24 - DENSITÉS D'UN VECTEUR ALÉATOIRE

(AJOUT DU § 6)

www.daniel-saada.eu

Un vecteur aléatoire est une variable aléatoire à valeurs dans un espace $\mathbb{R}^n$, définie sur un triplet $(\Omega, \mathcal{T}, p)$.
On désignera par $\mathcal{B}^n$ la tribu des boréliens de $\mathbb{R}^n$, par $\pi_i : (x_1,...,x_n) \to x_i$ les projections de $\mathbb{R}^n$ sur $\mathbb{R}$, par $< t.x >= \sum_1^n t_i x_i$ le produit scalaire dans $\mathbb{R}^n$, par $\| t \|$ la norme de t, ou $| t |$ s'il n'y a pas d'ambigüité.
Une variable aléatoire à valeurs dans $\mathbb{R}^n$ est une application mesurable X de $(\Omega, \mathcal{T}, p)$ dans $(\mathbb{R}^n, \mathcal{B}^n)$: X est donc une application de Ω dans $\mathbb{R}^n$ telle que $X^{-1}(B) \in \mathcal{T}$ pour tout borélien B de $\mathbb{R}^n$; il suffit que B soit un pavé borélien $B_1 \times ... \times B_n$ car ces pavés engendrent $\mathcal{B}^n$.

Si $X = (X_1,...,X_n)$ est un vecteur aléatoire, chacune des coordonnées X_i est mesurable car $X_i = \pi_i \circ X$.
Réciproquement, si $X_1,...,X_n$ sont n variables aléatoires réelles, $(X_1,...,X_n)$ est un vecteur aléatoire de $\mathbb{R}^n$ car si $B = B_1 \times ... \times B_n$ est le produit de n boréliens, $X^{-1}(B)$ est l'intersection des boréliens $X_i^{-1}(B_i)$.

1) Les notions de base

a) Loi de probabilité de $X = (X_1,...,X_n)$

On pose $p_X(B) = p\left(X^{-1}(B)\right)$ quand $B \in \mathcal{B}^n$: p_X est une probabilité sur $(\mathbb{R}^n, \mathcal{B}^n)$, on dit que c'est la loi de X.
La loi de chaque X_i est alors $p_{X_i}(B) = p\left(X_i^{-1}(B)\right) = p_X\left(\pi_i^{-1}(B)\right)$ puisque $X_i^{-1} = X^{-1} \circ \pi_i^{-1}$.
Si $Z = (X,Y)$, avec X et Y indépendantes : $p_Z = p_X \otimes p_Y$ (mon livre, chapitre 13).

b) Densité f_X de $X = (X_1,...,X_n)$

On dit que $X = (X_1,...,X_n)$ a une densité s'il existe f intégrable et positive sur $\mathbb{R}^n$ telle que

$$p_X(B) = p(X \in B) = \int_B f d\lambda = \int_{\mathbb{R}^n} 1_B(u) f(u) du$$

pour tout borélien B de $\mathbb{R}^n$, λ étant la mesure de Borel-Lebesgue sur $\mathbb{R}^n$.

Une densité f est donc une fonction mesurable positive de $(\mathbb{R}^n, \mathcal{B}^n)$ dans $(\mathbb{R}, \mathcal{B}^1)$ et d'intégrale 1.

Si un couple (X,Y) a une densité $f_{X,Y}$, X et Y ont des densités (elles sont donc diffuses) données par

$$f_X(x) = \int_{\mathbb{R}} f_{X,Y}(x,y) dy \text{ et } f_Y(y) = \int_{\mathbb{R}} f_{X,Y}(x,y) dx .$$

Quand X et Y sont indépendantes, $f_{X,Y}(x,y) = f_X(x) f_Y(y)$.

Un exemple de couple sans densité. $X = (\cos U, \sin U)$ n'a pas de densité : en effet, le cercle C d'équation $x^2 + y = 1$ est de mesure (surface) nulle, donc on devrait avoir $p(X \in C) = 0$, alors que $p(X \in C) = 1$.

c) Fonction de répartition d'un vecteur du plan

La fonction de répartition F d'un couple (X,Y) est définie par $F(x,y) = p(X \le x \text{ et } Y \le y)$; F est croissante en chacune de ses variables et $F(x,y) = p_{X,Y}\left(]-\infty, x] \times]-\infty, y]\right)$. Quand (X,Y) a une densité f,

$$F(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y}f(s,t)dsdt$$

ce qui prouve que $f(x,y)=\dfrac{\partial^2 F}{\partial x\partial y}(x,y)$ quand (x,y) est un point de continuité de f.

