tous } A \in \mathcal{E}, B \in \mathcal{F}, \qquad \{X \in A\} \text{ et } \{Y \in B\} \text{ sont indépendants.}$$

Ceci revient à $P(X \in A, Y \in B) = P(X \in A)P(Y \in B)$, ce qui s'écrit aussi $P_{(X,Y)}(A \times B) = P_X(A)P_Y(B)$, où $P_{(X,Y)}$ est la loi de (X, Y) (variable aléatoire $\Omega \to E \times F$).

$$X \text{ et } Y \text{ sont indépendantes ssi } P_{(X,Y)} = P_X \otimes P_Y.$$

Dans ce cas, on pourra "intégrer par tranches" et appliquer les théorèmes de Fubini.

Une variable aléatoire $X : \Omega \to E$ est **discrète** si E est dénombrable, ou plus généralement si $P(X \in D) = 1$, où $D \subset E$ est dénombrable.

Pour simplifier, supposons juste que E est dénombrable.
Dans ce cas, la loi de X est

$$P_X = \sum_{x \in E} P(X = x)\delta_x$$

En effet, si $A \subset E$, A est dénombrable donc

$$P_X(A) = P(X \in A) = P(\biguplus_{x \in A} \{X = x\}) = \sum_{x \in A} P(X = x) = \sum_{x \in E} P(X = x)\delta_x(A).$$

En particulier, par le théorème de transfert, pour toute fonction $\varphi : E \to \mathbb{R}$,

$$E[\varphi(X)] = \sum_{x \in E} P(X = x)\varphi(x),$$

à condition que $\varphi \geq 0$, ou que $\varphi(X)$ soit intégrable (c.-à-d. $\sum_{x \in E} P(X = x)|\varphi(x)| < \infty$).

Pour une variable aléatoire X à valeurs dans un ensemble E dénombrable, la loi de X est connue dès que l'on connaît $P(X = x)$ pour tout $x \in E$.

Une variable aléatoire $X : \Omega \to \mathbb{R}^d$ est **à densité** si sa loi P_X a une densité par rapport à la mesure de Lebesgue λ_d, autrement dit s'il existe une fonction $f_X : \mathbb{R}^d \to \mathbb{R}_+$ mesurable telle que

$$\text{pour tout } A \in \mathcal{B}(\mathbb{R}^d), \qquad P_X(A) = P(X \in A) = \int_A f_X(x)dx.$$

En particulier, par le théorème de transfert, pour toute fonction $\varphi : E \to \mathbb{R}$,

$$E[\varphi(X)] = \int \varphi(x) f_X(x)dx,$$

à condition que $\varphi \geq 0$, ou que $\varphi(X)$ soit intégrable (c.-à-d. $\int |\varphi(x)| f_X(x)dx < \infty$).