

MAP 311 - Aléatoire

Leçon 4

2016-2017

Vecteurs aléatoires indépendants

Définition

Soient \mathbf{Z} un vecteur aléatoire à valeurs dans \mathbb{R}^n et \mathbf{V} un vecteur aléatoire à valeurs dans \mathbb{R}^m .

\mathbf{Z} et \mathbf{V} sont indépendants si pour tous boréliens A et B de \mathbb{R}^n et \mathbb{R}^m , on a

$$\mathbb{P}(\mathbf{Z} \in A, \mathbf{V} \in B) = \mathbb{P}(\mathbf{Z} \in A)\mathbb{P}(\mathbf{V} \in B).$$

Il suffit de vérifier

$$\begin{aligned} & \mathbb{P}\left(\mathbf{Z} \in \prod_{i=1}^n]-\infty, z_i], \mathbf{V} \in \prod_{j=1}^m]-\infty, v_j]\right) \\ &= \mathbb{P}\left(\mathbf{Z} \in \prod_{i=1}^n]-\infty, z_i]\right)\mathbb{P}\left(\mathbf{V} \in \prod_{j=1}^m]-\infty, v_j]\right) \end{aligned}$$

car la fonction de répartition caractérise la loi.

Vecteurs aléatoires indépendants à densité

Proposition

Si \mathbf{Z} et \mathbf{V} ont des densités, alors \mathbf{Z} et \mathbf{V} sont indépendants si et seulement si pour (presque) tous $(z, v) \in \mathbb{R}^n \times \mathbb{R}^m$,

$$f_{\mathbf{Z}, \mathbf{V}}(z, v) = f_{\mathbf{Z}}(z)f_{\mathbf{V}}(v)$$

Preuve : Si la densité jointe se met sous forme produit, alors

$$\begin{aligned}\mathbb{P}(\mathbf{Z} \in A, \mathbf{V} \in B) &= \iint_{A \times B} f_{\mathbf{Z}, \mathbf{V}}(z, v) dz dv \\ &= \int_A f_{\mathbf{Z}}(z) dz \int_B f_{\mathbf{V}}(v) dv = \mathbb{P}(\mathbf{Z} \in A)\mathbb{P}(\mathbf{V} \in B)\end{aligned}$$

Donc \mathbf{Z} et \mathbf{V} sont indépendantes.

Vecteurs aléatoires indépendants à densité

Proposition

Si Z et V ont des densités, alors Z et V sont indépendants si et seulement si pour (presque) tous $(z, v) \in \mathbb{R}^n \times \mathbb{R}^m$,

$$f_{Z,V}(z, v) = f_Z(z)f_V(v)$$

Preuve de la réciproque (pour $n = m = 1$) : Si Z et V sont à densité et indépendants, alors en posant $f(z, v) = f_Z(z)f_V(v)$, le couple (X, Y) de densité f a pour fonction de répartition

$$\begin{aligned}\mathbb{P}(X \leq x, Y \leq y) &= \iint_{]-\infty, x] \times]-\infty, y]} f(z, v) dz dv \\ &= \int_{-\infty}^x f_Z(z) dz \int_{-\infty}^y f_V(v) dv \\ &= \mathbb{P}(Z \leq x) \mathbb{P}(V \leq y) = \mathbb{P}(Z \leq x, V \leq y)\end{aligned}$$

Donc (X, Y) et (Z, V) ont même loi.

Proposition

Soient $\mathbf{Z} \in \mathbb{R}^n$ et $\mathbf{V} \in \mathbb{R}^m$ des vecteurs aléatoires indépendants. Soient g et h mesurables positives ou bornées sur \mathbb{R}^n et \mathbb{R}^m , alors

$$\mathbb{E}(g(\mathbf{Z})h(\mathbf{V})) = \mathbb{E}(g(\mathbf{Z}))\mathbb{E}(h(\mathbf{V}))$$

Preuve :

- 1) Cas général: vrai pour g et h fonctions indicatrices d'événements, donc vrai pour les fonctions étagées. Vrai pour les fonctions positives par passage à la limite puis on écrit $g = g^+ - g^-$ et $h = h^+ - h^-$.
- 2) Si \mathbf{Z} et \mathbf{V} ont des lois à densité, preuve plus simple :

$$\begin{aligned}\mathbb{E}(g(\mathbf{Z})h(\mathbf{V})) &= \iint_{\mathbb{R}^n \times \mathbb{R}^m} g(\mathbf{z})h(\mathbf{v})f(\mathbf{z}, \mathbf{v})d\mathbf{z}d\mathbf{v} \\ &= \int_{\mathbb{R}^n} g(\mathbf{z})f_{\mathbf{Z}}(\mathbf{z})d\mathbf{z} \int_{\mathbb{R}^m} h(\mathbf{v})f_{\mathbf{V}}(\mathbf{v})d\mathbf{v} \\ &= \mathbb{E}(g(\mathbf{Z}))\mathbb{E}(h(\mathbf{V})) \text{ (par théorème de Fubini)}\end{aligned}$$

Proposition

Si X et Y sont deux variables aléatoires réelles indépendantes, alors

$$\text{Cov}(X, Y) = 0$$

Preuve immédiate :

$$\mathbb{E}((X - \mathbb{E}(X))(Y - \mathbb{E}(Y))) = \mathbb{E}(X - \mathbb{E}(X))\mathbb{E}(Y - \mathbb{E}(Y)) = 0$$

Attention : Réciproque fausse. Voir l'exemple des fléchettes.

