**PRODUIT DE VARIABLES ALÉATOIRES POSITIVES ET INDÉPENDANTES** (mise à jour)

www.daniel-saada.eu

Toutes les variables aléatoires considérées sont définies sur le même triplet $(\Omega,\mathcal{T},p)$ et positives. Si $X$ est une telle variable aléatoire, la fonction de répartition $F$ de $X$, définie par $F(t)=p(X\le t)$, est nulle sur $]-\infty,0[$ et $F(0)=p(X=0)$.

On dit que $X$ est à *densité* s'il existe une fonction $f$ positive et intégrable sur $[0,+\infty[$ telle que $F(t)=\int_0^t f(x)dx$ sur $\mathbb{R}^+$ ; il en résulte que $\int_0^{+\infty} f(x)dx=1$, que $F$ est continue et $F(0)=0$, que $p(X=t)=0$ pour tout $t$ ( $X$ est alors qualifiée de diffuse) et enfin que $p(a<X<b)=p(a\le X\le b)=\int_a^b f$ .

## A – Fonction de répartition d'un produit $XY$ quand l'une des variables a une densité

On suppose que $Y$ a une densité $g$ et on appelle $G$ la fonction de répartition de $Y$. On se propose de calculer $p(XY\le a)$. Pour $a=0$, $p(XY\le 0)=p(XY=0)=p(X=0)$ car $p(Y=0)=0$. On suppose désormais $a>0$.

**1) On suppose $Y$ majorée, $0\le Y<M$**

On introduit les réels $t_k=k\dfrac{M}{n}$, $k$ allant de $0$ à $n-1$, l'entier $n$ étant $\ge 2$ ; alors,

$$(XY\le a)=\sum_{k=0}^{n-1}(XY\le a \text{ et } t_k\le Y<t_{k+1}) \text{ et } p(XY\le a)=\sum_{k=0}^{n-1}p(XY\le a \text{ et } t_k\le Y<t_{k+1}).$$

Comme $X$ et $Y$ sont positives, et en convenant que $a/t_0=+\infty$ :

$$(X\le a/t_{k+1} \text{ et } t_k\le Y<t_{k+1})\subset(XY\le a \text{ et } t_k\le Y<t_{k+1})\subset(X\le a/t_k \text{ et } t_k\le Y<t_{k+1}).$$

Comme $X$ et $Y$ sont indépendantes, et en convenant que $p(X\le a/t_0)=1$ :

$$p(X\le a/t_{k+1}).\mathrm{p}(t_k\le Y<t_{k+1})\subset p(XY\le a \text{ et } t_k\le Y<t_{k+1})\subset p(X\le a/t_k).\mathrm{p}(t_k\le Y<t_{k+1}) \text{, d'où}$$

**(a)** $$\sum_{k=0}^{n-1}p(X\le a/t_{k+1}).\mathrm{p}(t_k\le Y<t_{k+1}) \le p(XY\le a)\le\sum_{k=0}^{n-1}p(X\le a/t_k).\mathrm{p}(t_k\le Y<t_{k+1}).$$

Comme $Y$ est diffuse, et en posant $F(a/t_0)=1$:

$$\sum_{k=0}^{n-1}F(a/t_{k+1})\big(G(t_{k+1})-G(t_k)\big) \le p(XY\le a)\le\sum_{k=0}^{n-1}F(a/t_k)\big(G(t_{k+1})-G(t_k)\big) .$$

**b)** $\sum_{k=0}^{n-1}F(a/t_k)\big(G(t_{k+1})-G(t_k)\big)-\sum_{k=0}^{n-1}F(a/t_{k+1})\big(G(t_{k+1})-G(t_k)\big)$ tend vers $0$ quand n infini.

