

## Vecteurs aléatoires Gaussiens

### Préliminaires

**Notation pour les vecteurs et les matrices.** Tout vecteur  $u \in \mathbb{R}^d$  est un vecteur colonne (autrement dit une matrice avec une colonne et  $d$  lignes, ou  $d \times 1$ ). Si  $A \in \mathbb{R}^{m \times n}$  est une matrice  $m \times n$  ( $m$  lignes et  $n$  colonnes) dont les éléments sont courants  $a_{ij}$  pour  $i = 1, \dots, m, j = 1, \dots, n$  alors  $A^T$  est la matrice transposée qui est une matrice  $n \times m$  dont les éléments sont courants  $a_{ji}$ ,  $i = 1, \dots, n, j = 1, \dots, m$ .

Si  $A$  est une matrice  $n \times m$  et  $B$  une matrice  $m \times k$  alors le produit (lignes par colonnes)  $AB$  est défini étant la matrice  $n \times k$  dont les éléments sont courants  $(AB)_{ij} = \sum_{k=1}^m A_{ik} B_{kj}$  pour  $i = 1, \dots, n$  et  $j = 1, \dots, k$ . On a que  $(AB)^T = B^T A^T$ . On note  $I_d$  la matrice identité  $d \times d$ , i.e.  $(I_d)_{ij} = 1$  si  $i = j = 1, \dots, d$  et  $(I_d)_{ij} = 0$  si  $i \neq j, i, j = 1, \dots, d$ .

La transposée d'un vecteur  $u \in \mathbb{R}^d$  est un vecteur ligne (ce qui revient au même à une matrice  $1 \times d$ ); si  $u \in \mathbb{R}^n$  et  $v \in \mathbb{R}^m$  alors le produit matriciel  $uv^T$  est une matrice  $n \times m$  dont les éléments sont courants  $(uv^T)_{ij} = u_{i1}(v^T)_{1j} = u_{i1}v_{j1} = u_i v_j$  pour  $i = 1, \dots, n$  et  $j = 1, \dots, m$ ; si  $u \in \mathbb{R}^n$  et  $v \in \mathbb{R}^n$  le produit matriciel  $u^T v$  est une matrice  $1 \times 1$  dont les éléments sont courants  $(u^T v)_1 = \sum_{i=1}^n (u^T)_1 i v_{i1} = \sum_{i=1}^n u_i v_i$  qui n'est rien d'autre que le produit scalaire des deux vecteurs.

Si  $X$  est un vecteur aléatoire de dimension  $d$  alors  $E[X]$  est le vecteur de dimension  $d$  tel que  $(E[X])_i = E[X_i]$ . Si  $A$  est une matrice aléatoire alors  $E[A]$  est la matrice dont les éléments sont courants  $(E[A])_{ij} = E[A_{ij}]$ .

**Matrice de covariance.** Soit  $X = (X_1, \dots, X_d)$  un vecteur aléatoire dans  $\mathbb{R}^d$  tel que  $E[X_j^2] < +\infty$  pour tout  $j = 1, \dots, d$ . On appelle matrice de covariance du vecteur  $X$ , et on la notera  $\Sigma$ , la matrice dont les éléments sont courants  $\Sigma_{ij} = \text{Cov}(X_i, X_j)$ ,  $i, j = 1, \dots, d$ .

### Proposition 1.

- $\Sigma = E[(X - E[X])(X - E[X])^T]$
- Les éléments diagonaux sont les variances des composantes de  $X$  de  $\Sigma_{ii} = \text{Var}(X_i)$ .
- Pour tout  $u \in \mathbb{R}^d$ :  $\text{Var}(u^T X) = u^T \Sigma u$ .
- $\Sigma$  est symétrique et semi-définie positive,  $\Sigma_{ji} = \Sigma_{ij}$  et  $u^T \Sigma u \geq 0$  pour tout  $u \in \mathbb{R}^d$ .