$\text{Cov}(X, Y) = 0$ mais X et Y ne sont pas indépendantes :

$$\mathbb{P}(X^2 \geq 3/4, Y^2 \geq 3/4) = 0 < \mathbb{P}(X^2 \geq 3/4)\mathbb{P}(Y^2 \geq 3/4)$$

Suite de variables aléatoires indépendantes

Définition

1) Les v.a. X_1, \dots, X_n sont indépendantes si pour tous boréliens A_1, \dots, A_n ,

$$\mathbb{P}(X_1 \in A_1, \dots, X_n \in A_n) = \prod_{i=1}^n \mathbb{P}(X_i \in A_i)$$

2) La suite $(X_n)_{n \in \mathbb{N}}$ est dite indépendante si pour tout n , la famille finie X_1, \dots, X_n est indépendante.

Méthode de simulation dite “du rejet”

Objectif : Simuler une variable aléatoire de densité $f(y)$, mais on ne connaît pas ou on ne sait pas inverser sa fonction de répartition F .

- On suppose que :
 - $f(y) \leq ag(y)$, avec g densité de probabilité et a constante (on a nécessairement $a \geq 1$. Pourquoi ?).
 - On sait simuler des variables aléatoires de densité g .
- On dispose de :
 - une suite X_1, \dots, X_n, \dots de variables aléatoires indépendantes de loi $\mathcal{U}(0, 1)$.
 - une suite Z_1, \dots, Z_n, \dots , de variables aléatoires indépendantes de densité g , indépendantes des $(X_n)_{n \in \mathbb{N}}$.

Pour chaque n , (X_n, Z_n) a pour densité jointe $\mathbf{1}_{[0,1]}(x)g(z)$.

Méthode de simulation dite “du rejet”

On considère le premier temps (aléatoire) N tel que $f(Z_n) > ag(Z_n)X_n$:

$$N = \inf\{n \in \mathbb{N}^* , f(Z_n) > ag(Z_n)X_n\}.$$

On pose $Y = Z_n$ si $N = n$.

Intéressons-nous à la loi du couple de variables aléatoires (N, Y) à valeurs dans $\mathbb{N}^* \times \mathbb{R}$.

Proposition

La variable aléatoire N suit une loi géométrique de paramètre $1/a$ (et donc d'espérance a).

La variable aléatoire Y a pour densité f .

N et Y sont indépendantes.

Application : Supposons que f est bornée par C et est à support borné dans $[b, c]$.

On peut prendre $g(z) = \frac{1}{c-b} \mathbf{1}_{[b,c]}(z)$, densité de la loi uniforme sur $[b, c]$ et $a = C(c - b)$.

Preuve : on a bien $f(z) \leq ag(z)$ pour tout z .

Méthode du rejet : On tire (X_n, Z_n) de loi uniforme sur $[0, 1] \times [b, c]$, jusqu'à ce que $f(Z_n) > ag(Z_n)X_n$, avec $ag(Z_n) = C$. Et alors on pose $Y = Z_n$.

De manière équivalente : On tire (\tilde{X}_n, Z_n) de loi uniforme sur $[0, C] \times [b, c]$, jusqu'à ce que $f(Z_n) > \tilde{X}_n$. Et alors on pose $Y = Z_n$.

Somme de variables aléatoires indépendantes

Proposition

1) Soient X_1, \dots, X_n des variables aléatoires de carré intégrable. Alors

$$\text{Var}\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n \text{Var}(X_i) + 2 \sum_{1 \leq i < j \leq n} \text{Cov}(X_i, X_j)$$

2) Si de plus, les X_i sont indépendantes, alors

$$\text{Var}\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n \text{Var}(X_i)$$

Preuve :

1) On utilise que $(X, Y) \mapsto \text{Cov}(X, Y)$ est une forme bilinéaire sur $L^2 \times L^2$.

2) X et Y indépendantes $\implies \text{Cov}(X, Y) = 0$.

Proposition

Si les X_i sont indépendantes et de même loi, alors

$$\text{Var}\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{\text{Var}(X_i)}{n}$$

Disparition des fluctuations aléatoires quand n tend vers l'infini !