Les différences $t_{k+1}-t_k$ valant $1/n$, à partir d'un certain rang $n_0$, $G(t_{k+1})-G(t_k)\le\varepsilon$ pour tout $k$ car G est uniformément continue sur $[0,M]$, et donc pour $n\ge n_0$ :

$$0\le\sum_{k=0}^{n-1}F(a/t_k)\big(G(t_{k+1})-G(t_k)\big)-\sum_{k=0}^{n-1}F(a/t_{k+1})\big(G(t_{k+1})-G(t_k)\big)\le\varepsilon\Big(\sum_{k=0}^{n-1}F(a/t_k)-\sum_{k=0}^{n-1}F(a/t_{k+1})\Big)\le\varepsilon .$$

**c)** $\sum_{k=0}^{n-1}F(a/t_k)\big(G(t_{k+1})-G(t_k)\big)$ a pour limite $\int_0^M F(a/x)g(x)dx$ .

Comme $\int_0^M F(a/x)g(x)dx=\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}}F(a/x)g(x)dx$,

$$0\le\sum_{k=0}^{n-1}F(a/t_k)\big(G(t_{k+1})-G(t_k)\big)-\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}}F(a/x)g(x)dx =\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}}(F(a/t_k)-F(a/x)g(x)dx.$$

Or, $\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}}(F(a/t_k)-F(a/x)g(x)dx \le \sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}}(F(a/t_k)-F(a/t_{k+1})g(x)dx$, et on retombe sur **b)**.

**d)** $p(XY \le a)=\int_0^M F(a/x)g(x)dx$

C'est la conséquence de **a)**, **b)**, **c)**.

**2) On ne suppose plus Y majorée**

$(XY \le a)=(XY \le a \text{ et } Y < M)+(XY \le a \text{ et } Y \ge M)$ et $p(XY \le a \text{ et } Y \ge M) \le p(Y \ge M) \to 0$ quand $M$ tend vers l'infini, donc $p(XY \le a)=\lim_{M\to+\infty} p(XY \le a \text{ et } Y < M)$.

On a appris en **1)** que $p(XY \le a \text{ et } Y < M)=\int_0^M F(a/x)g(x)dx$ ; en faisant tendre $M$ vers l'infini :

$$p(XY \le a)=\int_0^{+\infty} F(a/x)g(x)dx$$

car l'intégrale existe. En effet, $0 \le F(a/x)g(x) \le g(x)$ et $g$ est intégrable sur $\mathbb{R}^+$.

Si $X$ avait une densité $f$, on aurait trouvé $p(XY \le a)=\int_0^{+\infty} G(a/x)f(x)dx$. L'égalité des deux intégrales se prouve ainsi : $\int_0^{\infty} F(a/x)g(x)dx=\int_0^{\infty} g(x)\left(\int_0^{a/x} f(t)dt\right)dx=\int_0^{\infty} f(t)\left(\int_0^{a/t} g(x)dx\right)dt=\int_0^{\infty} f(t)G(a/t)dt$.

## B – <u>Densité de $XY$ quand $X$ et $Y$ ont chacune une densité</u>

Si $X$ a une densité $f$ et $Y$ a une densité $g$, $XY$ a pour densité $\int_0^{+\infty}\frac{f(a/x)g(x)}{x}dx$, ou $\int_0^{+\infty}\frac{g(a/x)f(x)}{x}dx$ (changer $x$ en $a/x$), *intégrales qui existent pour presque tout* $a>0$.

DÉMONSTRATION

Les fonctions $x \to \frac{f(a/x)g(x)}{x}$ sont mesurables et positives sur $]0,+\infty[$, donc soit $\int_0^{+\infty}\frac{f(a/x)g(x)}{x}dx$ existe, soit $\int_0^{+\infty}\frac{f(a/x)g(x)}{x}dx=+\infty$. Or,

$$\int_0^a\left(\int_0^{+\infty}\frac{f(u/x)g(x)}{x}dx\right)du=\int_0^{+\infty} g(x)\left(\int_0^a\frac{f(u/x)}{x}du\right)dx=\int_0^{+\infty} g(x)F(a/x)dx=p(XY \le a)$$

ce qui prouve que $\int_0^{+\infty}\frac{f(u/x)g(x)}{x}dx$ est finie pour presque tous les $u$ de $[0,a]$, et donc pour presque tous les $u$ de $]0,+\infty[$, et que cette intégrale est une densité de $XY$.

*Remarque*. Pour retenir cette formule, on dérive hardiment $p(XY \le a)=\int_0^{+\infty} F(a/x)g(x)dx$ par rapport à $a$, en utilisant que la densité $f$ est, *presque partout*, la dérivée de la fonction de répartition $F$.

*Exemple où* $\int_0^{+\infty}\frac{f(a/x)g(x)}{x}dx=+\infty$ *pour* $a=1$ (Guy BOIZET).

