23 - DENSITÉS D'UNE VARIABLE ALÉATOIRE RÉELLE (*mise à jour*)

www.daniel-saada.eu

Si X est une variable aléatoire réelle sur un triplet $(\Omega,\mathcal{T},p)$, la loi de probabilité de X, notée p_X, est la probabilité définie sur la tribu $\mathcal{B}(\mathbb{R})$ des boréliens B de $\mathbb{R}$ par $p_X(B)=p\left(X^{-1}(B)\right)$. La fonction de répartition de X, notée F_X, est la fonction de $\mathbb{R}$ dans $\mathbb{R}$ définie par $F_X(x)=p(X\leq x)=p_X\left(]-\infty,x]\right)$. La variable aléatoire X est dite *diffuse* quand $p(X=x)=0$ pour tout réel x : cette condition équivaut à F_X continue sur $\mathbb{R}$, F_X étant alors uniformément continue sur $\mathbb{R}$.

Si p est une probabilité sur $\mathcal{B}(\mathbb{R})$, la fonction de répartition de p, notée F_p, est définie par $F_p(x)=p\left(]-\infty,x]\right)$. Une probabilité p sur $\mathcal{B}(\mathbb{R})$ est dite diffuse si $p\{x\}=0$ pour tout réel x ; X diffuse équivaut à p_X diffuse.

Une *fonction de répartition continue* est une fonction F de $\mathbb{R}$ dans $\mathbb{R}$ croissante et continue, telle que $\lim\limits_{-\infty}F=0$ et $\lim\limits_{+\infty}F=1$; pour une telle fonction, il existe une probabilité p et une seule, diffuse, sur les boréliens de $\mathbb{R}$ telle que $p(a,b)=F(b)-F(a)$, (a,b) désignant l'un des quatre intervalles d'extrémités a et b. On la note p_F.

Si f est une fonction positive et intégrable sur $\mathbb{R}$ telle que $\int_{\mathbb{R}}f=1$, l'application $B\to\int_B f=\int 1_B.f$ est une probabilité diffuse sur les boréliens de $\mathbb{R}$, noté p_f ; λ étant la mesure de Lebesgue sur $\mathbb{R}$ et B un borélien, $\lambda(B)=0$ implique $p_f(B)=0$.

1) Mesures absolument continues devant une autre

Si p et q sont deux probabilités sur $(\Omega,\mathcal{T})$, les trois assertions qui suivent sont équivalentes :

(i) $T\in\mathcal{T}$ et $p(T)=0$ impliquent $q(T)=0$;

(ii) $\forall\varepsilon>0,\exists\eta>0$ tel que $p(T)\leq\eta$ implique $q(T)\leq\varepsilon$;

(iii) Il existe f de Ω dans $\mathbb{R}$, positive et intégrable pour p, telle que $q(T)=\int_T f\,dp$ pour tout T de $\mathcal{T}$.

On dit que q est absolument continue devant p et on écrit $q<<p$.

Remarque. Si Ω est un espace $\mathbb{R}^n$ et si $\mathcal{T}=\mathcal{B}(\mathbb{R}^n)$, il suffit que T soit ouvert dans l'implication **(ii)**.

2) Fonctions de répartition continues absolument continues

Une fonction de répartition continue F est dite absolument continue quand $p_F<<\lambda$. D'après ce qui précède, F est absolument continue si et seulement si il existe f positive et $\lambda-$intégrable sur $\mathbb{R}$ telle que

$$F(b)-F(a)=\int_{[a,b]}f\,d\lambda=\int_a^b f(x)dx\,.$$

Tout ouvert de $\mathbb{R}$ est une réunion au plus dénombrable d'intervalles ouverts disjoints ; il en résulte que F est absolument continue si et seulement si pour tout $\varepsilon>0$, il existe $\eta>0$ tel que pour toute famille finie ou dénombrable

d'intervalles ouverts disjoints $]a_n,b_n[$, $\sum_n (b_n - a_n) \le \eta$ implique $\sum_n F(b_n) - F(a_n) \le \varepsilon$.
C'est le cas si F est lipchitzienne sur $\mathbb{R}$.