La fonction de répartition F_X de X s'obtient par $F_X(x)=p(X\le x)=\lim\limits_{y\to+\infty}F(x,y)=F(x,+\infty)$; idem avec F_Y.

Le lecteur généralisera sans mal à la fonction de répartition d'un vecteur aléatoire $(X_1,...,X_n)$.

Exemple : fonction de répartition du couple $X=(\cos U,\sin U)$ *quand* U *est uniforme sur* $[0,2\pi[$.

$F(x,y)=p(\cos u\le x \text{ et } \sin u\le y)$ vaut évidemment 0 quand $(x$ *ou* y$)\le -1$ et 1 quand $(x$ *et* y$)\ge 1$;

$F(x,y)=\dfrac{1}{2}+\dfrac{1}{\pi}\arcsin y$ quand $x\ge 1$ et $-1\le y\le 1$; $F(x,y)=1-\dfrac{1}{\pi}\arccos x$ quand $y\ge 1$ et $-1\le x\le 1$.

On a donc élucidé F à l'extérieur de $]-1,1[^2$. A l'intérieur de ce carré, $F(x,y)$ vaut, sauf erreur :

	$y\in[-1,0]$	$y\in[0,1]$
$-1\le x\le -\sqrt{1-y^2}$	0	$\dfrac{\pi-arc\cos x}{\pi}$
$\lvert x\rvert\le\sqrt{1-y^2}$	$\dfrac{\pi/2+\arcsin x+\arcsin y}{2\pi}$	$\dfrac{2\pi+\arccos x-\arcsin y}{2\pi}$
$\sqrt{1-y^2}\le x\le 1$	$\dfrac{2\pi+2\arcsin y}{2\pi}$	$\dfrac{\pi-2\arccos x+2\arcsin y}{2\pi}$

Presque partout en (x,y), F est continue et donc $\dfrac{\partial^2 F}{\partial x\partial y}=0$: on retrouve que $X=(\cos U,\sin U)$ est sans densité.

2) Si un vecteur X a une densité, il en est de même de sa norme

a) Cas $n=2$

En passant aux coordonnées polaires et en utilisant le théorème de Fubini :

$$p\left(X\in B(0,R)\right)=\int_B f\,d\lambda=p(\|X\|\le R)=\int_0^R\int_0^{2\pi}rf(r\cos\theta,r\sin\theta)drd\theta.$$

et donc $\|X\|$ a pour densité $R\to R\int_0^{2\pi}f(R\cos\theta,R\sin\theta)d\theta$ (presque partout).

b) Cas $n=3$

En passant aux coordonnées sphériques et en utilisant le théorème de Fubini :

$$P(Y\le r)=\int_0^r\rho^2d\rho\int_{[0,2\pi]^2}f(\rho\cos\theta\cos\varphi,\rho\sin\theta\cos\varphi,\rho\sin\varphi\,|\cos\varphi|)d\theta d\varphi.$$

On dérive par rapport à r : la densité de $\|X\|$ est, presque partout :

$$r^2\int_{[0,2\pi]^2}f(r\cos\theta\cos\varphi,r\sin\theta\cos\varphi,r\sin\varphi\,|\cos\varphi|\,d\theta d\varphi.$$

c) Cas général

Un effort théorique est nécessaire.

E est un espace euclidien réel de dimension n, λ la mesure de Lebesgue sur E : λ est l'unique mesure $\sigma-finie$ sur la tribu des boréliens de E invariante par translation et valant 1 sur $[0,1]^n$ [1]. On désigne part B la boule unité et S sa sphère unité.

Si h est une homothétie de rapport k : $\lambda\big(h(A)\big)=|k|^n\ \lambda(A)$ pour tout borélien A de E (commencer avec les pavés et finir par les boréliens).