Soit $g$ une fonction positive et intégrable sur $[1,+\infty[$ telle que $g^2$ n'y soit pas intégrable : $f(x)=\frac{g(1/x)}{x^2}$ est définie et intégrable sur $]0,1]$ car

$$\int_0^1 \frac{g(1/x)}{x^2}dx = \int_1^{+\infty} g(x)dx \text{ et } \frac{f(1/x)g(x)}{x} = x.g^2(x) \text{ n'est pas intégrable sur } [1,+\infty[ .$$

En complétant $g$ par $0$ sur $[0,1]$, et $f$ par $0$ sur $]1,+\infty[$, nous tenons notre exemple.

*Exemple où* $\int_0^{+\infty} \frac{f(a/x)g(x)}{x}dx = +\infty$ *pour* *<u>tout rationnel</u>* $a > 0$.

On prend pour $g$ la fonction en escalier définie par : $g = n$ sur $[n, n+1/n^3[$, $g = 0$ sur $[n+1/n^3, n+1[$ ; $g$ est intégrable car $\int_1^{+\infty} g = \sum_1^{+\infty} 1/n^2 = \pi^2/6$ (l'intégrale ne vaut pas 1, c'est sans importance), $g^2$ ne l'est pas.

Pour $a > 0$, $\frac{f(a/x)g(x)}{x} = \frac{1}{a^2}x.g(x).g(x/a)$ ; si $a = p/q$,

$$\int_0^{+\infty} \frac{f(a/x)g(x)}{x}dx = \frac{1}{a^2}\int_0^{+\infty} x.g(x).g(x/a)dx = \frac{1}{a^2q^2}\int_0^{+\infty} x.g(x/p).g(x/q)dx .$$

Quand $n.(pq) \le x < n.(pq) + 1/(npq)^3$, $nq \le x/p < nq + 1/p(npq)^3 \le nq + 1/(nq)^3$ et $g(x/p) = nq$.

De même, $g(x/q) = np$, donc $\int_{npq}^{npq+1/(npq)^3} x.g(x/p).g(x/q)dx \ge \frac{1}{(npq)^3}.npq.nq.np \ge \frac{1}{pq}$ ; il en résulte que

$$\int_0^{+\infty} x.g(x/p).g(x/q)dx \ge \sum_n \int_{npq}^{npq+1/(npq)^3} x.g(x/p).g(x/q)dx = +\infty .$$

On voit donc que l'expression « $\int_0^{+\infty} \frac{f(a/x)g(x)}{x}dx$ existe presque partout » n'est pas un vain mot.

*Existence des intégrales* *<u>pour tout réel</u>* *$a$ quand $f$ et $g$ sont majorées sur* $]0,1]$.

Soit $K$ et $L$ des majorants de $f$ et $g$ :

-- sur $[a,+\infty[$, $0 \le \frac{f(a/x)g(x)}{x} \le K\frac{g(x)}{x} \le \frac{K}{a}g(x)$, donc l'intégrale converge sur $[a,+\infty[$ ;

– sur $]0,1]$, $0 \le \frac{f(a/x)g(x)}{x} \le L\frac{f(a/x)}{x}$, le changement de variable $x \leftarrow a/x$ montre l'existence de $\int_0^1 \frac{f(a/x)}{x}dx$ et donc de $\int_0^1 \frac{f(a/x)g(x)}{x}dx$.

En résumé :

> Si $X$ et $Y$ sont des variables aléatoires positives et indépendantes de fonctions de répartition respectives $F$ et $G$ et $a$ un réel positif :
>
> – si $X$ a une densité $f$, $p(XY \le a) = \int_0^{+\infty} G(a/x)f(x)dx$ ;
>
> – si $Y$ a une densité $g$, $p(XY \le a) = \int_0^{+\infty} F(a/x)g(x)dx$ ;
>
> – si $X$ et $Y$ sont de densités respectives $f$ et $g$, la densité en $a$ de $XY$ est, *pour presque tout* $a$, $\int_0^{+\infty} \frac{f(a/x)g(x)}{x}dx$ ou $\int_0^{+\infty} \frac{g(a/x)f(x)}{x}dx$.