On reviendra sur cette propriété dans le Cours 5.

- Quelle est la loi d'une somme de variables aléatoires indépendantes ?
Cas discret (variables aléatoires à valeurs dans \mathbb{Z}) : Convolution discrète.

$$\begin{aligned}\mathbb{P}(X + Y = i) &= \sum_{j \in \mathbb{Z}} \mathbb{P}(X = j) \mathbb{P}(Y = i - j) \\ &= \sum_{j \in \mathbb{Z}} \mathbb{P}(X = i - j) \mathbb{P}(Y = j) \quad \forall i \in \mathbb{Z}\end{aligned}$$

Somme de v.a. indépendantes et convolution

Proposition

Soient X et Y deux v.a. réelles indépendantes, de densités respectives f_X et f_Y . Alors $Z = X + Y$ admet la densité f_Z donnée par :

$$f_Z(z) = \int_{\mathbb{R}} f_X(z - y)f_Y(y)dy = \int_{\mathbb{R}} f_X(x)f_Y(z - x)dx$$

f_Z est appelée produit de convolution de f_X et f_Y et est souvent notée $f_X * f_Y$.

Preuve : Soit g une fonction continue bornée.

$$\begin{aligned}\mathbb{E}(g(Z)) &= \int_{\mathbb{R}} \int_{\mathbb{R}} g(x + y)f_X(x)f_Y(y)dxdy \\ &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} g(x + y)f_X(x)dx \right) f_Y(y)dy = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} g(z)f_X(z - y)dz \right) f_Y(y)dy \\ &= \int_{\mathbb{R}} g(z) \left(\int_{\mathbb{R}} f_X(z - y)f_Y(y)dy \right) dz\end{aligned}$$

Application : La somme de deux variables aléatoires normales indépendantes de lois normales $\mathcal{N}(m, s^2)$ et $\mathcal{N}(\mu, \sigma^2)$ est une variable aléatoire normale de loi $\mathcal{N}(m + \mu, s^2 + \sigma^2)$.

Preuve : calcul.

On verra un résultat général sur la stabilité de la loi gaussienne (Cours 5).

Calcul des lois de vecteurs aléatoires

Proposition

Soit \mathbf{V} un vecteur aléatoire n -dimensionnel. Si pour toute fonction $h : \mathbb{R}^n \rightarrow \mathbb{R}$ continue bornée

$$\mathbb{E}(h(\mathbf{V})) = \int_{\mathbb{R}^n} h(\mathbf{v})f(\mathbf{v})d\mathbf{v}$$

alors la loi de \mathbf{V} admet la densité f .

(méthode de la fonction muette)

Question : Soit $\mathbf{U} = (U_1, \dots, U_n)$ vecteur aléatoire de densité $f_{\mathbf{U}}$. Soit $\mathbf{g} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ et $\mathbf{V} = \mathbf{g}(\mathbf{U})$. Quelle est la loi de \mathbf{V} ?

Nous avons pour $h : \mathbb{R}^n \rightarrow \mathbb{R}$ continue bornée,

$$\mathbb{E}(h(\mathbf{V})) = \int_{\mathbb{R}^n} h(\mathbf{v}) P_{\mathbf{V}}(d\mathbf{v}) = \int_{\mathbb{R}^n} h(\mathbf{g}(\mathbf{u})) f_{\mathbf{U}}(\mathbf{u}) d\mathbf{u}$$

Utilisons la formule du changement de variable dans \mathbb{R}^n . Soient $\mathbf{u} = (u_1, \dots, u_n)$ et $\mathbf{v} = \mathbf{g}(\mathbf{u})$.

Proposition

Soit \mathbf{g} difféomorphisme de A sur $B = \mathbf{g}(A)$: \mathbf{g} bijective, \mathbf{g} et \mathbf{g}^{-1} continûment différentiables. Soit

$$J(\mathbf{g})(\mathbf{u}) = \det \left(\left(\frac{\partial g_i}{\partial u_j}(\mathbf{u}) \right)_{i,j=1}^n \right)$$

le jacobien de \mathbf{g} au point \mathbf{u} . Alors on

$$\int_A h \circ \mathbf{g}(\mathbf{u}) f_{\mathbf{U}}(\mathbf{u}) d\mathbf{u} = \int_B h(\mathbf{v}) f_{\mathbf{U}} \circ \mathbf{g}^{-1}(\mathbf{v}) |J(\mathbf{g}^{-1})(\mathbf{v})| d\mathbf{v}$$

Proposition

Si de plus la densité de \mathbf{U} est portée par A : $\int_A f_{\mathbf{U}}(\mathbf{u})d\mathbf{u} = 1$, alors $\mathbf{V} = \mathbf{g}(\mathbf{U})$ a pour densité

$$f_{\mathbf{V}}(\mathbf{v}) = f_{\mathbf{U}} \circ \mathbf{g}^{-1}(\mathbf{v})|J(\mathbf{g}^{-1})(\mathbf{v})|\mathbf{1}_B(\mathbf{v})$$

