Statistique

Corrigé du partiel du 27 mars 2002

Questionnaire à choix multiples

1. Si X et Y sont indépendants, $\text{var}(2X+3Y)$ vaut

   ○ $2\text{var}(X)+3\text{var}(Y)$ ○ $4\text{var}(X)+9\text{var}(Y)+12\text{var}(XY)$
   ⊗ $4\text{var}(X)+9\text{var}(Y)$ ○ $6\mathbb{E}[XY]$ ○ $12\mathbb{E}[XY]$

2. Si $F(x_1,x_2)$ est la fonction de répartition et $f(x_1,x_2)$ est la fonction densité du couple (X_1,X_2), ces deux fonctions sont reliées par

   ○ $f(x_1,x_2)=F(x_1,1)+F(1,x_2)$ ⊗ $f(x_1,x_2)=\frac{d}{dx_1}\left[\frac{dF(x_1,x_2)}{dx_2}\right]$
   ○ $f(x_1,x_2)=\int\int F(x_1,x_2)\,dx_1\,dx_2$ ○ $f(x_1,x_2)=\frac{dF(x_1,x_2)}{dx_1}\frac{dF(x_1,x_2)}{dx_2}$

3. La covariance $\text{cov}(X,Y)$ vérifie l'inégalité

   ○ $\text{cov}(X,Y)\geq|\text{var}(X)-\text{var}(Y)|$ ○ $|\text{cov}(X,Y)|\leq\min(\text{var}(X),\text{var}(Y))$
   ○ $|\text{cov}(X,Y)|\leq\text{var}(X)\text{var}(Y)$ ⊗ $2\text{cov}(X,Y)\leq\text{var}(X)+\text{var}(Y)$

4. La covariance $\text{cov}(X+Y,X-Y)$ est égale à

   ○ $\text{cov}(X,Y)$ ⊗ $\text{var}(X)-\text{var}(Y)$
   ○ $\text{var}(X)+\text{var}(Y)$ ○ $2(\text{var}(X)+\text{var}(Y))$

5. Si $X\sim\mathcal{N}_p(\mu,\Sigma)$, si B est une matrice (r,p), la loi de BX est

   ○ $\mathcal{N}_r(B\mu,\Sigma)$ ○ $\mathcal{N}_p(\mu B,\Sigma\,BB^T)$
   ○ $\mathcal{N}_r(B\mu,\Sigma+BB^T)$ ⊗ $\mathcal{N}_r(B\mu,B\Sigma B^T)$

EXERCICE 1

1. *$f(x_1,x_2)$ est bien intégrable* car, si on considère en premier l'intégrale en x_2 la fonction est intégrable en 0 [car bornée par σ^{-4}] et en $+\infty$ [car bornée par x_2^{-4}]. Ayant intégré en x_2, la fonction en x_1 est d'ordre x_1^{-3}, intégrable en $\pm\infty$, et la bornitude règle le problème en $x_1 = 0$.

2. Par un simple changement de variables,

$$\int\int \frac{1}{(x_1^2+(x_2+\sigma)^2)^2} dx_1 dx_2 = \sigma^{-2} \int\int \frac{1}{(y_1^2+(y_2+1)^2)^2} dy_1 dy_2$$

donc k *est de la forme* $\sigma^2 k'$.

3. On a

$$\begin{aligned}
\int_{-\infty}^{\infty} \frac{(y_2+1)^2}{(y_1^2+(y_2+1)^2)^2} dy_1 &= \int_{-\infty}^{\infty} \frac{1}{y_1^2+(y_2+1)^2} dy_1 - \int_{-\infty}^{\infty} \frac{y_1^2}{(y_1^2+(y_2+1)^2)^2} dy_1 \\
&= \int_{-\infty}^{\infty} \frac{1}{y_1^2+(y_2+1)^2} dy_1 - \frac{1}{2}\int_{-\infty}^{\infty} \frac{1}{y_1^2+(y_2+1)^2} dy_1 \\
&= \frac{1}{2}\frac{1}{(y_2+1)}\int_{-\infty}^{\infty} \frac{1}{y_1^2+1} dy_1 \\
&= \frac{\pi}{2}\frac{1}{(y_2+1)}
\end{aligned}$$

donc

$$\int\int \frac{\mathbb{I}_{y_2>0}}{(y_1^2+(y_2+1)^2)^2} dy_1 dy_2 = \frac{\pi}{2}\int_0^{\infty} \frac{1}{(y_2+1)^3} dy_2$$

4. Par conséquent,

$$\int\int \frac{\mathbb{I}_{y_2>0}}{(y_1^2+(y_2+1)^2)^2} dy_1 dy_2 = \frac{\pi}{2}\left[\frac{-1}{2(y_2+1)^2}\right]_0^{\infty} = \frac{\pi}{4}$$

et, par inversion, $k' = 4/\pi$ et $k = 4\sigma^{-2}/\pi$.

