T.P. I : Vecteurs gaussiens

Exercice 1. Simulation de la loi normale par la méthode de Box-Muller.

Soient X et Y deux variables indépendantes de loi normales centrées réduites.

1. Déterminer la loi du couple (X, Y). Puis montrer que la densité du couple (R^2, Θ), où $R^2 = X^2 + Y^2$ et $\Theta = \tan^{-1}(Y/X)$, est donnée par

$$f(d, \theta) = \frac{1}{4\pi} e^{-d/2}, \quad 0 < d < \infty, 0 < \theta < 2\pi$$

2. Montrer que les variables aléatoires R^2 et Θ sont indépendantes de lois respectivement exponentielle $\mathcal{E}(1/2)$ et uniforme sur $]0, 2\pi[$.

3. Réciproquement, en déduire que si U et V sont deux variables aléatoires indépendantes de loi uniformes sur $]0, 1[$ alors les variables aléatoires

$$X = \sqrt{-2\ln U}\cos(2\pi V) \quad Y = \sqrt{-2\ln U}\sin(2\pi V)$$

sont indépendantes de loi $\mathcal{N}(0, 1)$.

En déduire une méthode pour simuler un échantillon de la loi normale centrée réduite.

Vérifier que les variables ainsi obtenues ont bien les propriétés voulues (histogramme, fonction de répartition,plot2d).

Exercice 2.

Soient U_1, U_2 et U_3 trois variables indépendantes de loi uniforme sur $[0, 1]$. En utilisant les résultats précédents montrer que

$$X = \sqrt{-\frac{2\ln U_3}{S}}V_1 \quad Y = \sqrt{-\frac{2\ln U_3}{S}}V_2$$

avec $V_1 = 2U_1 - 1$, $V_2 = 2U_2 - 1$ et $S = V_1^2 + V_2^2$, sont indépendantes de loi $\mathcal{N}(0, 1)$.

En déduire une nouvelle méthode pour simuler un échantillon de la loi normale centrée réduite.

Exercice 3.

On garde les mêmes notations que précédemment. Montrer que si $\begin{pmatrix} X \\ Y \end{pmatrix}$ est un vecteur gaussien centré, réduit, alors

$$\begin{pmatrix} 1 & 0 \\ \rho & \sqrt{1-\rho^2} \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix}$$

est un vecteur gaussien centré de matrice de covariance $\begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$.

Utiliser dans Scilab la fonction *chol* en la testant sur une matrice 2x2.

Ecrire une fonction *vecgaus* de paramètre ρ qui renvoie un échantillon d'un vecteur gaussien et qui en donne une représentation graphique à l'aide de *plot2d*.

Procéder de même pour un vecteur gaussien $V = \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix}$ d'espérance $\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}$ et de covariance :

$$\begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho \\ \sigma_1\sigma_2\rho & \sigma_2^2 \end{pmatrix}$$

où $\sigma_i^2 = \mathrm{Var}(Y_i)$ et $\rho = \mathrm{cor}(Y_1, Y_2) = \mathrm{cov}(Y_1, Y_2)/\sigma_1\sigma_2$. On pourra utiliser

$$\begin{pmatrix} \sigma_1 & 0 \\ \sigma_2\rho & \sigma_2\sqrt{1-\rho^2} \end{pmatrix}.$$

Exercice 4.

On suppose que $V = \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix}$ est un vecteur gaussien centré et de matrice de covariance :

$$\Sigma = \begin{pmatrix} 2 & 3 \\ 3 & 5 \end{pmatrix}$$

On construit les variables suivante :

$$X_1 = 2Y_1 - Y_2,$$
$$X_2 = Y_1 - Y_2.$$

1. Montrer que X_1 et X_2 sont indépendantes et de même loi $\mathcal{N}(0,1)$.

2. Simuler un $n-$échantillon de V.

3. On suppose que Y_1 et Y_2 représente la résultat d'une expérience. Cette expérience est dite faussée si le premier résultat est inférieur à 0.4 et le deuxième est supérieur à 3. Calculer à l'aide d'une simulation la probabilité que l'expérience soit faussée.

4. Après la mise en place d'un autre procédé, les lois des variables Y_1 et Y_2 ont été un peu modifiée : V a maintenant pour moyenne $(-0.4, 2)$ et pour matrice de covariance Σ. Simuler un $n-$échantillon et comparer avec le rśultat de la question précédente.

Exercice 5. Loi du chi-deux et illustration du théorème de Cochran.

1. Pour $m = 2$, puis $m = 4$, simuler m échantillons de taille $n = 10000$ de la loi normale $\mathcal{N}(0,1)$. Ces échantillons seront placés dans une matrice X à n lignes et m colonnes. Calculer l'échantillon Y obtenu en calculant la somme des carrées de chaque ligne de cette matrice.

Représenter un histogramme des valeurs de Y. Superposer sur le même graphique la densité de la loi du chi-deux à m degrés de liberté (*dchisq*).

Représenter la fonction de répartition empirique de Y. Superposer sur le même graphique la fonction de répartition de la loi du chi- deux à m degrés de liberté (*pchisq*).

2. Calculer la moyenne empiriques M de chacune des lignes de X, retrancher M à chacune des colonnes de X et calculer l'échantillon Z obtenu en prenant la somme des carrées de chaque ligne.

Représenter un histogramme des valeurs de Z. Superposer sur le même graphique la densité de la loi du chi-deux à $m - 1$ degrés de liberté. Représenter la fonction de répartition empirique de Z. Superposer sur le même graphique la fonction de répartition de la loi du chi-deux à $m - 1$ degrés de liberté.

Justifier que le vecteur Z peut également être obtenu par la commande $Z = Y - m * M^2$.