Compléments sur les couples aléatoires

1. Couple image.

Dans ce paragraphe, on va s'intéresser à la loi d'un vecteur aléatoire (S,T) qui s'obtient comme image d'un vecteur (X,Y) par une application u. On connaît la loi du couple (X,Y) et on cherche la loi de $(S,T) = u(X,Y) = (u_1(X,Y), u_2(X,Y))$. On ne considèrera que des couples (S,T) et (X,Y) possédant une densité. La méthode est alors la même que dans le cas unidimensionnel : on cherche la densité du couple (S,T) en faisant un changement de variables mais dans $\mathbf{R}^2$. Nous commençons par la formule de changement de variables dans ce contexte.

1.1. Changement de variables.

Voici le cadre de l'étude. D est un « ouvert » de $\mathbf{R}^2$ et h est une application de D dans $\mathbf{R}$. On voudrait calculer l'intégrale $\iint_D h(x,y)\,dxdy$ en introduisant un nouveau jeu de coordonnées (s,t) c'est à dire en effectuant le changement de variables « $(s,t) = u(x,y)$ » où u est une application de D dans $\mathbf{R}^2$. Comme dans le cas réel, la fonction u doit être régulière.

C^1-difféomorphisme. La formule du changement de variables nécessite de la régularité sur la fonction qui permet de passer d'un jeu de coordonnées à l'autre : celle-ci doit être un C^1-difféomorphisme.

Définition. Soient D et D' deux ouverts de $\mathbf{R}^2$. $u : D \longrightarrow \mathbf{R}^2$ est un C^1-difféormphisme de D sur D' si u est une bijection de D sur D' telle que u et u^{-1} soient de classe C^1.

Si $u : D \longrightarrow \mathbf{R}^2$ est différentiable, sa différentielle $\mathrm{D}u$ est représentée par la matrice suivante – appelée matrice jacobienne –

$$\forall (x,y) \in D, \qquad \mathrm{D}u(x,y) = \begin{pmatrix} \partial_x u_1(x,y) & \partial_y u_1(x,y) \\ \partial_x u_2(x,y) & \partial_y u_2(x,y) \end{pmatrix}.$$

Le jacobien de u est le déterminant de cette matrice, soit

$$\forall (x,y) \in D, \qquad \mathrm{J}u(x,y) = \det\left(\mathrm{D}u(x,y)\right) = \partial_x u_1(x,y)\, \partial_y u_2(x,y) - \partial_x u_2(x,y)\, \partial_y u_1(x,y).$$

On peut retenir le critère suivant :

Proposition 1. *Soient D un ouvert de $\mathbf{R}^2$ et $u : D \longrightarrow \mathbf{R}^2$ une application injective de classe C^1. Si, pour tout $(x,y) \in D$, $\mathrm{J}u(x,y) \neq 0$ alors $u(D)$ est un ouvert de $\mathbf{R}^2$ et l'application u est un C^1–difféomorphisme de D sur $u(D)$.*

Si u est un C^1–difféomorphisme de D sur D' alors u^{-1} est un C^1–difféomorphisme de D' sur D. On a dans ce cas, comme dans le cas réel, $\mathrm{D}u^{-1} = \left(\mathrm{D}u \circ u^{-1}\right)^{-1}$ et en particulier,

$$\forall (s,t) \in D', \quad \mathrm{J}u^{-1}(s,t) = \frac{1}{\mathrm{J}u\left(u^{-1}(s,t)\right)}.$$

Un cas particulier important est celui des fonctions u linéaires :

$$\forall (x,y)\in \mathbf{R}^2, \qquad u(x,y)=\begin{pmatrix} u_1(x,y)\\ u_2(x,y)\end{pmatrix}=\begin{pmatrix} a & b\\ c & d\end{pmatrix}\begin{pmatrix} x\\ y\end{pmatrix}\ ;$$

si M est la matrice qui apparaît ci-dessus, alors $\mathrm{J}u(x,y)=\det(M)$ et donc l'application u est un C^1–difféomorphisme de $\mathbf{R}^2$ dans $\mathbf{R}^2$ si et seulement si $\det(M)\neq 0$.

