X2016 – MAP 311 – PC 3

RAPPELS : Vecteurs de variables aléatoires

Igor Kortchemski – igor.kortchemski@cmap.polytechnique.fr

1 Méthode de la fonction muette (Section 4.10 du poly)

Il s'agit d'une méthode très efficace pour déterminer la loi d'une variable aléatoire à valeurs dans *n'importe quel espace* (mais dans MAP 311 on se limitera à des variables aléatoires à valeurs dans $\mathbb{R}^n$). Il s'agit en quelque sorte de la réciproque du théorème de transfert.

Plus précisément, soit $X = (X_1, \dots, X_n)$ une variable aléatoire à valeurs dans $\mathbb{R}^n$ (sa loi $\mathbb{P}_X$ est donc une probabilité sur $\mathbb{R}^n$). Alors, d'après le théorème de transfert, pour toute fonction $h : \mathbb{R}^n \to \mathbb{R}$ continue bornée, on a

$$\mathbb{E}\left[h(X_1, \dots, X_n)\right] = \int_{\mathbb{R}^n} h(x_1, \dots, x_n)\mathbb{P}_X(dx_1, \dots, dx_n), \tag{1}$$

où $\mathbb{P}_X = \mathbb{P}_{(X_1,\dots,X_n)}$ est la loi du vecteur X (qui est donc une mesure de probabilité sur $\mathbb{R}^n$, par rapport à laquelle on fait l'intégrale). La méthode de la fonction muette consiste à utiliser le fait que, réciproquement, s'il existe une probabilité μ sur $\mathbb{R}$ telle que pour toute fonction $h : \mathbb{R}^n \to \mathbb{R}$ continue bornée on

$$\mathbb{E}\left[h(X_1, \dots, X_n)\right] = \int_{\mathbb{R}^n} h(x_1, \dots, x_n)\mu(dx_1, \dots, dx_n),$$

alors la loi de X est μ.

Cas particulier (variable aléatoire réelle à densité) : Soit Y une variable aléatoire réelle et f une fonction positive. On suppose que pour toute fonction continue bornée $h : \mathbb{R} \to \mathbb{R}$ on a

$$\mathbb{E}\left[h(Y)\right] = \int_{\mathbb{R}} h(x)f(x)dx.$$

Alors Y est une variable aléatoire réelle à densité et $f(x)$ est une densité de Y au point x.

Cas particulier (v.a. à densité dans $\mathbb{R}^n$) : Soit $X = (X_1, \dots, X_n)$ une variable aléatoire à valeurs dans $\mathbb{R}^n$ et $f : \mathbb{R}^n \to \mathbb{R}$ une fonction positive. On suppose que pour toute fonction continue bornée $h : \mathbb{R}^n \to \mathbb{R}$ on a

$$\mathbb{E}\left[h(X)\right] = \int_{\mathbb{R}^n} h(x_1, \dots, x_n)f(x_1, \dots, x_n)dx_1 \cdots dx_n.$$

Alors $X = (X_1, \dots, X_n)$ est à densité, et $f(x_1, \dots, x_n)$ est une densité de X au point $(x_1, \dots, x_n)$.

Attention. En pratique, on utilise presque **toujours** la méthode de la fonction muette pour déterminer des lois de variables aléatoires qui ne sont pas discrètes (sauf typiquement dans le cas de variables aléatoires réelles définies avec des minimums ou des maximums, etc.)

2 Théorèmes de Fubini (Section 4.8 du poly)

Lorsqu'on calcule des intégrales multiples, on a souvent recourt aux théorèmes de Fubini qui permettent (sous certaines hypothèses !) d'intervertir les intégrales. Informellement :

Pour des questions, demande de précisions ou explications, n'hésitez pas à m'envoyer un mail.

Théorème de Fubini (signe positif) : lorsqu'on intègre une fonction **positive** à plusieurs variables, on peut toujours intervertir les intégrales.

Théorème de Fubini (signe quelconque) : lorsqu'on intègre une fonction de signe à plusieurs variables, on peut intervertir les intégrales si l'intégrale est **absolument convergente**.

Attention. Il est très important de se souvenir qu'une espérance est une intégrale (voir par exemple (1)) : on peut donc (lorsque c'est licite) intervertir espérance et intégrales !

Remarque : en toute généralité, les théorèmes de Fubini sont en fait valables pour des intégrales par rapport à des mesures σ-finies.