### Démonstration.

- $E[(X - E[X])(X - E[X])^T]$  est une matrice dont les éléments sont courants  $E[(X - E[X])_{ij}(X - E[X])_{kl}] = \text{Cov}(X_i, X_j) = \Sigma_{ij}$ .
- $\Sigma_{ii} = \text{Cov}(X_i, X_i) = \text{Var}(X_i)$
- $\text{Var}(u^T X) = \text{Var}(u_1 X_1 + \dots + u_d X_d) = \text{Cov}(\sum_{i=1}^d u_i X_i, \sum_{j=1}^d u_j X_j) = \sum_{i,j=1}^d u_i u_j \text{Cov}(X_i, X_j) = \sum_{i,j=1}^d u_i u_j \Sigma_{ij} = u^T \Sigma u$ . La covariance étant une fonction bilinéaire.
- $\Sigma_{ij} = \text{Cov}(X_i, X_j) = \text{Cov}(X_j, X_i) = \Sigma_{ji}$ .  $u^T \Sigma u = \text{Var}(u^T X) \geq 0$ .

**Remarque 2.** Si  $X$  est tel que les composantes  $X_j$  sont indépendantes, alors la matrice de covariance est diagonale car  $\Sigma_{ij} = \text{Cov}(X_i, X_j) = 0$  si  $i \neq j$ . Et donc  $\text{Var}(u^T X) = \sum_{i=1}^d u_i^2 \text{Var}(X_i)$  pour tout  $u \in \mathbb{R}^d$ .

**Définition 3.** La fonction caractéristique  $\chi_X: \mathbb{R} \rightarrow \mathbb{C}$  d'une v.a. réelle est la transformée de Fourier de sa loi de probabilité :

$$\chi_X(t) = E[e^{itX}], \quad t \in \mathbb{R}$$

Si  $X$  admet une densité  $f_X$  alors  $\chi_X(t) = \int_{\mathbb{R}} e^{itx} f_X(x) dx$ .

La fonction caractéristique détermine de manière unique une loi de probabilité d'où son nom.

**Proposition 4.** Deux v.a. réelles  $X$  et  $Y$  telles que  $\chi_X(t) = \chi_Y(t)$  pour tout  $t \in \mathbb{R}$  ont la même loi, c-à-d  $P(X \in A) = P(Y \in A)$  pour tout Borel  $A$  de  $\mathbb{R}$ . Si  $X$  ou  $Y$  admet une densité alors l'autre v.a. admet une densité  $f_X(z)$  et  $f_Y(z)$  pour tout  $z \in \mathbb{R}$ .

Quelques propriétés calculatoires...

**Proposition 5.**

1. Pour tout  $t \in \mathbb{R}$ :  $\chi_X(-t) = \overline{\chi_X(t)}$ .
2. Pour tout  $a \in \mathbb{R}$ :  $\chi_{X+a}(t) = e^{ita} \chi_X(t)$ .
3. Soit  $\mu = E[X]$  et  $\sigma^2 = \text{Var}(X)$  et  $U = (X - \mu)/\sigma$  alors  $\chi_X(t) = \chi_U(t/\sigma) e^{it\mu}$  et  $\chi_U(t) = \chi_X(t/\sigma) e^{-it\mu}$ .
4. Si  $X$  et  $Y$  sont indépendantes alors  $\chi_{X+Y}(t) = \chi_X(t) \chi_Y(t)$ .

**Démonstration.**  $\chi_X(t) = E[e^{itX}] = \overline{\chi_X(-t)}$ .  $\chi_{X+a}(t) = E[e^{it(X+a)}] = e^{ita} E[e^{itX}] = e^{ita} \chi_X(t)$ .

$\chi_X(t) = E[e^{itX}] = E[e^{it(\mu + (X-\mu))}] = e^{it\mu} E[e^{it(X-\mu)}] = e^{it\mu} \chi_U(t/\sigma)$ . Pour  $X$  et  $Y$  indépendantes on a  $E[e^{it(X+Y)}] = E[e^{itX} e^{itY}] = E[e^{itX}] E[e^{itY}] = \chi_X(t) \chi_Y(t)$ .