Formule du changement de variables. Commençons par un énoncé précis.

Théorème 2. *Soient D' un ouvert de $\mathbf{R}^2$, v un C^1–difféomorphisme sur D'. Considérons une fonction $h : v(D')\longrightarrow \mathbf{R}$ borélienne. Si h est positive ou intégrable sur $v(D')$ alors*

$$\iint_{v(D')} h(x,y)\,dxdy=\iint_{D'} h(v(s,t))\,|\mathrm{J}v(s,t)|\,dsdt. \tag{1}$$

La signification de ce résultat est la suivante :

- si h est positive, les deux termes de l'égalité précédente sont soit tout deux infinis soit tout deux finis et égaux ;
- sinon h est intégrable sur $v(D')$ si et seulement si $h\circ v\,|\mathrm{J}v|$ est intégrable sur D' et dans ce cas on a l'égalité (1).

La formule (1) correspond à un changement de variables du type $(x,y)=v(s,t)$ où (x,y) sont les variables d'intégration de départ (dont on veut se débarasser). En pratique, cette situation n'est pas très fréquente ; il faut tout de même se souvenir du passage en coordonnées polaires. On souhaite calculer $\iint h(x,y)\,dxdy$ en posant $x=r\cos\theta$, $y=r\sin\theta$. Ceci correspond à la fonction $v(r,\theta)=(r\cos\theta, r\sin\theta)$ qui est un C^1–difféomorphisme de $]0,+\infty[\times]-\pi,\pi[$ sur $\mathbf{R}^2\backslash\mathbf{R}_-$ (le plan privé de la demi-droite $y=0$, $x\leq 0$). Notez que $v^{-1}(x,y)=(r,\theta)$ avec

$$r=\sqrt{x^2+y^2}, \quad \text{et,} \quad \theta = 2\arctan\left(\frac{y}{x+\sqrt{x^2+y^2}}\right).$$

Exemple. Montrons que $I=\int_{\mathbf{R}}\exp(-x^2/2)\,dx=\sqrt{2\pi}$. Pour cela, calculons I^2 :

$$I^2=\int_{\mathbf{R}}\exp(-x^2/2)\,dx\int_{\mathbf{R}}\exp(-y^2/2)\,dy=\iint_{\mathbf{R}^2}\exp\Big(-(x^2+y^2)/2\Big)\,dxdy,$$

puis comme la demi-droite $y=0$, $x\leq 0$ est négligeable,

$$I^2=\iint_{\mathbf{R}^2}\exp\Big(-(x^2+y^2)/2\Big)\,dxdy=\iint_{\mathbf{R}^2\backslash\mathbf{R}_-}\exp\Big(-(x^2+y^2)/2\Big)\,dxdy.$$

Effectuons le changement de variables $(x,y)=(r\cos\theta, r\sin\theta)=v(r,\theta)$. On a

$$\forall (r,\theta)\in]0,+\infty[\times]-\pi,\pi[, \qquad \mathrm{J}v(r,\theta)=\begin{vmatrix} \partial_r x(r,\theta) & \partial_\theta x(r,\theta)\\ \partial_r y(r,\theta) & \partial_\theta y(r,\theta)\end{vmatrix}=\begin{vmatrix} \cos\theta & -r\sin\theta\\ \sin\theta & r\cos\theta\end{vmatrix}=r.$$

La formule du changement de variables donne alors

$$I^2=\iint_{]0,+\infty[\times]-\pi,\pi[}\exp(-r^2/2)\,r\,drd\theta=\int_{|\theta|<\pi}d\theta\int_0^{+\infty} r\exp(-r^2/2)\,dr=2\pi.$$

I est l'intégrale d'une fonction positive, donc I est positive ; il vient $I=\sqrt{2\pi}$.

