NOTE 16 - LOIS UNIFORMES CORRÉLÉES

www.daniel-saada.eu

Source : http://www.les-mathematiques.net/phorum/read.php?13,812512

Soit F la fonction de répartition d'une loi normale X **standard** (centrée réduite) : $F(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-t^2/2}dt$.

La variable aléatoire $U=F(X)$ est à valeurs dans $]0,1[$; elle est uniforme car F est croissante et bijective :

$$p(U\le a)=p\left(X\le F^{-1}(a)\right)=F\left(F^{-1}(a)\right)=a\,,\ a\in]0,1[.$$

En posant $G(x)=\sqrt{12}(F(x)-0,5)$, la variable $G(X)$ est uniforme sur $]-\sqrt{3},\sqrt{3}[$ donc centrée et réduite.

$r(X,Y)=\dfrac{E(XY)-E(X)E(Y)}{\sigma(X)\sigma(Y)}$ est le coefficient de corrélation linéaire de deux variables aléatoires X et Y ; si X et Y sont centrés et réduites, $r(X,Y)=E(XY)$.

Si X et Z sont deux lois normales indépendantes centrées et réduites, si $Y=X\cos t+Z\sin t$, alors Y est aussi une loi normale centrée réduite et donc $V=G(Y)$ est uniforme comme $U=G(X)$. En outre :

a) $r(X,Y)=E(XY)=\cos t$ car $E(X^2)=V(X)=1$ et $E(XZ)=E(X)E(Z)=0$ par indépendance ;

b) $r(U,V)=\dfrac{6}{\pi}\arcsin\left(\dfrac{\cos t}{2}\right)$, comme cela va être démontré.

> Pour simuler U et V uniformes sur $]-\sqrt{3},\sqrt{3}[$ telles que $r(U,V)=r$, on prendra donc $\cos t=2\sin(r\pi/6)$, X et Z indépendantes de loi $\mathcal{N}(0,1)$, $Y=X\cos t+Z\sin t$, et enfin $U=G(X)$ et $V=G(Y)$; les variables $U/\sqrt{12}+0,5$ et $V/\sqrt{12}+0,5$ seront alors uniformes sur $]0,1[$ et de coefficient de corrélation r également.

Il nous faut donc établir le

Théorème : Si X et Z sont des lois normales centrée réduites indépendantes, $r\in[-1,1]$ et $Y=rX+Z\sqrt{1-r^2}$:

$$r\left(G(X),G(Y)\right)=E\left(G(X)G(Y)\right)=\frac{6}{\pi}\arcsin(r/2).$$

On sait que $E\left(G(X)G(Y)\right)=\int_{R^2}G(x)G(y)f_{X,Y}(x,y)dxdy$, où $f_{X,Y}$ est la densité du couple (X,Y).

Puisque $Y=rX+Z\sqrt{1-r^2}$, la matrice de covariance de (X,Y) est $C=\begin{pmatrix}1&r\\r&1\end{pmatrix}$, d'inverse $\dfrac{1}{1-r^2}\begin{pmatrix}1&-r\\-r&1\end{pmatrix}$.

La densité de (X,Y) en (x,y) est donc

$$f_{X,Y}(x,y)=\frac{1}{2\pi\sqrt{1-r^2}}\exp\left(\frac{-1}{2}(x,y)C^{-1}\begin{pmatrix}x\\y\end{pmatrix}\right)=\frac{1}{2\pi\sqrt{1-r^2}}\exp\left(\frac{-1}{2}\frac{x^2+y^2-2rxy}{1-r^2}\right).$$

Il nous faut donc établir

$$\frac{3}{\pi^2\sqrt{1-r^2}}\int_{R^2}\left(\int_0^x e^{-t^2/2}dt\right)\left(\int_0^y e^{-t^2/2}dt\right)\exp\left(\frac{-1}{2}\frac{x^2+y^2-2rxy}{1-r^2}\right)dxdy=\frac{6}{\pi}\arcsin(r/2).$$

Les trois démonstrations qui suivent, aussi diverses que remarquables, ont pour but de contourner le calcul direct de cette intégrale quadruple : **un logiciel de calcul formel est-il capable, au besoin en le guidant, d'établir cette identité ?**

La note se clôt par une simulation, pour contrôler le résultat théorique obtenu.

Dans toute la suite, on posera $s=\sqrt{1-r^2}$.

DÉMONSTRATION DE GÉRARD LETAC

1) Les polynômes d'Hermite

Le n ième polynôme d'Hermite H_n est défini par $H_n(x)=(-1)^n e^{x^2/2}\dfrac{d^n}{dx^n}(e^{-x^2/2})$.