Par conséquent, si A est un borélien de S et si A_r désigne, pour $r>0$, l'ensemble des $t.x$ pour x dans A et t dans $[0,r]$, $\lambda(A_r)=r^n\lambda(A_1)$ (A_r est un homothétique de A_1). On pose alors $\sigma(A)=\lambda(A_1)$: σ est une mesure sur S, de masse $\lambda(B)=\pi^{n/2}\int_0^{+\infty}t^{n/2}e^{-t}dt$.

Pour toute fonction h intégrable sur E, l'intégrale de h sur E est égale à l'intégrale double sur $R^+\times S$ (muni de $dt.d\sigma$) de la fonction qui à (t,x) associe $h(tx)nt^{n-1}$: $\int_E h\,d\lambda=n\int_{R^+\times S}h(tx)t^{n-1}dt\,d\sigma(x)$.

Soit φ l'application de $E-(0)$ sur $]0,+\infty[\times S$ définie par $\varphi(x)=\big(\|x\|,x/\|x\|\big)$: φ est bijective et continue, φ^{-1} étant $(r,s)\mapsto rs$. D'après le Théorème du transfert : $\int_E(h\circ X)d\lambda=\int_F h\,d(\lambda\circ X^{-1})$, sous réserve de convergence et X étant mesurable de E dans F.

Si h est intégrable sur E, alors $\int_E hd\lambda=\int_E(h\circ\varphi^{-1})\circ\varphi\,d\lambda=\int_{]0,+\infty[\times S}(h\circ\varphi^{-1})\,d(\lambda\circ\varphi^{-1})$.

Il ne reste plus qu'à vérifier que la mesure $\lambda\circ\varphi^{-1}$ est la mesure produit $\rho\otimes\sigma$, où ρ est la mesure sur $]0,+\infty[$ de densité nt^{n-1} sur R^+. Pour ce faire, il suffit d'établir $\lambda\circ\varphi^{-1}=\rho\otimes\sigma$ sur la famille des $I\times A$, où I est un intervalle $]s,r]$ de $]0,+\infty[$ et A un borélien de S, qui engendre les boréliens de $]0,+\infty[\times S$.
Si $s=0$, $I=]0,r]$ et alors

$$\lambda\circ\varphi^{-1}\big(]0,r]\times A\big)=r^n\lambda\circ\varphi^{-1}\big(]0,1]\times A\big)=r^n\sigma(A)=\rho\big(]0,r]\big)\sigma(A).$$

Si $s>0$, $I=]0,r]-]0,s]$ et on opère par additivité.

Soit X une variable aléatoire à valeurs dans R^n euclidien et $Y=\|X\|$ sa norme : si X a une densité f, $p(Y\le r)=p\big(X\in B(0,r)\big)=\int_{B'(0,r)}f$. Une démonstration analogue prouverait

$$\int_{B'(0,r)}f\,d\lambda=n\int_{]0,r]\times S}f(tx)t^{n-1}dt\,d\sigma(x)=n\int_0^r t^{n-1}\Big(\int_S f(tx)d\sigma(x)\Big)dt.$$

En dérivant par rapport à r, une densité de Y est $r\mapsto nr^{n-1}\int_S f(rx)d\sigma(x)$, presque partout.

3) Le théorème du transfert pour un vecteur aléatoire

Si u est une fonction mesurable (*a fortiori* continue) de $(\mathbb{R}^n,\mathcal{B}^n)$ dans $(\mathbb{R},\mathcal{B}^1)$, alors

$u\circ X$ est une variable aléatoire réelle et $\int_\Omega(u\circ X)dp=\int_{\mathbb{R}^n}u\,dp_X$

pourvu que l'un des deux membres soit intégrable (en valeur absolue).

$\int_\Omega(u\circ X)dp$ se note $E[u\circ X]$ ou $E[u(X_1,X_2)]$: c'est l'espérance de la variable aléatoire réelle $u\circ X$.

[1] http://math.univ-lyon1.fr/~villani/Cours/PDFFILES/chap4.pdf

Quand X a une densité f, $\int_{\mathbb{R}^n} u\,dp_X = \int_{\mathbb{R}^n} u(x)f(x)dx$; en particulier, $E(e^{i<t.X>}) = \int_{\mathbb{R}^n} e^{i<t.x>} f(x)dx$.