Changement de variables $(s,t) = u(x,y)$**.** En pratique, on veut calculer $\iint_D h(x,y)\,dxdy$ et on effectue le changement de variables $(s,t) = u(x,y)$ où u est un C^1–difféomorphisme de D sur $D' := u(D)$. On se ramène au cadre d'application de la formule (1) en considérant $v = u^{-1}$ puisque $(x,y) = u^{-1}(s,t)$. On a alors

$$\iint_D h(x,y)\,dxdy = \iint_{u^{-1}(D')} h(x,y)\,dxdy = \iint_{D'} h\left(u^{-1}(s,t)\right) \left|\mathrm{J}u^{-1}(s,t)\right| \,dsdt,$$

et comme $D' = u(D)$ et $\mathrm{J}u^{-1} = \left(\mathrm{J}u \circ u^{-1}\right)^{-1}$, on a finalement

$$\iint_D h\,(x,y)\;dxdy = \iint_{u(D)} \frac{h\left(u^{-1}(s,t)\right)}{\left|\mathrm{J}u\left(u^{-1}(s,t)\right)\right|}\,dsdt. \tag{2}$$

Pour appliquer cette formule, on doit calculer,

$$\forall (s,t) \in u(D), \qquad \frac{h\left(u^{-1}(s,t)\right)}{\left|\mathrm{J}u\left(u^{-1}(s,t)\right)\right|}.$$

En pratique, on détermine le rapport

$$\forall (x,y) \in D, \qquad \frac{h(x,y)}{\left|\mathrm{J}u(x,y)\right|}$$

que l'on exprime en fonction des coordonnées (s,t). On a ainsi $h(x,y)/|\mathrm{J}u(x,y)| = g(s,t)$ et

$$\iint_D h(x,y)\,dxdy = \iint_{u(D)} g(s,t)\,dsdt.$$

Cela permet dans certains cas d'éviter d'inverser la fonction u.

1.2. Application aux couples aléatoires.

Le point de départ est un vecteur aléatoire $Z = (X,Y)$ dont on connaît la loi $\mathbb{P}_Z$ au travers de sa densité $p_Z(x,y) = \mathbf{1}_D(x,y)\,p(x,y)$ où D est un ouvert de $\mathbf{R}^2$. On cherche la loi du vecteur aléatoire $U = (S,T)$ défini par $(S,T) = u(X,Y)$ où u est un C^1–difféomorphisme sur D. On essaie de déterminer la densité p_U de (S,T). Soit $f : \mathbf{R}^2 \longrightarrow \mathbf{R}$ une fonction borélienne et bornée ; calculons $\mathbb{E}[f(S,T)]$ de deux façons différentes. Tout d'abord,

$$\mathbb{E}[f(S,T)] = \iint_{\mathbf{R}^2} f(s,t)p_U(s,t)\,dsdt\ ;$$

mais aussi, comme la densité de (X,Y) est $\mathbf{1}_D(x,y)\,p(x,y)$,

$$\mathbb{E}[f(S,T)] = \mathbb{E}[f(u(X,Y))] = \iint_D f(u(x,y))p(x,y)\,dxdy.$$

Comme u est un C^1–difféomorphisme, on fait le changement de variables $(s,t) = u(x,y)$ qui conduit à – cf. (2) avec $h(x,y) = f(u(x,y))p(x,y)$ –

$$\mathbb{E}\left[f(S,T)\right] = \iint_{u(D)} f(s,t)\frac{p\left(u^{-1}(s,t)\right)}{\left|\mathrm{J}u\left(u^{-1}(s,t)\right)\right|}\,dsdt = \iint_{\mathbf{R}^2} f(s,t)\,\mathbf{1}_{u(D)}(s,t)\frac{p\left(u^{-1}(s,t)\right)}{\left|\mathrm{J}u\left(u^{-1}(s,t)\right)\right|}\,dsdt.$$