On trouve facilement $H_0=1$, $H_1(x)=x$, $H_{n+1}(x)=xH_n(x)-nH_{n-1}(x)$ et on prouve que

$$\int_R H_nH_m\frac{e^{-x^2/2}}{\sqrt{2\pi}}dx=0 \text{ quand } m\neq n \text{ et } \int_R H_n.H_n\frac{e^{-x^2/2}}{\sqrt{2\pi}}dx=n!$$

Soit p la probabilité définie par $dp=\dfrac{e^{-x^2/2}}{\sqrt{2\pi}}dx=F'(x)dx$; pour B borélien de R,

$$p(B)=\int_B dp=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}1_B(t)e^{-t^2/2}dt\,.$$

Soit $\mathcal{H}$ l'espace vectoriel des fonctions mesurables vérifiant $\int_R f^2dp<+\infty$:

il est bien connu que $\mathcal{H}$ est un espace de Hilbert pour le produit scalaire $(f/g)=\int_R f.g\,dp$, la norme $\|f\|$ de f étant définie par $\|f\|^2=(f/f)=\int_R f^2\,dp=\int_R f^2\dfrac{e^{-x^2/2}}{\sqrt{2\pi}}dx$; par exemple $\|H_n\|=\sqrt{n!}$.

E désignant l'espérance pour la probabilité p :

$$(f/g)=\int_R f(x)g(x)\frac{e^{-x^2/2}}{\sqrt{2\pi}}dx=E\big(f(X)g(X)\big) \text{ et } (f/H_0)=E\big(f(X)\big).$$

Il en résulte par exemple que $E\big(H_n(X)H_m(X)\big)=0$ si $n\neq m$ et $E\big(H_n(X)H_n(X)\big)=n!$; comme Y a même loi que X, $E\big(H_n(Y)H_m(Y)\big)=0$ si $n\neq m$ et $E\big(H_n(Y)H_n(Y)\big)=n!$

Les égalités $\int_R H_nH_mdp=0$ quand $m\neq n$ et $\int_R H_n.H_ndp=n!$ montrent que $(H_n)_{n\in\mathbb{N}}$ est une famille orthogonale de $\mathcal{H}$. On démontre[1] que cette famille est une base (hilbertienne) de $\mathcal{H}$, $(H_n/\sqrt{n!})_{n\in\mathbb{N}}$ étant orthonormée. Il en résulte que pour toute f de $\mathcal{H}$:

$$f=\sum_{n=0}^{\infty}\left(f/\frac{H_n}{\sqrt{n!}}\right)\frac{H_n}{\sqrt{n!}} \text{ \underline{au sens de la norme} et } \|f\|^2=(f/f)=\sum_{0}^{\infty}\left(f/\frac{H_n}{\sqrt{n!}}\right)^2.$$

2) Développement de la fonction G en série d'Hermite

Comme pour les séries de Fourier, on s'est demandé à quelles conditions on pouvait écrire, pour tout x réel :

$$f(x)=\sum_{n=0}^{\infty}\left(f/\frac{H_n}{\sqrt{n!}}\right)\frac{H_n(x)}{\sqrt{n!}} \text{ (Fourier-Hermite Series en anglais).}$$

On trouvera dans http://www.daniel-saada.eu/Notes/18-convergence-ponctuelle-des-series-d-Hermite.pdf

[1] http://www.daniel-saada.eu/Notes/17-Polynomes-d-Hermite.pdf

des conditions suffisantes, vérifiées par G, pour qu'il en soit ainsi :

La suite $a_n=(G/\frac{H_n}{\sqrt{n!}})$ vérifie donc $G(x)=\sum_1^\infty a_nH_n(x)/\sqrt{n!}$ pour tout x réel.

Cette suite a deux propriétés :

- $a_{2n}=0$; en effet, $\sqrt{(2n)!}\,a_n=(G/H_{2n})=E\left(G(X)H_{2n}(X)\right)$, G est impaire et H_{2n} est pair.
- $\sum_1^\infty a_n^2=1$ car $(G/G)=\int_R G(x)^2F'(x)dx=\frac{1}{6\sqrt{3}}\left[G(x)^3\right]_{x=-\infty}^{x=+\infty}=1$.

On peut donc écrire $G(x)=\sum_0^\infty a_{2n+1}H_{2n+1}(x)/\sqrt{(2n+1!}$, avec $\sum_0^\infty a_{2n+1}^2$.