Si de plus $n=2$, $\int_{\mathbb{R}^2} u\,dp_X = \int_{\mathbb{R}^2} u(x,y)f(x,y)\,dxdy$ et $E[\| X \|]$ s'obtient avec $u(x,y)=\sqrt{x^2+y^2}$:

$$E[\| X \|] = \int_{\mathbb{R}^2} \sqrt{x^2+y^2}\, f(x,y)\,dxdy.$$

Rappelons qu'on a également $E[\| X \|] = \int_0^{+\infty} \left(1-F_{\|X\|}(x)\right)dx$, $F_{\|X\|}$ étant la fonction de répartition de $\| X \|$.

Application. Soit (X,Y) de $(\Omega,\mathcal{T},p)$ dans $\mathbb{R}^2$ et (X',Y') de $(\Omega',\mathcal{T}',p')$ dans $\mathbb{R}^2$; on suppose X et X' de même loi, Y et Y' de même loi, X et Y indépendantes, X' et Y' indépendantes. Alors,

$$\int_\Omega \varphi(X,Y)dp = \int_{\Omega'} \varphi(X',Y')dp' \text{ quand } \varphi \text{ est intégrable de } \mathbb{R}^2 \text{ dans } \mathbb{R}.$$

En effet, les couples (X,Y) et (X',Y') ont même loi car, par exemple, $p_{X,Y} = p_X \otimes p_Y$; ensuite, $\varphi(X,Y)$ et $\varphi(X',Y')$ ont même loi également (revenir à la définition). Le théorème du transfert donne alors l'égalité cherchée.
Il en résulte que, quand U et V sont uniformes et indépendantes, à valeurs dans $[0,2\pi[$, on est en droit d'écrire

$$\int_\Omega \varphi(\cos U+\cos V, \sin U+\sin V)dp = \int_{[0,2\pi]^2} \varphi(\cos u+\cos v, \sin u+\sin v)\frac{dudv}{4\pi^2}$$

en choisissant $\Omega=[0,2\pi]^2$, $U(u,v)=u$, $V(u,v)=v$, $p=\lambda_2/4\pi^2$: U et V sont bien uniformes et indépendantes pour la probabilité p et $\varphi(\cos U+\cos V,\sin U+\sin V)$ est du type $\varphi(U,V)$.

4) La transformation de Fourier pour les fonctions numériques définies sur $\mathbb{R}^n$

Dans $\mathbb{R}^n$, le produit scalaire des vecteurs t et x est noté $<t.x>$: $<t.x> = \sum_1^n t_i x_i$; la norme $\| x \|$ de x sera abrégée en $|x|$ quand aucune confusion ne sera à craindre.

Si f est intégrable sur $\mathbb{R}^n$, et à valeurs réelles, sa transformée de Fourier, définie de $\mathbb{R}^n$ dans $\mathbb{R}$, est

$$\hat{f}(t) = \int_{\mathbb{R}^n} e^{i\varepsilon<t.x>} f(x)dx \text{ existe sur } \mathbb{R}^n \quad (\varepsilon = \pm 1).$$

Si $\hat{f}$ est à son tour intégrable, on a *presque partout* la formule d'inversion $f(x) = \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} e^{-i\varepsilon<t.x>}\hat{f}(t)dt$.

On peut donner une forme symétrique à cette formule d'inversion quand f et $\hat{f}$ sont intégrables :

$$\text{Si } \hat{f}(t) = \int_{\mathbb{R}^n} e^{2i\pi\varepsilon<t.x>} f(x)dx, \text{ alors } f(x) = \int_{\mathbb{R}^n} e^{-2i\pi\varepsilon<t.x>}\hat{f}(t)dt \quad (\varepsilon=\pm 1).$$

Quand une variable aléatoire X à valeurs dans $\mathbb{R}^n$ a une densité f, $E(e^{i<t.X>}) = \int_{\mathbb{R}^n} e^{i<t.x>} f(x)dx$;

si $t \to E(e^{i<t.X>})$ est intégrable sur $\mathbb{R}^n$, alors $\frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} e^{-i<t.x>}E(e^{i<t.X>})dt$ est une densité de X.