On obtient donc

$$\forall (s,t) \in \mathbf{R}^2, \qquad p_U(s,t) = \mathbf{1}_{u(D)}(s,t)\frac{p\left(u^{-1}(s,t)\right)}{\left|\mathrm{J}u\left(u^{-1}(s,t)\right)\right|}.$$

En pratique, on peut suivre le plan suivant :

- on montre que u est C^1 et injective sur D ;
- on montre que, pour $(x, y) \in D$, $\mathrm{J}u(x, y) \neq 0$ en calculant le jacobien ;
- on détermine le domaine $u(D)$: très souvent on doit inverser u i.e. exprimer (x, y) en fonction de (s, t) ;
- on utilise la méthode ci-dessus : il reste à exprimer $p(x, y)/|\mathrm{J}u(x, y)|$ en fonction de (s, t).

Exemple. Soient X et Y deux v.a. indépendantes, X de loi $\mathcal{E}xp(\lambda)$ et Y de loi $\mathcal{E}xp(\mu)$, $\lambda, \mu > 0$. Déterminons la loi du couple $(X/Y, Y)$.

Pour se ramener à la situation que l'on vient de décrire on introduit les variables $S = X/Y$ et $T = Y$. On cherche la loi de (S, T). Soit f une fonction borélienne et bornée ; calculons $\mathbb{E}[f(S, T)]$. On a

$$\mathbb{E}[f(S,T)] = \mathbb{E}\left[f\left(X/Y, Y\right)\right] = \iint_{]0,+\infty[^2} f(x/y, y)\, \lambda e^{-\lambda x}\, \mu e^{-\mu y}\, dxdy.$$

Considérons la fonction $u :]0, +\infty[^2 \longrightarrow \mathbf{R}^2$ définie par $u(x, y) = (x/y, y)$. u est de classe C^∞ sur l'ouvert $D =]0, +\infty[^2$ et on a

$$\forall (x, y) \in]0, +\infty[^2, \qquad \mathrm{J}u(x, y) = \begin{vmatrix} 1/y & -x/y^2 \\ 0 & 1 \end{vmatrix} = \frac{1}{y} \neq 0.$$

u est injective : si $(x/y, y) = (x'/y', y')$ alors $x = x'$ et $y = y'$. Déterminons $u\left(]0, +\infty[^2\right)$. Trivialement, $u\left(]0, +\infty[^2\right) \subset]0, +\infty[^2$; remarquons que $(s, t) = u(x, y)$ signifie que $x = st$ et $y = t$ (c'est u^{-1}). Il vient alors $u\left(]0, +\infty[^2\right) =]0, +\infty[^2$. La formule du changement de variables donne

$$\mathbb{E}[f(S,T)] = \lambda\mu \iint_{]0,+\infty[^2} f(s, t)\, te^{-\lambda st} e^{-\mu t}\, dsdt.$$

Le couple $U = (S, T)$ a donc pour densité $p_U(s, t) = \lambda\mu\, t \exp\left(-t(\lambda s + \mu)\right) \mathbf{1}_{s>0} \mathbf{1}_{t>0}$.

Remarquons que dans cet exemple (comme dans bien d'autres), on peut éviter de faire un changement de variables en dimension deux. En effet, on a, d'après le théorème de Fubini,

$$\mathbb{E}[f(S,T)] = \lambda\mu \iint_{]0,+\infty[^2} f(\tfrac{x}{y}, y)\, e^{-\lambda x} e^{-\mu y}\, dxdy = \lambda\mu \int_0^{+\infty} e^{-\mu y} \left(\int_0^{+\infty} f(\tfrac{x}{y}, y)\, e^{-\lambda x}\, dx\right) dy.$$

Pour $y > 0$ fixé, on fait le changement de variables réelles $s = x/y$, pour obtenir

$$\int_0^{+\infty} f(x/y, y)\, e^{-\lambda x}\, dx = \int_0^{+\infty} f(s, y)\, ye^{-\lambda sy}\, ds,$$

et par suite

$$\mathbb{E}\left[f(S,T)\right] = \lambda\mu \iint_{]0,+\infty[^2} f(s, y)\, ye^{-\lambda sy} e^{-\mu y}\, dsdy,$$

ce qui est le résultat que l'on avait trouvé.