**Exemple 6.** Si  $X \sim E(\lambda)$  alors

$$\chi_X(t) = E[e^{itX}] = \int_0^\infty e^{itx} \lambda e^{-\lambda x} dx = \frac{\lambda}{\lambda - it}.$$

En effet si  $F(x) = e^{-(\lambda x)/\lambda}$  alors  $F'(x) = e^{-(\lambda x)/\lambda}$  et

$$\lim_{x \rightarrow +\infty} F(x) = \lim_{x \rightarrow +\infty} \frac{\cos(t) e^{-(\lambda x)/\lambda}}{it\lambda} + i \lim_{x \rightarrow +\infty} \frac{\sin(t) e^{-(\lambda x)/\lambda}}{it\lambda} = 0$$

donc

$$\int_0^\infty e^{itx} \lambda e^{-\lambda x} dx = F(+\infty) - F(0) = -F(0) = -\frac{1}{\lambda - it}.$$

**Exemple 7.** Si  $Z \sim N(0, 1)$  alors

$$\chi_Z(t) = \int_{\mathbb{R}} e^{itx} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx = e^{-t^2/2}, \quad t \in \mathbb{R}.$$

Si  $X = \mu + \sigma Z$  alors  $X \sim N(\mu, \sigma^2)$  et donc

$$\chi_X(t) = e^{it\mu} \chi_Z(t/\sigma) = e^{it\mu} e^{-t^2/(2\sigma^2)} = \exp(itE[X] - \frac{\text{Var}(X)}{2} t^2).$$

**Exemple 8.** Soit  $Y \sim P(\lambda)$  (Poisson) et  $X|Y \sim \text{Bin}(Y, p)$  c-i-à-d

$$P(X = k | Y = n) = \binom{n}{k} p^k (1-p)^{n-k}$$

alors

$$E[e^{itX} | Y] = (1 + p(e^{it} - 1))^Y, \quad E[e^{itY}] = e^{-\lambda(1-p)} e^{p\lambda e^{it}}$$

et

$$\chi_X(t) = E[e^{itX}] = E[E[e^{itX} | Y]] = E[(1 + p(e^{it} - 1))^Y] = e^{-\lambda(1-p)} e^{p\lambda e^{it}} = e^{-\lambda(1-p)} e^{p\lambda e^{it}}$$

et donc la loi marginale de  $X$  est la loi de Poisson de paramètre  $\lambda p$ .

**Définition 9.** Soit  $X$  un vecteur aléatoire dans  $\mathbb{R}^d$ . La fonction caractéristique  $\chi_X: \mathbb{R}^d \rightarrow \mathbb{C}$  de  $X$  est donnée par

$$\chi_X(t) = E[e^{it^T X}] = E[e^{i(t_1 X_1 + \dots + t_d X_d)}] \quad t \in \mathbb{R}^d.$$

**Proposition 10.** Soit  $X$  un vecteur aléatoire dans  $\mathbb{R}^d$ . Les v.a.  $X_1, \dots, X_d$  sont indépendantes ssi

$$\chi_X(t) = \prod_{i=1}^d \chi_{X_i}(t_i) \quad t \in \mathbb{R}^d.$$

**Proposition 11.** Deux vecteurs aléatoires  $X, Y$  ont la même loi ssi  $\chi_X(t) = \chi_Y(t)$  pour tout  $t \in \mathbb{R}^d$ . En particulier si l'une des deux v.a. admet densité alors il en est de même pour l'autre v.a. et  $f_X = f_Y$ .

## Vecteurs Gaussiens

**Proposition 12.** (Stabilité) Soient  $X_1, \dots, X_d$  des v.a. Gaussiennes indépendantes. Pour tout  $u \in \mathbb{R}^d$  la combinaison linéaire  $Y = u_1 X_1 + \dots + u_d X_d$  est une v.a. Gaussienne réelle.