La formule $F(x) = \int_{-\infty}^{x} f(t)dt$ montre que F est déjà dérivable en chaque point de continuité x de f et qu'alors $F'(x) = f(x)$, mais cette condition suffisante n'est pas nécessaire. On démontre en effet que pour *presque tout réel* x, $F'(x)$ existe et vaut $f(x)$.

3) Variables aléatoires à densité

Une variable aléatoire réelle X, de fonction de répartition F, est dite à densité s'il existe une fonction f positive et intégrable sur $\mathbb{R}$ telle que $\int_a^b f = F(b) - F(a)$ pour tous réels, finis ou infinis, a et b. On dit que f est *une* densité de X car f peut être modifiée sur un ensemble de mesure nulle. X est alors diffuse et F absolument continue ; on a dit plus haut que, pour *presque tout réel* x, $f(x) = F'(x)$: une densité de X est donc F' et alors $\int_{\mathbb{R}} F' = 1$.

La fonction positive $f = F'$ est la limite, presque partout en x, de la suite des fonctions continues

$$f_n : x \to \frac{F(x+1/n) - F(x)}{1/n}.$$

Sur chaque $[a,b]$, les intégrales des f_n forment une suite convergente. En effet, par linéarité et changement de variable, $\int_a^b f_n(x)dx = n\left(\int_b^{b+1/n} F(x)dx - \int_a^{a+1/n} F(x)dx\right)$, dont la limite est $F(b) - F(a)$ parce que F est continue.

En vertu du lemme de Fatou, F' est intégrable sur $[a,b]$ et $\int_a^b F' \le F(b) - F(a)$. On en déduit $\int_{-\infty}^a F' \le F(a)$ et $\int_b^{+\infty} F' \le 1 - F(b)$. Si donc $\int_{\mathbb{R}} F' = 1$, c'est que $\int_a^b f = F(b) - F(a)$ pour tous a et b, finis ou infinis.

Une condition nécessaire et suffisante pour que X admette une densité est que la dérivée de sa fonction de répartition (qui existe presque partout et est intégrable) soit d'intégrale 1 sur $\mathbb{R}$ (elle est d'intégrale ≤ 1).

Il existe des variables aléatoires diffuses et sans densité. On peut en effet construire des fonctions croissantes, continues, non constantes, dont la dérivée est nulle presque partout : quand $F(b) \ne F(a)$, $\int_a^b F' = 0$ et non $F(b) - F(a)$ (escalier de Cantor).

4) La propriété fondamentale des variables à densité

Cette propriété s'obtient par deux extensions successives de la définition.

a) *Densité et loi*

$\int_a^b f = F(b) - F(a)$ s'interprète ainsi : $p(a \le X \le b) = \int_a^b f$; plus généralement, pour tout intervalle I de $\mathbb{R}$, $p(X \in I) = \int_I f$, autrement dit, $p(X \in I) = p\left(X^{-1}(I)\right) = p_X(I) = \int_I f$. Prouvons alors que

$$p_X(B) = \int_B f(x)dx \text{ pour tout borélien } B \text{ de } \mathbb{R}.$$

En effet, les probabilités p_X et $B \to \int_B f$ sont deux probabilités sur la tribu borélienne de $\mathbb{R}$. Comme elles coïncident sur tous les intervalles I, elles sont égales (donner référence) sur $\mathcal{B}(\mathbb{R})$.

Cette égalité sera prise comme définition d'une densité quand X est à valeurs dans $\mathbb{R}^n$.

On en déduit que $\lambda(B)=0$ implique $p_X(B)=p\left(X^{-1}(B)\right)=0$. C'est ce qui explique qu'une variable aléatoire discrète X ne peut avoir de densité f : $X(\Omega)$, qui est au plus dénombrable, est un borélien B de mesure nulle, donc $\int_B f=0$ alors que $p_X(B)=p\left(X^{-1}(B)\right)=p(\Omega)=1$.

b) *Densité et espérances*

$p_X(B)=\int_B f$ s'écrit aussi $\int_{\mathbb{R}} 1_B dp_X=\int_B f=\int_{\mathbb{R}} 1_B.f=\int_{\mathbb{R}} 1_B(x)f(x)dx$. Montrons alors que pour toute fonction h intégrable pour p_X, $h.f$ est intégrable sur $\mathbb{R}$ pour la mesure de Lebesgue et $\int_R hdp_X=\int_R h(x)f(x)\,dx$.