3) $E\left(e^{tX-t^2/2}e^{wY-w^2/2}\right)=\sum_{n=0}^\infty\sum_{m=0}^\infty\frac{t^n}{n!}\frac{w^m}{m!}E\left(H_n(X)H_m(Y)\right)$

On utilise la formule bien connue [2] $\sum_{n=0}^\infty t^nH_n(x)/n!=e^{tx-t^2/2}$: aussi,

$$e^{tX-t^2/2}e^{wY-w^2/2}=\sum_{n=0}^\infty t^nH_n(X)/n!\times\sum_{n=0}^\infty w^nH_n(Y)/n!$$

et donc $e^{tX-t^2/2}e^{wY-w^2/2}=\sum_{n=0}^\infty\sum_{m=0}^\infty\frac{t^n}{n!}\frac{w^m}{m!}H_n(X)H_m(Y)$.

On suppose t et w fixés dans R.

Pour prouver $E\left(e^{tX-t^2/2}e^{wY-w^2/2}\right)=\sum_{n=0}^\infty\sum_{m=0}^\infty\frac{t^n}{n!}\frac{w^m}{m!}E\left(H_n(X)H_m(Y)\right)$, on utilise la linéarité de l'espérance et on prouve que $E\left(\sum_{n>N}^\infty\sum_{m>N}^\infty\frac{t^n}{n!}\frac{w^m}{m!}H_n(X)H_m(Y)\right)\to 0$ quand $N\to+\infty$. Or,

$$\left|E\left(\sum_{n>N}^\infty\sum_{m>N}^\infty\frac{t^n}{n!}\frac{w^m}{m!}H_n(X)H_m(Y)\right)\right|\le\sum_{n>N}^\infty\sum_{m>N}^\infty\frac{|t|^n}{n!}\frac{|w|^m}{m!}\left|E\left(H_n(X)H_m(Y)\right)\right|.$$

Avec $\left(E(AB)\right)^2\le E(A^2)E(B^2)$, $E\left(H_n(X)H_n(X)\right)=n!$, $E\left(H_n(Y)H_n(Y)\right)=n!$, on obtient

$\sum_{n>N}^\infty\sum_{m>N}^\infty\frac{|t^n|}{n!}\frac{|w^m|}{m!}\left|E\left(H_n(X)H_m(Y)\right)\right|\le\sum_{n>N}^\infty\sum_{m>N}^\infty\frac{|t^n|}{n!}\frac{|w^m|}{m!}\sqrt{n!m!}=\left(\sum_N^\infty\frac{|t|^n}{\sqrt{n!}}\right)^2$, qui est le carré du reste d'une série convergente.

4) $\boxed{E\left(H_n(X)H_m(Y)\right)=0 \text{ si } n\ne m \text{ et } E\left(H_n(X)H_n(Y)\right)=n!r^n}$.

$E\left(e^{tX-t^2/2}e^{wY-w^2/2}\right)=e^{twr}$ par calcul direct avec la formule $E(e^{aX})=e^{a^2/2}$ (X est normale $\mathcal{N}(0,1)$) : en effet,

$$E(e^{aX})=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^{ax}e^{-x^2/2}dx=e^{a^2/2}\text{ (poser } t=x-a\text{).}$$

$E\left(e^{tX-t^2/2}e^{wY-w^2/2}\right)=e^{-(t^2+w^2)/2}E(e^{tX+wY})$; $tX+sY=tX+w(rX+sZ)=(t+wr)X+wsZ$

X et Z étant indépendantes : $E(e^{tX+wY})=E(e^{(t+wr)X})E(e^{wsZ})=e^{(t+wr)^2/2}e^{(ws)^2/2}=e^{(t^2+w^2+2twr)/2}$

[2] http://en.wikipedia.org/wiki/Hermite_polynomials

et $E\left(e^{tX-t^2/2}e^{wY-w^2/2}\right)=e^{twr}$.

On a donc $\sum_{n=0}^{\infty}\sum_{m=0}^{\infty}\frac{t^n}{n!}\frac{w^m}{m!}E\left(H_n(X)H_m(Y)\right)=e^{twr}=\sum_{0}^{\infty}\frac{t^n w^n r^n}{n!}=\sum_{0}^{\infty}\frac{t^n}{n!}\frac{w^n}{n!}n!r^n$.