**Démonstration.** Il suffit de montrer que la fonction caractéristique de  $u^T X$  est la f.c. d'une v.a. Gaussienne réelle, i.e.

$$\chi_Y(t) = \exp(i t E[Y] - \frac{1}{2} \text{Var}(Y) t^2)$$

(voir l'exemple 7). Or  $E[Y] = \sum_{i=1}^d u_i E[X_i]$  et  $\text{Var}(Y) = \sum_{i=1}^d u_i^2 \text{Var}(X_i)$  par l'indépendance des  $X_i$ . Donc

$$\begin{aligned} \chi_Y(t) &= E[e^{it(u_1 X_1 + \dots + u_d X_d)}] = E\left[\prod_{j=1}^d e^{it u_j X_j}\right] = \prod_{j=1}^d E[e^{it u_j X_j}] \\ &= \prod_{j=1}^d \exp(i t u_j E[X_j] - \frac{1}{2} \text{Var}(X_j) u_j^2 t^2) = \exp(i t E[Y] - \frac{1}{2} \text{Var}(Y) t^2) \end{aligned}$$

On en déduit par l'unicité de la fonction caractéristique que  $Y$  est une v.a. Gaussienne.

**Définition 13.** On appelle vecteur Gaussien de dimension  $d$  un vecteur aléatoire  $X = (X_1, \dots, X_d)$  tel que toute combinaison linéaire de ses composantes  $(\sum_{j=1}^d u_j X_j)_{j=1, \dots, d}$  est une v.a. réelle Gaussienne, c-i-à-d  $u^T X$  est Gaussienne. (En particulier, toutes ses composantes  $X_j$  sont Gaussiennes.)

Par la Proposition 12 de tels vecteurs existent bien. Si  $g: \mathbb{R}^d \rightarrow \mathbb{R}^m$  est une application linéaire et  $X$  est un vecteur Gaussien de dimension  $d$  alors la composée  $Y = g(X)$  est un vecteur Gaussien de dimension  $m$ . En effet si  $g_k(x) = \sum_{j=1}^d g_{kj} x_j$  alors pour tout  $u \in \mathbb{R}^m$ , la v.a.

$$u^T Y = \sum_{k=1}^m u_k Y_k = \sum_{k=1}^m u_k g_k(X) = \sum_{k=1}^m u_k \sum_{j=1}^d g_{kj} X_j = \sum_{j=1}^d \left( \sum_{k=1}^m u_k g_{kj} \right) X_j = (g^T u)^T X$$

est Gaussienne, car combinaison linéaire des composantes de  $X$ .

Si  $X$  est un vecteur Gaussien (de dimension  $d$ ) alors  $Y = X^T E[X]$  est encore un vecteur Gaussien:  $u^T Y = u^T X^T E[X] = (E[X]^T u)^T X$  est Gaussienne car somme d'une Gaussienne et une constante (il suffit de voir que  $E[X]$  est la fonction caractéristique).

**Théorème 14** Soit  $X$  un vecteur Gaussien de dimension  $d$ . Soit  $\mu = E[X] \in \mathbb{R}^d$  la moyenne de  $X$  et  $\Sigma$  la matrice de covariance de  $X$ . La fonction caractéristique du vecteur  $X$  est donnée par

$$\chi(t) = \exp \left( i t^T \mu - \frac{1}{2} t^T \Sigma t \right) \quad t \in \mathbb{R}^d.$$

**Démonstration.**  $Y = t^T X$  est une v.a. Gaussienne.  $E[Y] = t^T \mu$  et  $\text{Var}(Y) = t^T \Sigma t$ .

$$\chi(t) = \chi_Y(1) = e^{i E[Y] - \text{Var}(Y)/2} = e^{i t^T \mu - t^T \Sigma t / 2}.$$

Remarque: la loi du vecteur Gaussien ne dépend que de son espérance et de sa covariance. La quantité  $\chi(t)$  est toujours 1.

**Définition 15.** Soit  $X$  un vecteur Gaussien de dimension  $d$  et  $\Sigma$  sa matrice de covariance. On note sa loi par  $N_d(\mu, \Sigma)$ .

**Lemme 16.** Soit  $\Sigma$  une matrice  $d \times d$ , symétrique et semi-définie positive. Alors il existe une matrice carrée de dimension  $d$  telle que  $\Sigma = A A^T$ . On dit que  $A$  est une racine carrée de  $\Sigma$ . De plus si  $\Sigma$  est inversible alors il en est de même de  $A$ .