3 Indépendance de variables aléatoires (Section 4.9 du poly)

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé. Si $X_1 : (\Omega, \mathcal{A}) \to (E_1, \mathcal{E}_1), \ldots, X_n : (\Omega, \mathcal{A}) \to (E_n, \mathcal{E}_n)$ sont des variables aléatoires, on dit qu'elles sont (mutuellement) indépendantes (sous $\mathbb{P}$) si

$$\text{pour tous } B_1 \in \mathcal{E}_1, \ldots, B_n \in \mathcal{E}_n, \qquad \mathbb{P}\left(X_1 \in B_1, \ldots, X_n \in B_n\right) = \prod_{i=1}^{n} \mathbb{P}\left(X_i \in B_i\right).$$

Si I est un ensemble quelconque, on dit que des variables aléatoires $(X_i)_{i \in I}$ sont indépendantes si pour tout $J \subset I$, avec $\mathrm{Card}(J) < \infty$, les variables aléatoires $(X_j)_{j \in J}$ sont indépendantes.

Attention. En pratique, dans MAP 311, on prendra les espaces d'arrivée E_i de la forme $\mathbb{R}$ (variable aléatoire réelle) ou de la forme $\mathbb{R}^n$ (vecteurs de variables aléatoires réelles).

On utilise très souvent les principes suivants, qui construisent d'autres familles de variables aléatoires indépendantes :

Principe de composition. Soient $X_1 : (\Omega, \mathcal{A}) \to (E_1, \mathcal{E}_1), \ldots, X_n : (\Omega, \mathcal{A}) \to (E_n, \mathcal{E}_n)$ des variables aléatoires indépendantes. Si $(f_i : E_i \to E'_i)_{1 \leq i \leq n}$ sont des fonctions mesurables, alors les variables aléatoires $f_1(X_1), \ldots, f_n(X_n)$ sont indépendantes.

(Par exemple, si X, Y, Z sont des variables aléatoires réelles indépendantes, alors $X^2, 1/Y, e^{-Z}$ sont indépendantes).

Principe des coalitions. Soient $X_1 : (\Omega, \mathcal{A}) \to (E_1, \mathcal{E}_1), \ldots, X_n : (\Omega, \mathcal{A}) \to (E_n, \mathcal{E}_n)$ des variables aléatoires indépendantes. Soient $1 = i_1 < i_2 < \cdots < i_{p-1} < n$ des nombres entiers. On pose

$$Y_1 = (X_{i_1}, \ldots, X_{i_2 - 1}), Y_2 = (X_{i_2}, \ldots, X_{i_3 - 1}), \ldots, Y_{p-1} = (X_{i_{p-1}}, \ldots, X_n).$$

Alors $Y_1, \ldots, Y_{p-1}$ sont indépendantes.

(Par exemple, si X, Y, Z sont des variables aléatoires réelles indépendantes, alors (X, Y) et Z sont indépendantes).

Attention. En pratique, on combine très souvent les deux principes. Par exemple si X, Y, Z sont des variables aléatoires réelles indépendantes, alors $X^2 Y$ et Z sont indépendantes. En effet, (X, Y) et Z sont indépendantes (principe des coalitions), puis $X^2 Y$ et Z sont indépendantes (principe de composition appliqué à (X, Y) avec la fonction $f(x, y) = x^2 y$ et à Z avec la fonction identité).

Cas des variables aléatoires réelles à densité.

Soient $X_1, \ldots, X_n$ des variables aléatoires réelles à densité. On note f_i une densité de X_i. Les variables aléatoires $X_1, \ldots, X_n$ sont indépendantes si et seulement une densité de $(X_1, \ldots, X_n)$ au point $(x_1, \ldots, x_n)$ est (presque partout) $f_1(x_1) \cdots f_n(x_n)$.

Version fonctionnelle de l'indépendance. Dans les calculs, on utilise très souvent le résultat suivant.

Soient $X_1, \ldots, X_n$ des variables aléatoires **réelles** indépendantes. Alors

* **(Cas positif)** On suppose que $\mathbb{P}(X_i \geq 0) = 1$ pour tout $1 \leq i \leq n$. Alors

$$\mathbb{E}\left[\prod_{i=1}^{n} X_i\right] = \prod_{i=1}^{n} \mathbb{E}[X_i] \ \in \ \mathbb{R}_+ \cup \{+\infty\}.$$

* **(Cas de signe quelconque)** Si $X_1, \ldots, X_n$ sont intégrables, alors $X_1 X_2 \cdots X_n$ est intégrable. Dans ce cas, on a alors $\mathbb{E}[\prod_{i=1}^{n} X_i] = \prod_{i=1}^{n} \mathbb{E}[X_i]$.

Informellement, l'espérance transforme les produits (de variables aléatoires réelles indépendantes) en produit d'espérances.

4 Propriétés générales de l'espérance (Section 4.3 du poly)

On utilise très souvent les propriétés suivantes de l'espérance pour des variables aléatoires réelles.

Cas positif. Soient X et Y deux variables aléatoires réelles telles que $\mathbb{P}(X \geq 0) = \mathbb{P}(Y \geq 0) = 1$.

* (**Linéarité**) On a $\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y] \in \mathbb{R}_+ \cup \{+\infty\}$.
* (**Positivité**) On a $\mathbb{E}[X] \geq 0$. De plus $\mathbb{E}[X] = 0$ si et seulement si $\mathbb{P}(X = 0) = 1$.
* (**Croissance**) Si $\mathbb{P}(X \leq Y) = 1$, alors $\mathbb{E}[X] \leq \mathbb{E}[Y]$.