Rappel : $\mathbf{g} \circ \mathbf{g}^{-1} = \mathbf{I}$ donc $J(\mathbf{g}^{-1})(\mathbf{v}) = \frac{1}{J(\mathbf{g})(\mathbf{g}^{-1}(\mathbf{v}))}$

Lois exponentielles et loi Gamma

Proposition

Une somme S_n de n variables aléatoires indépendantes de loi exponentielle $\mathcal{E}(\beta)$ a une loi de densité

$$f_{S_n}(s) = \frac{\beta^n}{(n-1)!} s^{n-1} e^{-\beta s} \mathbf{1}_{]0, +\infty[}(s)$$

Preuve : récurrence + calcul de la convolution par intégration par parties.
Plus généralement,

Définition

X suit la loi Gamma de paramètres $\alpha > 0$ et $\beta > 0$ si X est à valeurs dans \mathbb{R}^+ et P_X admet la densité :

$$f(x) = \frac{1}{\Gamma(\alpha)} \beta^\alpha x^{\alpha-1} e^{-\beta x} \mathbf{1}_{]0, +\infty[}(x)$$

On note $X \sim \Gamma(\alpha, \beta)$.

Ici $\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{-x} dx$ (fonction Gamma d'Euler).

Exercice : Vérifier que $\int_0^{\infty} f(x) dx = 1$.

Proposition

La loi d'une somme de deux v.a. X et Y indépendantes de loi $\Gamma(a, \theta)$ et $\Gamma(b, \theta)$ est une loi $\Gamma(a + b, \theta)$.

Indication de preuve :

- 1) si a et b sont entiers : facile en utilisant la représentation avec des v.a. exponentielles.
- 2) cas général : on étudie la loi du couple $(X + Y, X/(X + Y))$.

Loi du χ^2

Définition

On appelle loi du chi-deux à n degrés de liberté et on note $\chi^2(n)$ la loi de

$$X_1^2 + \dots + X_n^2,$$

où X_1, \dots, X_n sont n variables aléatoires normales centrées réduites, indépendantes. C'est la loi $\Gamma(\frac{n}{2}, \frac{1}{2})$.

Preuve : Nous avons vu dans le Cours 3 que la loi de $\chi^2(1)$ est une loi $\Gamma(\frac{1}{2}, \frac{1}{2})$. Par la proposition précédente, nous en déduisons la loi $\chi^2(n)$. Cette loi interviendra dans les cours 7-9, en statistique.

Coordonnées polaires

- Soit $U = (X, Y)$ vecteur aléatoire à valeurs dans \mathbb{R}^2 de densité f_U et $V = (R, \Theta)$ ses coordonnées polaires.
- Loi de V : Introduisons $g : (x, y) \rightarrow (r, \theta)$ difféomorphisme de $\mathbb{R}^2 \setminus \{(0, 0)\}$ dans $]0, +\infty[\times]0, 2\pi[$, avec

$$g^{-1}(r, \theta) = (r \cos \theta, r \sin \theta) \text{ et } J(g^{-1})(r, \theta) = r.$$

Le vecteur aléatoire V a pour densité

$$f_V(r, \theta) = f_U(r \cos \theta, r \sin \theta) r \mathbf{1}_{]0, +\infty[}(r) \mathbf{1}_{]0, 2\pi[}(\theta)$$

- Considérons $U = (X, Y)$ où X et Y indépendantes de loi $\mathcal{N}(0, 1)$.
Quelle est la loi de V ? Quelles sont les lois marginales de R et de Θ ? la loi de R^2 ?

- Considérons $U = (X, Y)$ où X et Y indépendantes de loi $\mathcal{N}(0, 1)$.
Quelle est la loi de V ? Quelles sont les lois marginales de R et de Θ ? la loi de R^2 ?

Réponse :

$$f_V(r, \theta) = \frac{1}{2\pi} \exp\left(-\frac{r^2}{2}\right) r \mathbf{1}_{]0, +\infty[}(r) \mathbf{1}_{[0, 2\pi[}(\theta)$$

Donc R et Θ sont indépendantes, Θ de loi $\mathcal{U}(0, 2\pi)$ et R de loi a densité

$$f_R(r) = \exp\left(-\frac{r^2}{2}\right) r \mathbf{1}_{]0, +\infty[}(r)$$

R^2 a pour densité

$$f_{R^2}(\rho) = \frac{1}{2} \exp\left(-\frac{\rho}{2}\right) \mathbf{1}_{]0, +\infty[}(\rho)$$

On reconnaît la loi exponentielle de paramètre $1/2$.