Par identification[3] , il vient

$$E\left(H_n(X)H_m(Y)\right)=0 \text{ si } n\neq m \text{ et } E\left(H_n(X)H_n(Y)\right)=n!r^n.$$

5) $\boxed{E\left(G(X)G(Y)\right)=\sum_1^{\infty}a_n^2r^n=\sum_0^{\infty}a_{2n+1}^2r^{2n+1}}$

$G(X)G(Y)=\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\frac{a_n}{\sqrt{n!}}\frac{a_m}{\sqrt{m!}}H_n(X)H_m(Y)$ et donc

$$E\left(G(X)G(Y)\right)=\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\frac{a_n}{\sqrt{n!}}\frac{a_m}{\sqrt{m!}}E\left(H_n(X)H_m(Y)\right)=\sum_{1}^{\infty}\frac{a_n^2}{n!}.n!r^n \text{, CQFD.}$$

6) $\boxed{E\left(F(X)e^{tX-t^2/2}\right)=F(t/\sqrt{2})}$

Prouver que $E\left(F(X)e^{tX-t^2/2}\right)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}F(x)e^{tx-t^2/2}e^{-x^2/2}dt=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{t/\sqrt{2}}e^{-x^2/2}dt=F(t/\sqrt{2})$

équivaut à prouver que $\int_{-\infty}^{+\infty}F(x)e^{tx-t^2/2}e^{-x^2/2}dt=\int_{-\infty}^{t/\sqrt{2}}e^{-x^2/2}dt$.

$\int_{-\infty}^{+\infty}F(x)e^{tx-t^2/2}e^{-x^2/2}dt=\int_{-\infty}^{+\infty}F(x)e^{-(x-t)^2/2}dt=\int_{-\infty}^{+\infty}F(u+t)e^{-u^2/2}du$ (avec $u=x-t$)

Posons $I(t)=\int_{-\infty}^{+\infty}F(u+t)e^{-u^2/2}du$: $I(0)=\frac{\sqrt{2\pi}}{2}$ et $\frac{dI}{dt}=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^{-(u+t)^2/2}e^{-u^2/2}du$

On utilise $\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^{ax}e^{-x^2/2}dx=e^{a^2/2}$:

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^{-(u+t)^2/2}e^{-u^2/2}du=\frac{e^{-t^2/2}}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^{-u^2-ut}du=\frac{e^{-t^2/2}}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^{-x^2/2-xt/\sqrt{2}}\frac{dx}{\sqrt{2}}=\frac{e^{-t^2/2}}{\sqrt{2\pi}}\sqrt{2\pi}\frac{e^{t^2/4}}{\sqrt{2}}=\frac{e^{-t^2/4}}{\sqrt{2}}$$

Donc $I(t)=\frac{\sqrt{2\pi}}{2}+\frac{1}{\sqrt{2}}\int_0^t e^{-x^2/4}dx=\sqrt{2\pi}\left(\frac{1}{2}+\frac{1}{2\sqrt{\pi}}\int_0^t e^{-x^2/4}dx\right)=\sqrt{2\pi}\left(\frac{1}{2}+\frac{1}{2\sqrt{\pi}}\int_0^{t/\sqrt{2}}e^{-x^2/2}\sqrt{2}dx\right)$

$I(t)=\sqrt{2\pi}\left(\frac{1}{2}+\frac{1}{\sqrt{2\pi}}\int_0^{t/\sqrt{2}}e^{-x^2/2}dx\right)=\sqrt{2\pi}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{t/\sqrt{2}}e^{-x^2/2}dx$, CQFD.

7) Calcul de a_{2n+1} **et preuve de** $\boxed{E\left(G(X)G(Y)\right)=\frac{6}{\pi}\arcsin(r/2)}$

Comme $E\left(F(X)e^{tX-t^2/2}\right)=F(t/\sqrt{2})=\frac{1}{2}+\frac{1}{\sqrt{2\pi}}\int_0^{t/\sqrt{2}}e^{-x^2/2}dx$ et $e^{-x^2/2}=\sum_0^{\infty}\frac{(-1)^n x^{2n}}{2^n n!}$,

$$E\left(F(X)e^{tX-t^2/2}\right)=\frac{1}{2}+\frac{1}{\sqrt{2\pi}}\sum_0^{\infty}\frac{(-1)^n(t/\sqrt{2})^{2n+1}}{2^n n!(2n+1)}=\frac{1}{2}+\frac{1}{2\sqrt{\pi}}\sum_0^{\infty}\frac{(-1)^n t^{2n+1}}{4^n n!(2n+1)}.$$

[3] Il aurait suffi que les deux séries entières soient égales pour t et w voisins de 0.

Par conséquent, $E\left(G(X)e^{tX-t^2/2}\right)=\sqrt{12}E\left(F(X)e^{tX-t^2/2}\right)-\sqrt{3}E\left(e^{tX-t^2/2}\right)=\sqrt{12}E\left(F(X)e^{tX-t^2/2}\right)-\sqrt{3}$,

et $E\left(G(X)e^{tX-t^2/2}\right)=\dfrac{\sqrt{3}}{\sqrt{\pi}}\sum_0^\infty \dfrac{(-1)^n t^{2n+1}}{4^n n!(2n+1)}$.