Ainsi, si X est une variable aléatoire de signe quelconque et Y une variable aléatoire positive intégrable telle que $\mathbb{P}(|X| \leq Y) = 1$ (c'est-à-dire $|X(\omega)| \leq Y(\omega)$ presque sûrement), alors X est intégrable.

En particulier, si X est une variable aléatoire réelle bornée (c'est-à-dire qu'il existe $M > 0$ tel que $\mathbb{P}(|X| \leq M) = 1$), alors X est intégrable.

Cas de signe quelconque. Soient X et Y deux variables aléatoires réelles.

* (**Linéarité**) Si X et Y sont intégrables, alors $X + Y$ est intégrable et on a $\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$.
* (**Croissance**) Si $\mathbb{P}(X \leq Y) = 1$ et si X, Y sont intégrables, alors $\mathbb{E}[X] \leq \mathbb{E}[Y]$.
* (**Inégalité triangulaire**) Si X est intégrable, alors $|\mathbb{E}[X]| \leq \mathbb{E}[|X|]$.

À partir de ces propriétés, on démontre d'autres inégalités très utiles (cf Section 4.7 du poly) comme :

— l'inégalité de **Markov** : si $X \in L^p$ avec $p \geq 1$ (c'est-à-dire $\mathbb{E}[|X|^p] < \infty$), alors pour $a > 0$ on a

$$\mathbb{P}(|X| \geq a) \leq \frac{\mathbb{E}[|X|^p]}{a^p}.$$

— l'inégalité de **Bienaymé-Tchebyshev** : si $X \in L^2$, alors pour tout $a > 0$ on a

$$\mathbb{P}(|X - \mathbb{E}[X]| \geq a) \leq \frac{\mathrm{Var}(X)}{a^2}.$$

— l'inégalité de **Cauchy-Schwarz** : si $X, Y \in L^2$, alors $XY \in L^1$ et

$$|\mathbb{E}[XY] \leq \mathbb{E}[|XY|] \leq \sqrt{\mathbb{E}[X^2]\,\mathbb{E}[Y^2]}.$$

— l'inégalité de **Jensen** : si X est une variable aléatoire réelle telle que $X \in L^1$ et f est une fonction convexe telle que $f(X) \in L^1$, alors

$$f(\mathbb{E}[X]) \leq \mathbb{E}[f(X)].$$

5 Utilisation des fonctions indicatrices

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé. Si $A \in \mathcal{A}$, on note

$$\mathbb{1}_A$$

la variable aléatoire définie par

$$\mathbb{1}_A(\omega) = \begin{cases} 1 & \text{si } \omega \in A \\ 0 & \text{si } \omega \notin A. \end{cases}$$

Calcul d'une probabilité avec le théorème de transfert.

On a $\mathbb{P}(A) = \mathbb{E}[\mathbb{1}_A]$. En particulier, si X est une variable aléatoire à valeurs dans $\mathbb{R}^n$ et B est un borélien de $\mathbb{R}^n$, on a

$$\mathbb{P}(X \in B) = \mathbb{E}[\mathbb{1}_{X \in B}].$$

Cas particulier. Si $X = (X_1, \ldots, X_n)$ a une densité $f(x_1, \ldots, x_n)$ au point $(x_1, \ldots, x_n)$, on peut calculer $\mathbb{P}(X \in B)$ avec le théorème de transfert par une intégrale multiple :

$$\mathbb{P}(X \in B) = \mathbb{E}[\mathbb{1}_{X \in B}] = \int_{\mathbb{R}^n} \mathbb{1}_{(x_1, \ldots, x_n) \in B} f(x_1, \ldots, x_n) dx_1 \cdots dx_n.$$

Petites astuces. Soit X une variable aléatoire réelle.

* Si $X \geq 0$ (c'est-à-dire $X(\omega) \geq 0$ pour tout $\omega \in \Omega$), on a $X\mathbb{1}_A \leq X$ (en particulier $\mathbb{E}[X\mathbb{1}_A] \leq \mathbb{E}[X]$).
* Si $A_1, \ldots, A_n$ sont des événements disjoints qui forment une partition de Ω, on a

$$1 = \mathbb{1}_{A_1} + \cdots + \mathbb{1}_{A_n}$$

(c'est-à-dire que pour tout $\omega \in \Omega$ on a $1 = \mathbb{1}_{A_1}(\omega) + \cdots + \mathbb{1}_{A_n}(\omega)$). En particulier, si X est intégrable, on a

$$\mathbb{E}[X] = \sum_{i=1}^{n} \mathbb{E}[X\mathbb{1}_{A_i}],$$

ce qui permet de calculer des espérances en faisant des disjonctions de cas.