Simulation de variable aléatoire normale

- Nous allons simuler un vecteur (X, Y) de variables aléatoires normales centrées réduites indépendantes.
- On sait que

$$(X, Y) = (R \cos \Theta, R \sin \Theta) = (\sqrt{E} \cos \Theta, \sqrt{E} \sin \Theta)$$

où E et Θ sont indépendantes et suivent respectivement une loi exponentielle de paramètre $\frac{1}{2}$ et une loi uniforme sur $[0, 2\pi[$.

- Nous simulons deux variables aléatoires U_1 et U_2 indépendantes de loi uniforme sur $[0, 1]$. Alors

$$(\sqrt{-2 \log U_1} \cos(2\pi U_2), \sqrt{-2 \log U_1} \sin(2\pi U_2))$$

a même loi que (X, Y) .

- On en déduit facilement par changement de variable que

$$\sigma \sqrt{-2 \log U_1} \cos(2\pi U_2) + m$$

a pour loi $\mathcal{N}(m, \sigma^2)$.

(Méthode de Box-Müller)

Vecteur gaussien n -dimensionnel

- Si Z_1, \dots, Z_n sont des variables aléatoires indépendantes normales centrées réduites, alors la loi de $\mathbf{Z} = (Z_1, \dots, Z_n)$ a pour densité le produit des densités

$$f_{\mathbf{Z}}(\mathbf{z}) = \frac{1}{(2\pi)^{n/2}} \exp\left(-\frac{\sum_{i=1}^n z_i^2}{2}\right) = \frac{1}{(2\pi)^{n/2}} \exp\left(-\frac{\|\mathbf{z}\|^2}{2}\right)$$

- Soient $\mathbf{m} \in \mathbb{R}^n$ et \mathbf{C} une matrice de taille $n \times n$ symétrique définie positive.

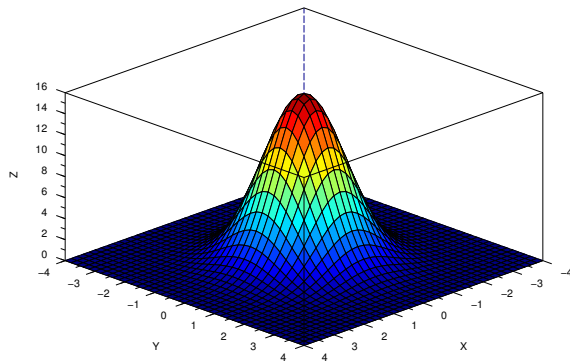
Définition

\mathbf{Z} suit une loi gaussienne de moyenne \mathbf{m} et de matrice de covariance \mathbf{C} si sa densité a la forme

$$f_{\mathbf{Z}}(\mathbf{z}) = \frac{1}{(2\pi)^{n/2} \sqrt{\det \mathbf{C}}} \exp\left(-\frac{1}{2}(\mathbf{z} - \mathbf{m})^t \mathbf{C}^{-1}(\mathbf{z} - \mathbf{m})\right)$$

On écrit $\mathbf{Z} \sim \mathcal{N}(\mathbf{m}, \mathbf{C})$.

Vecteur gaussien bi-dimensionnel



Densité gaussienne en dimension deux.

Vecteur gaussien bi-dimensionnel

Considérons un vecteur aléatoire (X, Y) de loi gaussienne tel que

$$\mathbb{E}(X) = m_X, \quad \mathbb{E}(Y) = m_Y, \quad \text{Var}(X) = \sigma_X^2, \quad \text{Var}(Y) = \sigma_Y^2,$$

et $\rho(X, Y) = \rho \notin \{-1, 1\}$. Alors

$$\mathbf{C} = \begin{pmatrix} \sigma_X^2 & \sigma_X \sigma_Y \rho \\ \sigma_X \sigma_Y \rho & \sigma_Y^2 \end{pmatrix}$$

et

$$\mathbf{C}^{-1} = \frac{1}{\sigma_X^2 \sigma_Y^2 (1 - \rho^2)} \begin{pmatrix} \sigma_Y^2 & -\sigma_X \sigma_Y \rho \\ -\sigma_X \sigma_Y \rho & \sigma_X^2 \end{pmatrix}$$

et donc la densité de (X, Y) est

$$f_{(X,Y)}(x, y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \exp \left\{ -\frac{1}{2(1-\rho^2)} \left(\left(\frac{x-m_X}{\sigma_X} \right)^2 + \left(\frac{y-m_Y}{\sigma_Y} \right)^2 - 2\rho \left(\frac{x-m_X}{\sigma_X} \right) \left(\frac{y-m_Y}{\sigma_Y} \right) \right) \right\}$$

Résultat : Soit (X, Y) un vecteur aléatoire de loi gaussienne de moyenne nulle et tel que $\text{Cov}(X, Y) = \rho\sigma_X\sigma_Y$.