5. Par le même changement de variables,

$$\int\int \frac{x_2}{(x_1^2+(x_2+\sigma)^2)^2} dx_1 dx_2 = \sigma^{-1} \int\int \frac{y_2}{(y_1^2+(y_2+1)^2)^2} dy_1 dy_2$$

6. Pour les mêmes raisons que ci-dessus, intégrant d'abord en y_1,

$$\int\int \frac{y_2}{(y_1^2+(y_2+1)^2)^2}\mathbb{I}_{y_2>0}\, dy_1 dy_2 = \frac{\pi}{2}\int_0^\infty \frac{y_2}{(y_2+1)^3}dy_2$$

Et

$$\int_0^\infty \frac{y_2}{(y_2+1)^3}dy_2 = \left[\frac{-1}{(y_2+1)}\right]_0^\infty - \left[\frac{-1}{2(y_2+1)^2}\right]_0^\infty = 1-\frac{1}{2}=\frac{1}{2}$$

7. Le résultat précédent montre que $\mathbb{E}[X_2]=\sigma$. De la symétrie de $f(x_1,x_2)$ en x_1, on déduit que $\mathbb{E}[X_1]=0$.

8. *La matrice de variance-covariance du vecteur* (X_1,X_2) n'existe pas car

$$\int\int \frac{\mathbb{I}_{y_2>0}}{(y_1^2+(y_2+1)^2)}dy_1dy_2\,,$$

formellement proportionnelle à $\mathbb{E}[X_1^2]+\mathbb{E}[(X_2+1)^2]$, n'est pas intégrable.

9. X_1 *et* X_2 *ne sont pas indépendants* car la densité jointe ne factorise pas. Par contre, X_1 *et* X_2 *ne sont pas corrélés* car

$$\int\int \frac{x_1x_2\mathbb{I}_{x_2>0}}{(x_1^2+(x_2+1)^2)}\,dx_1dx_2 = \int_0^\infty x_2 \int_{-\infty}^\infty \frac{x_1}{(x_1^2+(x_2+1)^2)}dx_1dx_2 = 0$$

Exercice 2

1. La loi marginale de X_1 est encore une loi normale [cf. cours] de paramètres

$$\mathbb{E}[\mathbb{E}[X_1|X_2]] = \mu_1 + \alpha\mathbb{E}[X_2-\mu_2] = \mu_1$$

et

$$\text{var}(\mathbb{E}[X_1|X_2]) + \mathbb{E}[\text{var}(X_1|X_2)] = \alpha^2\sigma_2^2+\beta$$

2. La loi jointe de (X_1,X_2) s'obtient comme produit de la loi marginale (de X_2) et de la loi conditionnelle (de X_1), donc préserve le format

exponentiel d'un polynôme du second degré, qui est le propre de la loi normale. On a ainsi

$$\begin{aligned} f(x_1,x_2) &= \frac{1}{2\pi\sqrt{\beta}\sigma_2}\exp\left\{\frac{-1}{2}\left(\beta^{-1}(x_1-\mu_1-\alpha(x_2-\mu_2))^2+\sigma_2^{-2}(x_2-\mu_2)^2\right)\right\} \\ &= \frac{1}{2\pi\sqrt{\beta}\sigma_2}\exp\left\{\frac{-1}{2}\left(\beta^{-1}(x_1-\mu_1)^2-2\beta^{-1}\alpha(x_1-\mu_1)(x_2-\mu_2)\right.\right. \\ &\qquad \left.\left.+(\sigma_2^{-2}+\beta^{-1}\alpha^2)(x_2-\mu_2)^2\right)\right\} \end{aligned}$$