**Démonstration.** Du fait que  $\Sigma$  est symétrique on déduit que elle est diagonalisable et donc qu'il existe une matrice orthogonale  $O$  (i.e. telle que  $O^T = O^{-1}$ ) et une matrice diagonale avec  $\lambda_{ii} \geq 0$  telles que  $\Sigma = O^T \Lambda O$ . Puisque  $\Sigma$  est semi-définie positive on a que les valeurs propres de  $\Sigma$  sont  $\geq 0$  et donc que  $\lambda_i \geq 0$  pour tout  $i = 1, \dots, d$ . Soit  $\Lambda^{1/2}$  la matrice diagonale telle que  $(\Lambda^{1/2})_{ii} = \sqrt{\lambda_i}$ , donc  $\Lambda^{1/2} \Lambda^{1/2} = \Lambda$  et si on pose  $A = O^T \Lambda^{1/2}$  on a que  $A A^T = O^T \Lambda^{1/2} (\Lambda^{1/2} O)^T = O^T \Lambda^{1/2} \Lambda^{1/2} O = O^T \Lambda O = \Sigma$ . Si  $\Sigma$  est inversible alors  $\lambda_i > 0$  pour tout  $i = 1, \dots, d$  et donc  $\Lambda^{1/2}$  est inversible et  $(\Lambda^{1/2})^{-1}$  est la matrice diagonale avec  $1/\sqrt{\lambda_i}$  sur la diagonale. Donc  $A^{-1} = (\Lambda^{1/2})^{-1} O^T = (\Lambda^{1/2})^{-1} O^T \Lambda^{1/2} O = (\Lambda^{1/2})^{-1} \Lambda^{1/2} O = O$  qui montre que  $A$  est inversible (étant le produit de deux matrices inversibles).

**Exemple 17.** Si  $X \sim N_d(\mu, \Sigma)$ ,  $A$  est une matrice  $n \times d$  et  $v \in \mathbb{R}^n$  alors

$$v + A X \sim N_n(v + A \mu, A \Sigma A^T).$$

En effet on remarque que  $t^T A X = (A^T t)^T X$  et donc

$$\begin{aligned} v + A X(t) &= e^{i t^T (v + A X)} = \exp \left( i t^T v + (A^T t)^T X \right) = \exp \left( i t^T v + \frac{1}{2} (A^T t)^T \Sigma (A^T t) \right) \\ &= \exp \left( i t^T [v + A \mu] + \frac{1}{2} t^T A \Sigma A^T t \right). \end{aligned}$$

$$Z = A^{\mathbb{P}_{\text{od}}^1}(X/\mathbb{P}_{\text{od}}^1/\mathbb{P}_{\text{od}}^1)$$
$$f_{\text{free}}(t) = \exp(i t^T \mathcal{P}_{\text{free}} t / 2)$$

**Démonstration.** On a dit que si  $X \sim \mathcal{N}(\mu, \Sigma)$  alors  $X(t) = f_{\Sigma}(t)$ . Il nous reste donc de montrer que si  $\Sigma$  est semi-définie positive et symétrique alors il existe bien un vecteur aléatoire Gaussien  $X$  de moyenne  $\mu$  et matrice de covariance  $\Sigma$  (et donc tel que  $X(t) = f_{\Sigma}(t)$ ). Mais par l'exemple 17 et le lemme 16 on a que il existe une matrice  $A$  telle que  $\Sigma = AA^T$  et que si  $Z \sim \mathcal{N}(0, I_d)$  alors  $X = \mu + AX$  est un vecteur aléatoire gaussien de moyenne  $\mu$  et matrice de covariance  $AA^T = \Sigma$ .