Alors la loi conditionnelle de Y sachant $X = x$ est une loi normale

$$\mathcal{N}\left(\frac{\sigma_Y}{\sigma_X}\rho x, \sigma_Y^2(1 - \rho^2)\right)$$

Preuve : $f_{Y|X=x}(y) = f_{X,Y}(x, y)/f_X(x)$.

Note :

- si $\rho > 0$, la moyenne est attirée par x .
- la variance est réduite, d'autant plus que les variables sont corrélées.

De plus,

$$\mathbb{E}(Y|X) = \frac{\sigma_Y}{\sigma_X}\rho X$$

L'espérance conditionnelle est linéaire en X .

Indépendance d'un vecteur gaussien

Proposition

Si V est un vecteur gaussien, ses composantes sont indépendantes ssi sa matrice de covariance est diagonale.

Preuve : \implies Evident car alors $\text{Cov}(V_i, V_j) = 0, \forall i \neq j$.

\impliedby Si la matrice de covariance est diagonale, on a alors une forme produit de la fonction densité. On en déduit l'indépendance.

- Ce résultat peut être faux si V n'est pas gaussien : voir Cours 3. On a construit un couple de variables aléatoires (X, Y) non indépendantes telles que $\text{Cov}(X, Y) = 0$ (fléchette).

Transformation affine d'un vecteur gaussien indépendant

- Considérons un vecteur $\mathbf{U} = (U_1, \dots, U_n)$, où les coordonnées sont des v.a. indépendantes de loi normale centrées réduites.
- Soient $\mathbf{m} \in \mathbb{R}^n$ et \mathbf{A} matrice carrée $n \times n$ inversible. On considère le vecteur

$$\mathbf{V} = \mathbf{A}\mathbf{U} + \mathbf{m} = \mathbf{g}(\mathbf{U})$$

La matrice $\mathbf{C} = \mathbf{A}\mathbf{A}^t$ est symétrique définie positive. La matrice jacobienne est égale à $J(\mathbf{g}) = \mathbf{A}$, d'où $\det(J(\mathbf{g}^{-1})) = 1/\sqrt{\det \mathbf{C}}$.

De plus,

$$\text{Cov}(V_i, V_j) = \sum_{k,l=1}^n a_{ik} a_{jl} \text{Cov}(U_k, U_l) = \sum_{k=1}^n a_{ik} a_{jk} = (\mathbf{A}\mathbf{A}^t)_{ij},$$

et \mathbf{C} est la matrice de covariance de \mathbf{V} . En appliquant la formule de changement de variable, nous obtenons que \mathbf{V} a la loi de densité

$$f_{\mathbf{V}}(\mathbf{v}) = \frac{1}{(2\pi)^{n/2} \sqrt{\det \mathbf{C}}} \exp\left(-\frac{1}{2}(\mathbf{v} - \mathbf{m})^t \mathbf{C}^{-1}(\mathbf{v} - \mathbf{m})\right)$$

Transformation affine d'un vecteur gaussien

- Considérons un vecteur $\mathbf{U} = (U_1, \dots, U_n)$ gaussien de moyenne \mathbf{m}_U et de matrice de covariance \mathbf{C}_U .
- Soient $\mathbf{m} \in \mathbb{R}^n$ et \mathbf{A} matrice carrée $n \times n$ inversible. On considère le vecteur

$$\mathbf{V} = \mathbf{A}\mathbf{U} + \mathbf{m} = g(\mathbf{U})$$

En appliquant la formule de changement de variable, nous obtenons que \mathbf{V} est un vecteur gaussien de moyenne

$$\mathbf{m}_V = \mathbf{A}\mathbf{m}_U + \mathbf{m}$$

et de matrice de covariance

$$\mathbf{C}_V = \mathbf{A}\mathbf{C}_U\mathbf{A}^t$$

Morale : Une transformée affine d'un vecteur gaussien est un vecteur gaussien.

Représentation d'un vecteur gaussien

- Soit V un vecteur de loi gaussienne $\mathcal{N}(m, C)$, avec C une matrice symétrique définie positive.

Il existe une matrice orthogonale P et une matrice diagonale strictement positive D telles que $C = PDP^{-1}$.

On pose $A = PD^{-1/2}P^{-1}$.