Mais $E\left(G(X)e^{tX-t^2/2}\right)=\sum_0^\infty E\left(G(X)\dfrac{t^n H_n(X)}{n!}\right)=\sum_0^\infty \dfrac{t^n}{n!}E\left(G(X)H_n(X)\right)=\sum_0^\infty \dfrac{t^{2n+1}}{(2n+1)!}E\left(G(X)H_{2n+1}(X)\right)$

car $E\left(G(X)H_{2n}(X)\right)=0$. On en déduit : $\dfrac{E\left(G(X)H_{2n+1}(X)\right)}{(2n+1)!}=\dfrac{\sqrt{3}}{\sqrt{\pi}}\dfrac{(-1)^n}{4^n n!(2n+1)}$ par identification des deux

séries entières. Comme $a_{2n+1}=E\left(G(X)\dfrac{H_{2n+1}(X)}{\sqrt{(2n+1)!}}\right)$, $a_{2n+1}=\dfrac{\sqrt{3}}{\sqrt{\pi}}\dfrac{(-1)^n\sqrt{(2n+1)!}}{4^n n!(2n+1)}$.

En conclusion, $E\left(G(X)G(Y)\right)=\sum_0^\infty a_{2n+1}^2 r^{2n+1}=\dfrac{3}{\pi}\sum_0^\infty \dfrac{(2n+1)!\,r^{2n+1}}{8^n (n!)^2(2n+1)^2}=\dfrac{6}{\pi}\arcsin(r/2)$.

DÉMONSTRATION DE SÉBASTIEN DA VEIGA

Rappelons que la densité de (X,Y) en (x,y) est $f_{X,Y}(x,y)=\dfrac{1}{2\pi\sqrt{1-r^2}}\exp\left(\dfrac{-1}{2}\dfrac{x^2+y^2-2rxy}{1-r^2}\right)$.

1) Montrons avec Maple que $\dfrac{\partial f_{X,Y}(x,y)}{\partial r}=\dfrac{\partial^2 f_{X,Y}(x,y)}{\partial x\partial y}$.

```
> f:=(x,y,r)->exp(-(x^2+y^2-2*r*x*y)/2/(1-r^2))/(2*Pi*sqrt(1-r^2));
```

$$f := (x,y,r) \rightarrow \frac{1}{2}\frac{e^{\left(-1/2\,\frac{x^2+y^2-2\,r\,x\,y}{1-r^2}\right)}}{\pi\sqrt{1-r^2}}$$

```
> simplify(diff(f(x,y,r),r));
```

$$-\frac{1}{2}\frac{(r^3-x\,y\,r^2+r\,x^2+r\,y^2-r-x\,y)\,e^{\left(\frac{x^2+y^2-2\,r\,x\,y}{2\,(r-1)\,(r+1)}\right)}}{\pi\,(1-r^2)^{(5/2)}}$$

```
> simplify(diff(f(x,y,r),x,y));
```

$$-\frac{1}{2}\frac{(r^3-x\,y\,r^2+r\,x^2+r\,y^2-r-x\,y)\,e^{\left(\frac{x^2+y^2-2\,r\,x\,y}{2\,(r-1)\,(r+1)}\right)}}{\pi\,(1-r^2)^{(5/2)}}$$

```
> diff(f(x,y,r),x);  diff(f(x,y,r),y);
```

$$-\frac{1}{4}\frac{(2x-2ry)\,\mathrm{e}^{\left(-\frac{x^2+y^2-2rxy}{2(1-r^2)}\right)}}{(1-r^2)^{(3/2)}\,\pi} \qquad -\frac{1}{4}\frac{(2y-2rx)\,\mathrm{e}^{\left(-\frac{x^2+y^2-2rxy}{2(1-r^2)}\right)}}{(1-r^2)^{(3/2)}\,\pi}$$

2) On prouve ensuite que $\frac{\partial}{\partial r}E\left(F(X)F(Y)\right)=\int_{R^2} f_X(x)f_Y(y)f_{X,Y}(x,y)dxdy$.