$$f_X(x) = (2\pi)^{-\frac{p}{2}} \det(\Sigma)^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} (x - \mu)^T \Sigma^{-1} (x - \mu) \right\} \quad x \in \mathbb{R}^d \quad (1)$$
$$f_Z(z) = f_{Z_1}(z_1) f_{Z_a}(z_d) = \frac{1}{(2\pi)^{d/2}} \exp \left\{ -\frac{1}{2} (z_1^2 + z_d^2) \right\} = \frac{1}{(2\pi)^{d/2}} \exp \left\{ -\frac{t^T t}{2} \right\}$$
$$f_X(x) = f_Z(\mathbb{E}^{\text{old}}(x)) \mathcal{J}_{\mathbb{E}^{\text{old}}(x)} = \frac{1}{(2)^{d/2}} \det(A^{\text{old}}) \exp\left(\frac{\mathbb{E}^{\text{old}}(x)^T \mathbb{E}^{\text{old}}(x)}{2}\right).$$
$$\begin{aligned} (\mathbb{E}^{\text{old}}(x))^T \mathbb{E}^{\text{old}}(x) &= [\mathbb{A}^{\text{old}}(x \mathbb{P}_{\text{old}} \mathbb{P}_{\text{old}}^T) A^{\text{old}}(x \mathbb{P}_{\text{old}} \mathbb{P}_{\text{old}}^T) = (x \mathbb{P}_{\text{old}} \mathbb{P}_{\text{old}}^T (A^{\text{old}})^T A^{\text{old}}(x \mathbb{P}_{\text{old}} \mathbb{P}_{\text{old}}^T) \\ &= (x \mathbb{P}_{\text{old}} \mathbb{P}_{\text{old}}^T (A^T)^{\text{old}} \mathbb{A}^{\text{old}}(x \mathbb{P}_{\text{old}} \mathbb{P}_{\text{old}}^T) = (x \mathbb{P}_{\text{old}} \mathbb{P}_{\text{old}}^T (A A^T)^{\text{old}}(x \mathbb{P}_{\text{old}} \mathbb{P}_{\text{old}}^T) = (x \mathbb{P}_{\text{old}} \mathbb{P}_{\text{old}}^T)^{\text{old}}(x \mathbb{P}_{\text{old}} \mathbb{P}_{\text{old}}^T) \end{aligned}$$

et  $\det(\Sigma) = \det(AA^T) = \det(A) \det(A^T) = [\det(A)]^2$  et  $\det(A^{-1}) = 1/\det(A)$ . Donc  $\det(\Sigma^{-1}) = 1/[\det(\Sigma)]^{1/2}$  et on obtient la formule (1).

**Lemme 20.** Soit  $X \sim N_d(\mu, \Sigma)$ . Les composantes  $(X_j)_{j=1,\dots,d}$  de  $X$  sont *mutuellement indépendantes* (c-à-d il n'existe pas de vecteur  $\mathbf{R}^d$  et  $c \in \mathbb{R}$  tels que  $u \neq 0$  et  $u^T X = c$ ) ssi  $\Sigma$  est inversible.

**Démonstration.** Montrons que si  $\Sigma$  est inversible les composantes de  $X$  sont *mutuellement indépendantes*: en effet si tel vecteur existait alors pour tout  $j = 1, \dots, d$ :

$$0 = \text{Cov}(c, X_j) = \text{Cov}(u^T X, X_j) = \sum_{k=1}^d u_k \text{Cov}(X_k, X_j) = \sum_{k=1}^d u_k \Sigma_{jk} = (\Sigma u)_j$$

et donc  $\Sigma u = 0$  qui montre que  $\Sigma$  a une valeur propre nulle et donc ne peut pas être inversible. Réciproquement si  $\Sigma$  est singulière (c-à-d elle n'est pas inversible) alors il existe  $u \in \mathbb{R}^d$ ,  $u \neq 0$  tel que  $\Sigma u = 0$  et donc  $\text{Var}(u^T X) = u^T \Sigma u = 0$  ce qui implique que la v.a.  $u^T X$  est constante et égale à  $\mathbb{E}[u^T X] = u^T \mu$  donc  $u^T X = c = u^T \mu$  qui montre que les composantes de  $X$  sont *mutuellement dépendantes*.