2. Rudiments sur les vecteurs gaussiens.

Dans ce paragraphe, nous examinons très brièvement un cas particulier de couples aléatoires, celui des vecteurs gaussiens. Les vecteurs gaussiens sont souvent utilisés car, d'un point de vue pratique, ils conduisent à des calculs relativement simples.

Pour pouvoir utiliser certaines notations matricielles, convenons de représenter les couples aléatoires comme des vecteurs colonnes i.e. $X = (X_1, X_2)^t$ où ^{t} désigne la transposition. De plus, si x et y sont deux vecteurs de $\mathbf{R}^2$ on note $x.y$ leur produit scalaire $x.y = x^t y = x_1 y_1 + x_2 y_2$.

Définition. Un couple aléatoire $X = (X_1, X_2)^t$ suit la loi normale centrée réduite de dimension 2, $\mathcal{N}(0, I_2)$, si X_1 et X_2 sont deux v.a.r. normales centrées réduites indépendantes.

Si X est un tel couple alors X possède une densité qui est :

$$\forall x = (x_1, x_2) \in \mathbf{R}^2, \qquad p(x) = p(x_1, x_2) = \frac{1}{2\pi} \exp\left(-\frac{x_1^2 + x_2^2}{2}\right) = \frac{1}{2\pi} \exp\left(-\frac{x^t x}{2}\right).$$

Comme dans le cas de variables aléatoires réelles, la loi d'un couple X est déterminée par sa fonction caractéristique qui est la fonction de $\mathbf{R}^2$ dans $\mathbf{C}$ définie par

$$\forall \lambda \in \mathbf{R}^2, \qquad \varphi_X(\lambda) = \mathbb{E}\left[e^{i\lambda^t X}\right] = \mathbb{E}\left[e^{i\lambda_1 X_1}\, e^{i\lambda_2 X_2}\right].$$

Si X suit la loi $\mathcal{N}(0, I_2)$, l'indépendance de X_1 et X_2 conduit immédiatement à :

$$\forall \lambda \in \mathbf{R}^2, \qquad \varphi_X(\lambda) = \varphi_{X_1}(\lambda_1)\, \varphi_{X_2}(\lambda_2) = e^{-\lambda_1^2/2}\, e^{-\lambda_2^2/2} = \exp\left(-\lambda^t\lambda/2\right).$$

Définition. Un vecteur aléatoire Y est gaussien si $Y = AX + \mu$ où μ est un vecteur de $\mathbf{R}^2$, A est une matrice réelle de taille 2×2 et X un vecteur normal centré réduit.

Calculons la fonction caractéristique de Y. On a, pour $\lambda \in \mathbf{R}^2$,

$$\varphi_Y(\lambda) = \mathbb{E}\left[\exp\left(i\lambda^t(AX + \mu)\right)\right] = e^{i\lambda^t\mu}\, \mathbb{E}\left[\exp\left(i\lambda^t AX\right)\right] = e^{i\lambda^t\mu}\, \mathbb{E}\left[\exp\left(i\left(A^t\lambda\right)^t X\right)\right],$$

et par suite,

$$\varphi_Y(\lambda) = e^{i\lambda^t\mu}\, \varphi_X\left(A^t\lambda\right) = \exp\left(i\lambda^t\mu - \frac{\lambda^t AA^t\lambda}{2}\right).$$