On considère

$$U = A(V - m)$$

C'est un vecteur gaussien de moyenne 0 et de matrice de covariance $ACA^t = PD^{-1/2}P^{-1}PDP^{-1}PD^{-1/2}P^{-1} = I$. Donc $U = (U_1, \dots, U_n)$, où les coordonnées sont des v.a. indépendantes de loi normale centrée réduite, et on a

$$V = A^{-1}U + m$$

- En appliquant la formule des marginales, on peut montrer en intégrant $f_V(\mathbf{v})$ par rapport à $dv_2 \cdots dv_n$ que V_1 suit une loi normale et de même pour chaque V_i : les coordonnées de \mathbf{V} suivent des lois normales.
- Nous remarquons aussi qu'un changement de variable linéaire sur \mathbf{V} :

$$\mathbf{W} = \mathbf{B}\mathbf{V} = \mathbf{B}\mathbf{A}^{-1}\mathbf{U} + \mathbf{B}\mathbf{m}$$

donne encore la même forme de densité.

- Donc toute combinaison linéaire d'un vecteur aléatoire de loi gaussienne suit une loi normale.

Nous proposons une nouvelle définition d'un vecteur gaussien.

Définition

Un vecteur aléatoire $\mathbf{V} = (V_1, \dots, V_n)$ est appelé un vecteur gaussien si toute combinaison linéaire $\sum_{j=1}^n b_j V_j$ suit une loi normale.

Reste à montrer que cette nouvelle définition implique la première (dans le cas où la matrice de covariance est définie positive). Voir cours 6.

Attention à la nouvelle définition : Il faut réclamer que toutes les combinaisons linéaires sont normales, et pas seulement que les lois marginales sont normales.

Un vecteur de v.a. gaussiennes n'est pas forcément gaussien !

Exemple : X de loi $\mathcal{N}(0, 1)$. Alors

$$Y = X\mathbf{1}_{|X|\leq 1} - X\mathbf{1}_{|X|>1}$$

suit la loi normale $\mathcal{N}(0, 1)$. En effet, si $h \in \mathcal{C}_b(\mathbb{R})$,

$$\begin{aligned}\mathbb{E}(h(Y)) &= \mathbb{E}(h(X)\mathbf{1}_{|X|\leq 1}) + \mathbb{E}(h(-X)\mathbf{1}_{|X|>1}) \\ &= \mathbb{E}(h(X)\mathbf{1}_{|X|\leq 1}) + \mathbb{E}(h(X)\mathbf{1}_{|X|>1}) = \mathbb{E}(h(X))\end{aligned}$$

(On a utilisé la symétrie de la loi normale).

Mais (X, Y) n'est pas gaussien :

$$\mathbb{P}(X + Y = 0) = \mathbb{P}(|X| > 1) \in]0, 1[$$

Invariance de la loi gaussienne par rotation

Proposition

Si \mathbf{K} est une matrice de rotation ($\mathbf{K}^{-1} = \mathbf{K}^t$) et si $\mathbf{V} = \mathbf{K}\mathbf{U}$, avec \mathbf{U} vecteur aléatoire à composantes indépendantes de loi normale centrée réduite, alors \mathbf{V} est un vecteur aléatoire avec composantes indépendantes, de loi normale centrée réduite.

Preuve : \mathbf{V} est un vecteur gaussien de moyenne nulle et de covariance $\mathbf{K}\mathbf{K}^t = \mathbf{I}$.

Espérance conditionnelle et régression linéaire

Soit (\mathbf{X}, Y) un vecteur aléatoire, \mathbf{X} à valeurs dans \mathbb{R}^n et Y à valeurs dans \mathbb{R} .

- L'espérance de Y est la constante qui approche au mieux la variable aléatoire Y :

$$\mathbb{E}(Y) = \operatorname{argmin}_{a \in \mathbb{R}} \mathbb{E}((Y - a)^2)$$

- Espérance conditionnelle.

On observe \mathbf{X} . Quelle est la meilleure approximation de Y sachant \mathbf{X} ?

On cherche donc à résoudre :

$$\psi_0 = \operatorname{argmin}_{\psi: \mathbb{R}^n \rightarrow \mathbb{R}} \mathbb{E}((Y - \psi(\mathbf{X}))^2)$$

↔ Problème de minimisation (infini-dimensionnelle) parfois complexe.

Résultat : $\psi_0(\mathbf{x}) = \mathbb{E}[Y | \mathbf{X} = \mathbf{x}]$.

- **Régression linéaire.**

On observe \mathbf{X} . Quelle est la meilleure combinaison affine de \mathbf{X} qui approche au mieux Y ?

On cherche donc à résoudre :

$$(\alpha_j^{reg})_{j=0}^n = \underset{(\alpha_j)_{j=0}^n \in \mathbb{R}^{n+1}}{\operatorname{argmin}} \mathbb{E} \left[\left(Y - \alpha_0 - \sum_{j=1}^n \alpha_j X_j \right)^2 \right]$$

et on obtient $Y^{reg} = \alpha_0^{reg} + \sum_{j=1}^n \alpha_j^{reg} X_j$.

↪ Problème de minimisation (fini-dimensionnelle, quadratique) beaucoup plus simple que celui correspondant à l'espérance conditionnelle.