D'abord $E\left(F(X)F(Y)\right)=\int_{R^2} F(x)F(y)f_{X,Y}(x,y)dxdy$, puis

$$\frac{\partial}{\partial r}E\left(F(X)F(Y)\right)=\int_{R^2} F(x)F(y)\frac{\partial^2 f_{X,Y}(x,y)}{\partial x\partial y}dxdy=\int_{-\infty}^{+\infty}F(x)dx\left(\int_{-\infty}^{+\infty}F(y)\frac{\partial^2 f_{X,Y}(x,y)}{\partial x\partial y}dy\right).$$

En intégrant par parties : $\int_{-\infty}^{+\infty}F(y)\frac{\partial^2 f_{X,Y}(x,y)}{\partial x\partial y}dy=\frac{-1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^{-y^2/2}\frac{\partial f_{X,Y}(x,y)}{\partial x}dy$ car les limites aux infinis de $y\mapsto\frac{\partial f_{X,Y}(x,y)}{\partial x}$ sont nulles. On obtient alors

$$\frac{\partial}{\partial r}E\left(F(X)F(Y)\right)=\frac{-1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}F(x)dx\left(\int_{-\infty}^{+\infty}e^{-y^2/2}\frac{\partial f_{X,Y}(x,y)}{\partial x}dy\right)=\frac{-1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^{-y^2/2}dy\left(\int_{-\infty}^{+\infty}F(x)\frac{\partial f_{X,Y}(x,y)}{\partial x}dx\right)$$

Une deuxième intégration par parties donne $\frac{\partial}{\partial r}E\left(F(X)F(Y)\right)=\frac{1}{2\pi}\int_{-\infty}^{+\infty}e^{-y^2/2}e^{-y^2/2}f_{X,Y}(x,y)dxdy$,

et donc $\frac{\partial}{\partial r}E\left(F(X)F(Y)\right)=\frac{1}{4\pi^2\sqrt{1-r^2}}\int_{-\infty}^{+\infty}e^{-x^2/2}e^{-y^2/2}\exp\left(\frac{-1}{2}\frac{x^2+y^2-2rxy}{1-r^2}\right)dxdy$.

3) Calculons cette intégrale double avec Maple.

\> **assume(a<0);assume(b<0);**

\> **e:=(x,y)->exp(a*x^2+b*y^2+c*x*y);**

$$e:=(x,y)\rightarrow \mathrm{e}^{(a\,x^2+b\,y^2+c\,x\,y)}$$

\> **int(int(e(x,y),y=-infinity..infinity),x=-infinity..infinity);**

$$\begin{cases}\frac{2\,\pi}{\sqrt{4\,a\sim\,b\sim-c^2}} & \mathrm{csgn}\left(-a\sim+\frac{c^2}{4\,b\sim}\right)=1\\ \infty & \mathit{otherwise}\end{cases}$$

On en déduit $\frac{\partial}{\partial r}E\left(F(X)F(Y)\right)=\frac{1}{4\pi^2\sqrt{1-r^2}}\frac{2\pi\sqrt{1-r^2}}{\sqrt{4-r^2}}=\frac{1}{2\pi\sqrt{4-r^2}}$, d'où

$E\left(F(X)F(Y)\right)=\frac{1}{2\pi}\arcsin\left(\frac{r}{2}\right)+k$; on détermine la constante k en faisant $r=0$:

$E\left(F(X)F(Y)\right)=E\left(F(X)F(Z)\right)=E\left(F(X)\right)E\left(F(Z)\right)=1/4$, donc $k=1/4$.

Enfin, $E\left(G(X)G(Y)\right)=12E\left(F(X)F(Y)\right)-3=\frac{6}{\pi}\arcsin(r/2)$.

DÉMONSTRATION DE « EGOROFF» (qui souhaite rester anonyme)

Dans toute la suite on supposera $0<r<1$**, on posera toujours** $s=\sqrt{1-r^2}$ et on emploiera la **notation** $P(A,B)$ pour $P(A\cap B)$.

étape 1 : on écrit $E\left(F(X)F(Y)\right)=E\left(F(X)F(rX+sZ)\right)=\int_{R^2}F(x)F(rx+sz)f_{X,Z}(x,y)dxdz$

et donc $E\left(F(X)F(rX+sZ)\right)=\int_{R^2}\left(\int_{-\infty}^{x}\frac{e^{-t^2/2}}{\sqrt{2\pi}}dt\right)\left(\int_{-\infty}^{rx+sz}\frac{e^{-w^2/2}}{\sqrt{2\pi}}dw\right)\frac{e^{-(x^2+z^2)/2}}{2\pi}dxdz$

d'où $E\left(F(X)F(Y)\right)=\int_{t\le x,w\le rx+sz}\frac{e^{-(x^2+z^2+t^2+w^2)/2}}{4\pi^2}dxdz$.