## C – <u>Deux applications</u>

**1) Loi d'un produit de variables uniformes indépendantes**

$X_1, X_2, ..., X_n$ sont des variables aléatoires uniformes et indépendantes sur $[0,1]$. Nous allons prouver que la fonction de répartition $F_n$ de leur produit est $F_n(a) = p(X_1 X_2 ... X_n \leq a) = \frac{1}{(n-1)!}\int_0^a (-\ln x)^{n-1} dx$ quand $a \in [0,1]$, et donc que la densité du produit est $f_n(a) = \frac{(-\ln a)^{n-1}}{(n-1)!}$ sur $]0,1]$, $f_n$ n'étant pas définie en $a = 0$.

Lorsque $n = 1$, on posera $F = F_1$ et $f = f_1$.

Démonstration par récurrence sur $n$ :

– quand $n = 2$, on sait que $p(X_1 X_2 \leq a) = \int_0^{+\infty} F(a/x) f(x) dx = \int_0^1 F(a/x) dx$ car $f$ est nulle sur $]1,+\infty[$ et vaut $1$ sur $[0,1]$ : d'autre part, $F(a/x) = 1$ si $x \leq a$ et $F(a/x) = a/x$ si $x \geq a$,

$$p(XY \leq a) = \int_0^1 F(a/x) dx = \int_0^a 1 dx + \int_0^1 (a/x) dx = a - \ln a$$

or, $a - \ln a = \int_0^a (-\ln x) dx$.

– supposons la formule vraie pour $n$ et prouvons la pour $n+1$ ; comme $X_1 X_2 ... X_n$ et $X_{n+1}$ sont indépendantes,

$$p(X_1 X_2 ... X_n X_{n+1} \leq a) = p(X_{n+1} X_1 ... X_n \leq a) = \int_0^{+\infty} F(a/x) f_n(x) dx$$

et

$$\int_0^{+\infty} F(a/x) f_n(x) dx = \int_0^a 1. f_n(x) dx + \int_a^1 (a/x) f_n(x) dx = \int_0^a f_n(x) dx + a\left[-f_{n+1}(x)\right]_a^1 = \int_0^a f_n(x) dx + a f_{n+1}(a) \ ;$$

reste à vérifier que $\int_0^a f_n(x) dx + a f_{n+1}(a) = \int_0^a f_{n+1}(x) dx$, ce qui se fait en intégrant par parties.

**2) Densité d'une somme de deux variables aléatoires indépendantes à densité**

Soit $f$ et $g$ les densités des variables aléatoires indépendantes $X$ et $Y$ ; $e^X$ et $e^Y$ sont des variables aléatoires positives et indépendantes et $X + Y \leq a$ équivaut à $e^X . e^Y \leq e^a$.

La densité de $e^X$ est $f(\ln x) \times \frac{1}{x}$ pour $x > 0$ (et 0 ailleurs) car $e^X \leq t \Leftrightarrow X \leq \ln t$ ; de même, la densité de $e^Y$ est $g(\ln x) \times \frac{1}{x}$ pour $x > 0$ et 0 ailleurs. La densité de $e^X . e^Y$ en $u > 0$ est donc d'après la partie **B**

$$h(u) = \frac{1}{u}\int_0^{+\infty} \frac{1}{x} f\left(\ln \frac{u}{x}\right) g(\ln x) dx$$

et $p(X + Y \leq a) = \int_0^{e^a} h(u) du = \int_{-\infty}^a h(e^v) e^v dv$.

Or $h(e^v) e^v = \int_0^{+\infty} \frac{1}{x} f(v - \ln x) g(\ln x) dx = \int_{-\infty}^{+\infty} f(v-t) g(t) dt$. On retrouve la formule bien connue qui donne la densité de $X + Y$ en un réel $v$ :

$$\int_{-\infty}^{+\infty} f(v-t) g(t) dt \text{, ou } \int_{-\infty}^{+\infty} f(t) g(v-t) dt .$$

Un petit calcul donne alors

$$p(X + Y \leq a) = \int_{-\infty}^{+\infty} F(a-t) g(t) dt \text{, ou } p(X + Y \leq a) = \int_{-\infty}^{+\infty} G(a-t) f(t) dt ,$$

$F$ et $G$ étant les fonctions de répartition de $X$ et $Y$.