Exercice : Déterminer $(\alpha_j)_{j=0}^n$ en fonction de $\mathbb{E}(X)$, $\mathbb{E}(Y)$, $\operatorname{Cov}(X_j, Y)$, $\operatorname{Cov}(X_i, X_j)$.

Mais la régression est sous-optimale du point de vue de l'approximation : la meilleure combinaison affine de \mathbf{X} n'est pas forcément la meilleure approximation de Y sachant \mathbf{X} . Sauf que :

- **Résultat** : Si (\mathbf{X}, Y) est gaussien, alors l'espérance conditionnelle $\mathbb{E}[Y|\mathbf{X}]$ est la régression linéaire de Y sur \mathbf{X} .

Preuve de : le ψ optimal est $\psi_0(\mathbf{x}) = \mathbb{E}(Y|\mathbf{X} = \mathbf{x})$.

En notant $\tilde{\psi}(\mathbf{x}) = \psi(\mathbf{x}) - \psi_0(\mathbf{x})$, on a

$$\begin{aligned}\mathbb{E}((Y - \psi(\mathbf{X}))^2) &= \mathbb{E}((Y - \psi_0(\mathbf{X}))^2) - 2\mathbb{E}(\tilde{\psi}(\mathbf{X})(Y - \psi_0(\mathbf{X}))) \\ &\quad + \mathbb{E}(\tilde{\psi}(\mathbf{X})^2)\end{aligned}$$

Or

$$\begin{aligned}\mathbb{E}(\tilde{\psi}(\mathbf{X})\psi_0(\mathbf{X})) &= \int_{\mathbb{R}^n} \tilde{\psi}(\mathbf{x})\psi_0(\mathbf{x})f_{\mathbf{X}}(\mathbf{x})d\mathbf{x} \\ &= \int_{\mathbb{R}^n} \tilde{\psi}(\mathbf{x})\left(\int_{\mathbb{R}} yf_{Y|\mathbf{X}=\mathbf{x}}(y)dy\right)f_{\mathbf{X}}(\mathbf{x})d\mathbf{x} \\ &= \iint_{\mathbb{R}^n \times \mathbb{R}} \tilde{\psi}(\mathbf{x})yf_{\mathbf{X},Y}(\mathbf{x},y)d\mathbf{x}dy = \mathbb{E}(\tilde{\psi}(\mathbf{X})Y)\end{aligned}$$

Donc

$$\mathbb{E}((Y - \psi(\mathbf{X}))^2) = \mathbb{E}((Y - \psi_0(\mathbf{X}))^2) + \mathbb{E}(\tilde{\psi}(\mathbf{X})^2)$$

dont le minimum (en $\tilde{\psi}$) est atteint pour $\tilde{\psi} = 0$.

Preuve de la méthode de rejet

Notons $A_n = \{f(Z_n) > ag(Z_n)X_n\}$.

Les événements A_n sont indépendants et ont même probabilité :

$$\mathbb{P}(A_n) = \alpha, \text{ d'où } \mathbb{P}(N > n) = \mathbb{P}\left(\bigcap_{k=1}^n A_k^c\right) = (1 - \alpha)^n$$

Soit h une fonction continue bornée.

Comme $\{N = n\} = A_n \cap \{N > n-1\}$:

$$\begin{aligned}\mathbb{E}(h(Y)) &= \sum_{n=1}^{\infty} \mathbb{E}(h(Z_n)\mathbf{1}_{A_n}\mathbf{1}_{N>n-1}) \\ &= \sum_{n=1}^{\infty} \mathbb{E}(h(Z_n)\mathbf{1}_{A_n})(1 - \alpha)^{n-1}\end{aligned}$$

Comme (X_n, Z_n) a pour densité $\mathbf{1}_{[0,1]}(x)g(z)$:

$$\begin{aligned}\mathbb{E}(h(Z_n)\mathbf{1}_{A_n}) &= \int_{\mathbb{R}} h(z)g(z) \left(\int_0^1 \mathbf{1}_{f(z) > ag(z)x} dx \right) dz \\ &= \int_{\mathbb{R}} h(z)g(z) \frac{f(z)}{ag(z)} dz \\ &= \frac{1}{a} \int_{\mathbb{R}} h(z)f(z) dz\end{aligned}$$

D'où $\alpha = 1/a$ (prendre $h = 1$). Ainsi $\alpha \in]0, 1[$ et N suit la loi géométrique de paramètre α .

Comme $\sum_{n=1}^{\infty} \alpha(1 - \alpha)^{n-1} = 1$, nous avons

$$\mathbb{E}(h(Y)) = \sum_{n=1}^{\infty} \mathbb{E}(h(Z_n)\mathbf{1}_{A_n})(1 - \alpha)^{n-1} = \int_{\mathbb{R}} h(z)f(z) dz$$

ce qui montre que Y a pour densité f .