étape 2 : il existe T et W normales standard telles que X,Z,T,W soient (mutuellement) indépendantes

Le lecteur pourra se reporter à : http://www.daniel-saada.eu/fichiers/22-existence_de_lois_uniformes.pdf, p. 5.
Toutefois, voici une démonstration directe. Soit $(\Omega,\mathcal{T},P)$ l'espace probabilisé sur lequel sont définies X et Z.
On élargit l'univers Ω en $\Omega'=\Omega\times R^2$ que l'on munit de la tribu produit $\mathcal{T}\otimes Bor(R^2)$ et de la probabilité $P\otimes\mu$, μ étant définie par $d\mu=\frac{1}{2\pi}e^{-(x^2+y^2)/2}dxdy$ (si B est un borélien de R^2, $\mu(B)=\frac{1}{2\pi}\int_B e^{-(x^2+y^2)/2}dxdy$).
Les quatre variables aléatoires définies sur Ω' par

$$X'(\omega,x,y)=X(\omega),\ Z'(\omega,x,y)=Z(\omega),\ T(\omega,x,y)=x,\ W(\omega,x,y)=y$$

suivent la loi normale standard et sont indépendantes. En assimilant X' à X et Z' à Z, $\frac{e^{-(x^2+z^2+t^2+w^2)/2}}{4\pi^2}$ est la densité de (X,Z,T,W) et $E\left(F(X)F(Y)\right)=P(T\le X,W\le rX+sZ)$.

étape 3 : transformations de $P(T\le X,W\le rX+sZ)$

a) $N=\frac{X-T}{\sqrt2}$ est une loi normale standard ; comme $W\le rX+sZ$ équivaut à $X\ge\frac{W-sZ}{r}$, on a

$$P(T\le X,W\le rX+sZ)=P(N\ge0,X\ge\frac{W-sZ}{r}).$$

b) $\frac{W-sZ}{r}$ est une loi normale centrée de variance $c^2=\frac{1+s^2}{r^2}$: $\frac{W-sZ}{r}=cN'$, avec N' normale standard ;

$$P(T\le X,W\le rX+sZ)=P(N\ge0,X-cN'\ge0).$$

c) Il existe deux réels a et b tels que $X-cN'=b\left(a.N+\sqrt{1-a^2}.N''\right)$, avec N'' normale standard et indépendante de N.

$X-cN'-baN$ valant $X\left(1-\frac{ab}{\sqrt2}\right)-cN'+ab\frac{T}{\sqrt2}$ et X , N', T étant indépendantes, $X-cN'-baN$ est une loi normale centrée et on peut poser $X-cN'-baN=dN''$, N'' étant normale standard.
En remplaçant N par $\frac{X-T}{\sqrt2}$, il vient

$$dN''=X\left(1-\frac{ab}{\sqrt2}\right)-cN'+ab\frac{T}{\sqrt2},\text{ d'où } d^2=\left(1-\frac{ab}{\sqrt2}\right)^2+c^2+\frac{a^2b^2}{2}.$$

Traduisons l'indépendance voulue entre N et N'' en imposant $\mathrm{cov}(N,N'')=E(N.N'')=0$:

$$E(dN''.N)=E(XN-cN'N-abN^2)=1/\sqrt{2}-ab$$

On obtient alors $d^2=b^2(1-a^2)=1/2+c^2$ et $b^2=1+c^2$: nous choisirons $b=\sqrt{1+c^2}=\sqrt{2}/r$, d'où $a=r/2$.

On a donc $P(T\le X,W\le rX+sZ)=P\left(N\ge 0,rN+\sqrt{4-r^2}N''\ge 0\right)$ avec N et N'' normales standard indépendantes.

étape 4 : calcul de $P\left(N\ge 0,rN+\sqrt{4-r^2}N''\ge 0\right)$

Par **indépendance**, $P\left(N\ge 0,rN+\sqrt{4-r^2}N''\ge 0\right)=\dfrac{1}{2\pi}\displaystyle\int_{x\ge 0,rx+y\sqrt{4-r^2}\ge 0}e^{-(x^2+y^2)/2}dxdy$.

On intègre $e^{-(x^2+y^2)/2}$ sur le cône de sommet O délimité par les demi-droites $(x=0,y\ge 0)$ et $(rx+y\sqrt{4-r^2}=0,x\ge 0)$. En passant en polaires,

$$\frac{1}{2\pi}\int_{x\ge 0,rx+y\sqrt{4-r^2}\ge 0}e^{-(x^2+y^2)/2}dxdy=\frac{1}{2\pi}\left(\int_0^{+\infty}\rho e^{-\rho^{2}/2}d\rho\right)\left(\int_\theta^{\pi/2}d\theta\right)=\frac{\pi/2-\theta}{2\pi}$$, où $\theta=\arcsin(-r/2)$.