**Exemple 21.** Soit  $(X, Y)$  un couple aléatoire Gaussien de matrice de covariance égale à

$$\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}.$$

1. Condition nécessaire et suffisante pour que soit vraiment la matrice de covariance de  $(X, Y)$  est que  $\Sigma$  soit symétrique et semi-définie positive ( $\det(\Sigma) \geq 0$  et  $\text{Tr}(\Sigma) \geq 0$ ), i.e.  $1 - \rho^2 \geq 0$  /  $|\rho| \leq 1$ .
2. Condition nécessaire et suffisante sur  $\rho$  pour qu'en plus le couple  $(X, Y)$  admette une densité est que  $\Sigma$  soit inversible (i.e. définie positive  $\det(\Sigma) > 0$  et  $\text{Tr}(\Sigma) > 0$ ). Donc  $1 - \rho^2 > 0$  /  $|\rho| < 1$ .
3. Si  $|\rho| < 1$  et  $(X, Y)$  est supposé centré ( $\mathbb{E}[X] = \mathbb{E}[Y] = 0$ ), alors  $(X, Y)$  admet pour densité

$$f_{(X,Y)}(x, y) = \frac{e^{-\frac{1}{2(1-\rho^2)}(x^2 - 2\rho xy + y^2)}}{2\pi \det(\Sigma)^{1/2}}$$

et on a

$$\begin{aligned} \det(\Sigma) &= \begin{vmatrix} 1 & \rho \\ \rho & 1 \end{vmatrix} = 1 - \rho^2 \\ f_{(X,Y)}(x, y) &= \frac{1}{2\pi(1-\rho^2)} e^{-\frac{1}{2(1-\rho^2)}(x^2 - 2\rho xy + y^2)} \\ &= \frac{1}{2\pi(1-\rho^2)} e^{-\frac{1}{2(1-\rho^2)}(x^2 - 2\rho xy + y^2)}, \quad x, y \in \mathbb{R} \end{aligned}$$

**Proposition 22.** Soit  $X$  un vecteur Gaussien de dimension  $d$ . Les v.a. Gaussiennes  $X_i$  et  $X_j$  sont indépendantes ssi  $\text{Cov}(X_i, X_j) = \delta_{ij} = 0$ . D'une manière générale, les  $X_{i_1}, \dots, X_{i_k}$  sont indépendantes ssi  $\text{Cov}(X_{i_a}, X_{i_b}) = 0$  pour tout  $a \neq b$  et  $a, b = 1, \dots, k$ .

**Démonstration.** Il est clair que si  $X_i$  et  $X_j$  sont indépendantes alors  $\delta_{ij} = 0$ . Montrons alors que  $\delta_{ij} = 0$  implique l'indépendance de  $X_i$  et  $X_j$ . Soit  $t = (t_1, t_2) \in \mathbb{R}^2$  et  $\phi = E[X]$ . La v.a.  $Y = t_1 X_i + t_2 X_j$  est une v.a. Gaussienne de moyenne  $E[Y] = t_1 \phi_i + t_2 \phi_j$  et variance  $\text{Var}(Y) = t_1^2 \sigma_{ii} + t_2^2 \sigma_{jj}$  et donc

$$\phi_{(X_i, X_j)}(t) = \phi_Y(1) = E[e^{i(t_1 X_i + t_2 X_j)}] = e^{i(t_1 \phi_i + t_2 \phi_j)} e^{-(t_1^2 \sigma_{ii} + t_2^2 \sigma_{jj})/2} = \phi_{X_i}(t_1) \phi_{X_j}(t_2)$$

ce qui implique l'indépendance. (En effet, par la propriété fondamentale des fonctions caractéristiques, on a montré que le couple  $(X_i, X_j)$  a une densité égale à la densité d'un couple de v.a. indépendantes.)

## Autres distributions classiques

### Loi Gamma

La v.a.  $X$  suit une loi gamma de paramètres  $\alpha > 0$  et  $\lambda > 0$  ssi sa densité est donnée par

$$f_X(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} \mathbf{1}_{x>0}$$

et on note  $X \sim G(\alpha, \lambda)$ . Le paramètre est la forme de la loi Gamma et  $\lambda$  son intensité. On rappelle que

$$\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} dt, \quad \alpha > 0$$

et que

1.  $\Gamma(\alpha + 1) = \alpha \Gamma(\alpha)$ .  $\Gamma(1) = 1$ ,  $\Gamma(n+1) = n!$  si  $n \in \mathbb{N}$ .
2.  $E[X] = \alpha / \lambda$ ,  $\text{Var}(X) = \alpha / \lambda^2$ .