On note Γ la matrice AA^t. Γ est une matrice symétrique, semi-définie positive. On montre facilement que Γ est la matrice de variance-covariance de Y c'est à dire

$$\Gamma = \begin{pmatrix} \mathbb{V}[Y_1] & \mathbb{C}\mathrm{ov}[Y_1, Y_2] \\ \mathbb{C}\mathrm{ov}[Y_1, Y_2] & \mathbb{V}[Y_2] \end{pmatrix},$$

où $\mathbb{C}\mathrm{ov}[Y_1, Y_2] = \mathbb{E}\left[(Y_1 - \mathbb{E}[Y_1])\,(Y_2 - \mathbb{E}[Y_2])\right]$. D'autre part le vecteur μ est simplement le vecteur des moyennes de Y i.e. $\mu = (\mathbb{E}[Y_1], \mathbb{E}[Y_2])^t$.

Comme dans le cas réel, on voit que la loi d'un couple gaussien est déterminée par la moyenne μ et la matrice de variance-covariance Γ. On dit que Y suit la loi $\mathcal{N}(\mu, \Gamma)$.

Remarque(s). Si Y est un vecteur gaussien, alors les marginales sont des gaussiennes. En effet, si $t \in \mathbf{R}$, on a, posant $\lambda = (t, 0)^t$,

$$\varphi_Y(\lambda) = \varphi_{Y_1}(t) = \exp\left(it\mu_1 - \Gamma_{1,1}t^2/2\right),$$

ce qui montre que Y_1 suit la loi $\mathcal{N}(\mu_1, \Gamma_{1,1})$.

On aimerait à présent savoir si le vecteur Y possède une densité. En fait cela n'est vrai que si la matrice Γ est inversible, ce qui revient à dire que la matrice A est inversible puisque $\det\Gamma = (\det A)^2$. Supposons donc que $\det\Gamma > 0$. Soit f une fonction borélienne et bornée de $\mathbf{R}^2$ dans $\mathbf{R}$. On a

$$\mathbb{E}[f(Y)] = \mathbb{E}[f(AX + \mu)] = \frac{1}{2\pi} \iint_{\mathbf{R}^2} f(Ax + \mu) \exp\left(-\frac{x^t x}{2}\right)\, dx_1 dx_2.$$

Effectuons le changement de variables $y = Ax + \mu = u(x)$ soit $x = A^{-1}(y - \mu) = v(y)$, pour obtenir comme $\mathrm{J}v(y) = \det\left(A^{-1}\right)$,

$$\mathbb{E}\left[f(Y)\right] = \frac{1}{2\pi} \iint_{\mathbf{R}^2} f(y) \exp\left(-\frac{\left(A^{-1}(y-\mu)\right)^t \left(A^{-1}(y-\mu)\right)}{2} \right) \left| \det\left(A^{-1}\right) \right| \, dy_1 dy_2.$$

On a $\left(A^{-1}\right)^t = \left(A^t\right)^{-1}$ et $\left(A^t\right)^{-1} A^{-1} = \left(AA^t\right)^{-1} = \Gamma^{-1}$; on en déduit tout d'abord que $\left(A^{-1}(y-\mu)\right)^t \left(A^{-1}(y-\mu\right)$ $(y-\mu)^t\Gamma^{-1}(y-\mu)$, puis notant que $\left|\det\left(A^{-1}\right)\right| = (\det\Gamma)^{-1/2}$,

$$\mathbb{E}[f(Y)] = \frac{1}{2\pi\sqrt{\det\Gamma}} \iint_{\mathbf{R}^2} f(y) \exp\left(-\frac{(y-\mu)^t\Gamma^{-1}(y-\mu)}{2} \right) dy_1 dy_2.$$

Y admet donc pour densité la fonction p_Y définie par

$$\forall y \in \mathbf{R}^2, \qquad p_Y(y) = \frac{1}{2\pi\sqrt{\det\Gamma}} \exp\left(-\frac{(y-\mu)^t\Gamma^{-1}(y-\mu)}{2} \right).$$