i) *Les deux assertions sont vraies quand h est étagée sur la tribu des boréliens de $\mathbb{R}$.*

Soit $h=\sum_1^n a_i.1_{B_i}$ avec $\mathbb{R}=\sum_1^n B_i$, chaque B_i étant un borélien. La fonction $h.f$ valant $a_i f$ sur B_i, il en résulte que $|h.f|$ est majorée sur $\mathbb{R}$ par $\max\limits_{1\le i\le n}|a_i|.f$ et donc que $h.f$ est intégrable sur $\mathbb{R}$ comme f. Deuxièmement, comme $\int_R hdp_X=\sum_1^n a_i.p_X(B_i)$ et $\int_R h(x)f(x)\,dx=\sum_1^n\int_{B_i} h(x)f(x)\,dx=\sum_1^n a_i\int_{B_i} f(x)\,dx$, l'égalité des intégrales provient de la formule vue en **a)** : $p_X(B)=\int_B f$ pour tout borélien B.

ii) *Les deux assertions sont vraies quand h est positive*.

Il existe une suite croissante de fonctions *étagées* positives h_n de limite h et $\int_R hdp_X=\lim\limits_n\int_R h_n dp_X$.

La suite des fonctions *intégrables* $h_n f$ tend en croissant vers hf parce que f est positive ; d'autre part, la suite des intégrales $\int_R h_n(x)f(x)\,dx$ est majorée par $\int_R hdp_X$ car $\int_R h_n(x)f(x)\,dx=\int_R h_n dp_X$. Le Théorème de Beppo-Levi (**11.8, p 185**) assure alors que $\int_R h(x)f(x)\,dx$ existe et que $\int_R h(x)f(x)\,dx=\lim\limits_n\int_R h_n(x)f(x)\,dx$. Comme $\int_R h_n(x)f(x)\,dx=\int_R h_n\,dp_X$ d'après **i)**, on obtient bien $\int_R h(x)f(x)\,dx=\int_R hdp_X$.

iii) *Enfin, si h n'est pas de signe constant*, la décomposition $h=h^+-h^-$ et la linéarité de l'intégrale suffisent pour conclure.

Comme le théorème du transfert (Chapitre 14) affirme que $E(h\circ X)=\int_\Omega (h\circ X)dp=\int_R hdp_X$, on peut énoncer :

> Si une variable aléatoire réelle X a une densité f et si h est une fonction de $\mathbb{R}$ dans $\mathbb{R}$ intégrable pour p_X :
>
> $$E(h\circ X)=\int_\Omega (h\circ X)dp=\int_R hdp_X=\int_R h(x)f(x)\,dx.$$

En particulier, pour $h=id_{\mathbb{R}}$, $E(X)=\int_{\mathbb{R}} x.f(x)dx$, à condition bien sûr que $\int_{\mathbb{R}}|x|.f(x)dx<+\infty$.

Exemple. La formule fondamentale s'étend, en séparant les parties réelle et imaginaire, aux fonctions h de $\mathbb{R}$ dans $\mathbb{C}$ qui sont p_X - intégrables. Soit X uniforme sur $[0,2\pi[$; vérifions que pour les fonctions continues

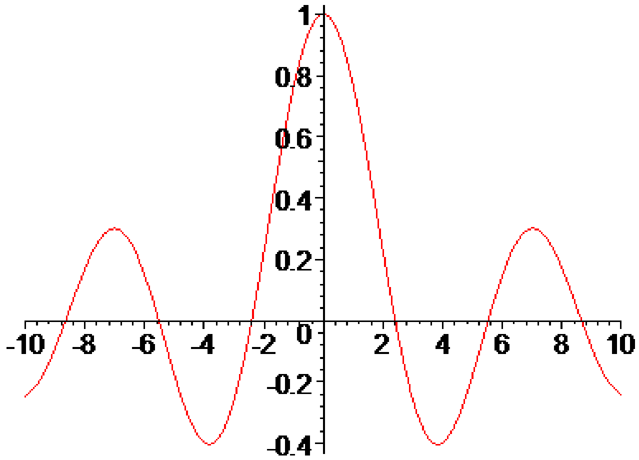
$h_t : x \to e^{it\cos x}$, $E(e^{it\cos X}) = \int_{\mathbb{R}} (x \to e^{it\cos x}) dp_X = \int_{\mathbb{R}} e^{itx}.f_{\cos X} dx$ pour tout réel t.