Pour un exemple d'emploi voir mon article sur les promenades aléatoires :

http://www.daniel-saada.eu/fichiers/21-promenades_aleatoires_planes.pdf

5) Formule des probabilités totales

On veut définir $p(Y \le a / X = b)$: par exemple limite quand $h \to 0^+$ de $p(Y \le a / b \le X \le b+h)$ et aboutir, quand le couple (X,Y) a une densité f, à la *formule des probabilités totales dans le cas continu* :

$$p(Y \le a) = \int_{-\infty}^{+\infty} p(Y \le a / X = x).f_X(x)dx$$

(on remplace dans la formule discrète $p(X = x)$ par $f_X(x)dx$ et on intègre au lieu de sommer).

Par hypothèse, $p(Y \le a \text{ et } \mathrm{b} \le \mathrm{X} \le b+h) = \int_{x=b}^{b+h}\int_{-\infty}^{a} f(x,y)dxdy$; par définition,

$$p(Y \le a / b \le X \le b+h) = \frac{p(Y \le a \text{ et } \mathrm{b} \le X \le b+h)}{p(b \le X \le b+h)} = \frac{\int_{v=-\infty}^{a}\int_{u=b}^{b+h} f(u,v)dudv}{\int_{b}^{b+h} f_X(u)du},$$ à condition que le dénominateur reste non nul dans un voisinage de b. Il en est ainsi quand $f_X(b) > 0$: en effet, $\frac{1}{h}\int_{b}^{b+h} f_X(u)du$ tend vers $f_X(b)$. Comme $\frac{1}{h}\int_{b}^{b+h}\left(\int_{-\infty}^{a} f(u,v)dv\right)du \to \int_{-\infty}^{a} f(b,v)dv$, $p(Y \le a / \mathrm{b} \le \mathrm{X} \le b+h)$ a pour limite $\frac{\int_{-\infty}^{a} f(b,v)dv}{f_X(b)}$ si $f_X(b) > 0$. Montrons maintenant que $p(Y \le a / X = b).f_X(x) = \int_{-\infty}^{a} f(b,v)dv$ pour *tout* réel b.

Puisque $\frac{p(Y \le a \text{ et } \mathrm{b} \le X \le b+h)}{p(b \le X \le b+h)} \le 1$ et $h > 0$, $\frac{1}{h}p(Y \le a \text{ et } b \le X \le b+h) \le \frac{1}{h}p(b \le X \le b+h)$ et leurs limites sont dans le même ordre : $\int_{-\infty}^{a} f(b,v)dv \le f_X(b)$. Il en résulte que $f_X(b) = 0$ implique $\int_{-\infty}^{a} f(b,v)dv = 0$: $p(Y \le a / X = b).f_X(x) = \int_{-\infty}^{a} f(b,v)dv$ est donc vraie partout ; reste à prouver que

$$p(Y \le a) = \int_{-\infty}^{+\infty}\left(\int_{-\infty}^{a} f(x,v)dv\right)dx,$$

or $\int_{-\infty}^{+\infty}\left(\int_{-\infty}^{a} f(x,v)dv\right)dx = \int_{-\infty}^{a} dv\int_{-\infty}^{\infty} f(x,v)dv = \int_{-\infty}^{a} f_Y(v)dv = p(Y \le a)$, la preuve est complète.

6) Une formule pour $E[X / Y]$ quand Y est positive

Nous allons prouver que $E[X / Y] = \int_{0}^{+\infty} E[X.e^{-tY}]dt$.

En effet, comme $E[X / Y] = \int_{x=-\infty}^{+\infty}\int_{y=0}^{+\infty} xe^{-ty} f(x,y)dxdy$:

$$\int_{0}^{+\infty} E[X.e^{-tY}]dt = \int_{t=0}^{+\infty}\left(\int_{x=-\infty}^{+\infty}\int_{y=0}^{+\infty} xe^{-ty} f(x,y)dxdy\right)dt$$

Or, $\int_{0}^{+\infty} e^{-ty}dt = 1/y$, d'où $\int_{0}^{+\infty} E[X.e^{-tY}]dt = \int_{x=-\infty}^{+\infty}\int_{y=0}^{+\infty} \frac{x}{y} f(x,y)dxdy = E[X / Y]$.

(Source : **Gérard Letac**, http://www.les-mathematiques.net/phorum/read.php?12,836884)

****