On retrouve bien $E\left(G(X)G(Y)\right)=12E\left(F(X)F(Y)\right)-3=\dfrac{6}{\pi}\arcsin(r/2)$.

CONTRÔLE PAR SIMULATION

Rappelons comment fabriquer deux lois normales X et Z indépendantes centrées et réduites :

$X=\cos(2\pi U)\sqrt{-2\ln(V)}$ et $Z=\sin(2\pi U)\sqrt{-2\ln(V)}$, U et V étant uniformes et indépendantes sur $[0,1]$.

Calcul numérique de F : on utilise la fonction spéciale erf définie par $erf(x)=\dfrac{2}{\sqrt{\pi}}\displaystyle\int_0^x e^{-t^2}dt$ et présente dans tous les logiciels mathématiques ; $F(x)=0,5+0,5erf(x/\sqrt{2})$ et $G(x)=\sqrt{3}.erf(x/\sqrt{2})$.

Algorithme pour la construction de deux lois U et V uniformes sur $]-\sqrt{3},\sqrt{3}[$ telles que $r(U,V)=\rho$:

1. $r\leftarrow 2\sin(\pi\rho/6)$
2. $U\leftarrow random$ et $V\leftarrow random$
3. $X\leftarrow\cos(2\pi U)\sqrt{-2\ln(V)}$ et $Z\leftarrow\sin(2\pi U)\sqrt{-2\ln(V)}$
4. $Y\leftarrow rX+Z\sqrt{1-r^2}$
5. $U=\sqrt{3}.erf(X/\sqrt{2})$ et $V\leftarrow\sqrt{3}.erf(Y/\sqrt{2})$.

Programme Maple pour $\rho=0,5$ et mille simulations :

```
r:=2*sin(Pi/12) ; digits:=5 ; S:=0 ;
for k  from 1 to 1000  do
U:=rand()/10^(12) ; V:=rand()/10^(12) ;
X:=cos(2*Pi*U)*sqrt(-2*ln(V)) ; Z:=sin(2*Pi*U)*sqrt(-2*ln(V)) ; Y:=r*X+sqrt(1-r*r)*Z ;
U:=sqrt(3.)*erf(X/sqrt(2.));V:=sqrt(3.)*erf(Y/sqrt(2.));S:=S+U*V;
od:evalf(S)/1000;
```

Résultats sur cinq fois 1000 simulations : 0,472 0,531 0,467 0,534 0,468 , ce qui est satisfaisant.

Voyons pour finir à quoi ressemblent deux lois uniformes U et V à valeurs entre 0 et 1 et de coefficient de corrélation $\rho = 0,5$: on observera que les valeurs sont plus rapprochées que lorsque U et V sont indépendantes, ce qui peut se quantifier par $E\left(|U-V|\right)$, ou plus simplement par $E\left((U-V)^2\right) = 2/3 - 2E(UV) = \dfrac{1-r(U,V)}{6}$.

U et V sont indépendantes	U et V sont corrélées avec $\rho = 0,5$
0.3957188605 , 0.1931398164	0.0752269711 , 0.5796960366
0.02242417046 , 0.8001874845	0.7457194802 , 0.6636425296
0.4275520569 , 0.8426226844	0.2995761723 , 0.4792852283
0.4122862858 , 0.9964172142	0.4712289339 , 0.5002377320
0.3864083074 , 0.6946071893	0.2593513494 , 0.5572973958
0.7730129800 , 0.7306162929	0.5454457804 , 0.2703617728
0.1065070537 , 0.3964127230	0.8570045342 , 0.8987312090
0.9449133500 , 0.2109364289	0.9515000096 , 0.6356964073
0.7500720722 , 0.4547443970	0.5002267990 , 0.1414496665
0.7366026223 , 0.3298445918	0.4501724224 , 0.0909992912
0.6157320692 , 0.8470187654	0.3334179523 , 0.2909569652
0.4710778522 , 0.3925545928	0.0893094790 , 0.3139184385
0.8364047111 , 0.4742592553	0.7359917562 , 0.2849573024
0.2240850446 , 0.07860611833	0.6426737564 , 0.9818437140
0.7206838296 , 0.5597053223	0.4217858276 , 0.1566472188
0.7596612973 , 0.4301371693	0.5314055036 , 0.1426356281
0.8675579182 , 0.8668184741	0.6405654058 , 0.4396346752
0.9669052975 , 0.4972015204	0.8763028140 , 0.6517101240
0.5561578451 , 0.3636274169	0.0909745320 , 0.1331766186
0.9883832000 , 0.7353422487	0.7829011072 , 0.6390357798