Comme $p_X([0,2\pi[) = 1$, $\int_{\mathbb{R}} (x \to e^{it\cos x}) dp_X = \int_{[0,2\pi[} (x \to e^{it\cos x}) dp_X = \int_0^{2\pi} e^{it\cos x} \frac{dx}{2\pi}$, d'une part.

D'autre part, $F_{\cos X} = \frac{2\pi - 2\arccos x}{2\pi}$ (faire un dessin) et la densité $f_{\cos X}$ vaut donc $\frac{1}{\pi\sqrt{1-x^2}}$ sur $]-1,1[$ et 0 ailleurs, d'où $\int_{\mathbb{R}} e^{itx}.f_{\cos X} dx = \frac{1}{\pi}\int_{-1}^{1} \frac{e^{itx}}{\sqrt{1-x^2}}$. On vérifie que effectivement, en posant $u = \cos x$:

$$\int_0^{2\pi} e^{it\cos x} \frac{dx}{2\pi} = \frac{1}{\pi}\int_0^{\pi} e^{it\cos x} dx = \frac{1}{\pi}\int_{-1}^{1} \frac{e^{itu}}{\sqrt{1-u^2}} du .$$

La fonction sinus étant impaire, on a aussi $E(e^{it\cos X}) = \frac{1}{\pi}\int_0^{\pi} \cos(t\cos x) dx = \frac{1}{\pi}\int_{-1}^{1} \frac{\cos(tu)}{\sqrt{1-u^2}} du$. La fonction $t \to \frac{1}{\pi}\int_0^{\pi} \cos(t\cos x) dx$ se note $J_0(t)$, J_0 étant la fonction de Bessel d'indice 0. Voici le graphe de la fonction oscillante J_0, fonction paire, majorée par 1, de limite 0 en l'infini :



5) La formule des probabilités totales dans le cas continu

La formule discrète $p(A) = \sum_n p(A/X=n).p(X=n)$ se généralise quand X (diffuse) a une densité f_X :

$$p(A) = \int_{\mathbb{R}} p(A/X=t) f_X(t) dt$$ (se rappeler que $p(X=t)=0$).

Exemple. On se promène dans un plan en faisant des pas de longueur 1, les directions de chaque pas étant aléatoires mais uniformes et indépendantes entre 0 et 2π radians. Si le point de départ de la promenade est l'origine O du plan rapporté à un repère orthonormé, la variable aléatoire égale à la position du n[ième] pas est

$$S_n = \sum_1^n X_i \, , X_i = (\cos\theta_i, \sin\theta_i),$$

les angles θ_i étant uniformes et indépendants sur $[0,2\pi[$.

Prouvons que $p(\| S_3 \| \le 1) = 1/4$. On utilise la formule ci-dessus :

$$p(\| S_3 \| \le 1) = \int_0^2 p(\| S_3 \| \le 1 / \| S_2 \| = R).f_{\|S_2\|}(R) dR$$

(on a sommé de 0 à 2 parce que il est impossible que $|S_3|\leq 1$ si $\| S_2 \| > 2$).

Par symétrie, $p\left(\| S_2 \| \leq R\right) = m\left\{\theta \in [0,2\pi[: \cos\theta \leq (R^2-2)/2\right\} = 1-\dfrac{1}{\pi}\arccos\left(\dfrac{R^2-2}{2}\right)$; en dérivant, il vient

$f_{|S_2|}(R) = 2/\pi\sqrt{4-R^2}$. D'autre part, $p\left(\| S_3 \| \leq 1 / \| S_2 \| = R\right) = arc\cos(R/2)$, d'où

$$p\left(\| S_3 \| \leq 1\right) = \frac{2}{\pi^2}\int_0^2 arc\cos(R/2).\frac{1}{\sqrt{4-R^2}}dR = \frac{2}{\pi^2}\int_0^1 arc\cos(x).\frac{1}{\sqrt{1-x^2}}dx = \frac{2}{\pi^2}\left[\frac{(\arccos x)^2}{2}\right]_1^0 = \frac{1}{4}.$$

****