Donc les paramètres de la loi $\mathcal{N}(\mu,\Sigma)$ sont

$$\mu=(\mu_1,\mu_2) \qquad \text{et} \qquad \Sigma^{-1}=\begin{pmatrix}\beta^{-1} & -\beta^{-1}\alpha \\ -\beta^{-1}\alpha & (\sigma_2^{-2}+\beta^{-1}\alpha^2)\end{pmatrix}$$

Par conséquent,

$$\begin{aligned} \Sigma &= \frac{1}{\beta^{-1}(\sigma_2^{-2}+\beta^{-1}\alpha^2)-\beta^{-2}\alpha^2}\begin{pmatrix}(\sigma_2^{-2}+\beta^{-1}\alpha^2) & \beta^{-1}\alpha \\ \beta^{-1}\alpha & \beta^{-1}\end{pmatrix} \\ &= \beta\sigma_2^2\begin{pmatrix}(\sigma_2^{-2}+\beta^{-1}\alpha^2) & \beta^{-1}\alpha \\ \beta^{-1}\alpha & \beta^{-1}\end{pmatrix}=\begin{pmatrix}\beta+\sigma_2^2\alpha^2 & \sigma_2^2\alpha \\ \sigma_2^2\alpha & \sigma_2^2\end{pmatrix} \end{aligned}$$

On retrouve bien les termes de variance $\beta+\sigma_2^2\alpha^2$ et σ_2^2 sur la diagonale.

3. Comme

$$\mathbb{E}[X_1]=\mu_1+\alpha\left[\mathbb{E}[X_2]-\mu_2\right] \qquad \text{et} \qquad \mathbb{E}[X_2]=\mu_2+\delta\left[\mathbb{E}[X_1]-\mu_1\right],$$

il vient

$$\mathbb{E}[X_1]-\mu_1=\alpha\delta\left[\mathbb{E}[X_1]-\mu_1\right]$$

et *μ_1 et μ_2 sont les moyennes respectives de X_1 et X_2.*

4. Par le théorème de Bayes,

$$f(x_2)f(x_1|x_2)=f(x_1)f(x_2|x_1)$$

et, donc,

$$\frac{f(x_1|x_2)}{f(x_2|x_1)}=\frac{f(x_1)}{f(x_2)}=f(x_1)\,\frac{1}{f(x_2)},$$

produit de la densité de x_1 et d'une fonction de x_2.

5. Partant des densités des deux lois (1), on peut simplifier en posant $\mu_1 = \mu_2 = 0$ et

$$
\begin{aligned}
\frac{f(x_1|x_2)}{f(x_2|x_1)} &= \sqrt{\frac{\gamma}{\beta}} \exp \frac{1}{2} \left\{ \gamma^{-1}(x_2 - \delta x_1)^2 - \beta^{-1}(x_1 - \alpha x_2)^2 \right\} \\
&= \sqrt{\frac{\gamma}{\beta}} \exp \frac{1}{2} \left\{ (\gamma^{-1} - \beta^{-1}\alpha^2) x_2^2 - (\beta^{-1} - \gamma^{-1}\delta^2) x_1^2 \right. \\
&\qquad \left. + 2(\beta^{-1}\alpha - \gamma^{-1}\delta) x_1 x_2 \right\} .
\end{aligned}
$$

Par conséquent, ce rapport factorise en une fonction de x_1 par une fonction de x_2 si, et seulement si, le coefficient de $x_1 x_2$ est nul, soit

$$\beta^{-1}\alpha = \gamma^{-1}\delta \,.$$

Il donne de plus une fonction intégrable en x_1 si, et seulement si, le coefficient de x_1^2 est négatif, soit

$$\beta^{-1} \geq \gamma^{-1}\delta^2 \,.$$

Par symétrie, on doit aussi avoir

$$\gamma^{-1} \geq \beta^{-1}\alpha^2 \,,$$

mais cette condition est automatiquement vérifiée, étant donnée les deux premières.

6. *Les deux lois conditionnelles correspondent à une loi jointe normale bidimensionnelle* $\mathcal{N}_2(\mu, \Sigma)$ puisque la fonction de x_1 exhibée dans la question précédente est

$$\exp \left\{ \frac{-1}{2} x_1^2 (\beta^{-1} - \gamma^{-1}\delta^2) \right\} ,$$

proportionnelle à la densité de la loi normale

$$\mathcal{N}(0, (\beta^{-1} - \gamma^{-1}\delta^2)^{-1} \,.$$

On est donc ramené à la question 2.