### Loi Beta

On dit que  $X$  suit une loi beta de paramètres  $a > 0$  et  $b > 0$  ssi la densité de  $X$  est donnée par

$$f_X(x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1} (1-x)^{b-1} \mathbf{1}_{0 < x < 1}$$

et on note  $X \sim B(a, b)$ .

**Proposition 23.** On a  $E[X] = a/(a+b)$ ,  $\text{Var}(X) = ab/((a+b)^2(a+b+1))$ . Si  $X$  et  $Y$  sont indépendantes et  $X \sim G(\alpha, \lambda)$  et  $Y \sim G(\beta, \lambda)$  alors

$$S = X + Y \sim G(\alpha + \beta, \lambda), \quad R = \frac{X}{X+Y} \sim B(\alpha, \beta)$$

et  $S$  et  $R$  sont indépendantes.

**Démonstration.** Si  $X \sim B(a, b)$

$$\begin{aligned} E[X] &= \frac{(a+b)}{(a-b)} \int_0^1 x x^{a-1} (1-x)^{b-1} dx = \frac{(a+b)}{(a-b)} \int_0^1 x^{(a+1)-1} (1-x)^{b-1} dx \\ &= \frac{(a+b)}{(a-b)} \frac{(a+1-b)}{(a+b+1)} = \frac{a}{a+b} \end{aligned}$$

$$E[X^2] = \frac{(a+b)}{(a-b)} \int_0^1 x^{(a+2)-1} (1-x)^{b-1} dx = \frac{(a+b)}{(a-b)} \frac{(a+2-b)}{(a+b+2)} = \frac{a(a+1)}{(a+b)(a+b+1)}.$$

Si  $X \sim G(\alpha, \lambda)$  et  $Y \sim G(\beta, \lambda)$  alors on considère le changement de variables  $(x, y) = (s, r)$  avec  $s = x + y$  et  $r = x/(x + y)$ .  $\alpha, \beta \in \mathbb{R}_+^2$ ,  $\lambda \in \mathbb{R}_+ \setminus \{0, 1\}$  et  $\mathcal{D}(s, r) = (x, y)$  avec  $x = rs$  et  $y = s(1-r)$ . La matrice Jacobienne de  $\mathcal{D}$  est donnée par

$$J_{\mathcal{D}} = \frac{D \mathcal{D}(s, r)}{D(s, r)} = \begin{vmatrix} \frac{x}{s} & \frac{x}{r} \\ \frac{y}{s} & \frac{y}{r} \end{vmatrix} = \begin{vmatrix} r & s \\ (1-r)s & s \end{vmatrix} = rs - s^2(1-r) = rs$$

et donc

$$\begin{aligned} f_{(S,R)}(s, r) &= f_{(X,Y)}(x, y) |J_{\mathcal{D}}| = \frac{1}{\Gamma(\alpha)\Gamma(\beta)} s^{\alpha-1} r^{\alpha-1} (1-r)^{\beta-1} e^{-s/\lambda} \mathbf{I}_{x>0, y>0} \\ &= \frac{1}{\Gamma(\alpha)\Gamma(\beta)} s^{\alpha+\beta-1} r^{\alpha-1} (1-r)^{\beta-1} \mathbf{I}_{s>0} \mathbf{I}_{0<r<1} = f_R(r) f_S(s) \end{aligned}$$

avec

$$f_R(r) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} r^{\alpha-1} (1-r)^{\beta-1} \mathbf{I}_{0<r<1}, \quad f_S(s) = \frac{1}{\Gamma(\alpha+\beta)} s^{\alpha+\beta-1} e^{-s/\lambda} \mathbf{I}_{s>0}$$

et donc  $R \sim B(\alpha, \beta)$  et  $S \sim G(\alpha+\beta, \lambda)$  et  $S$  et  $R$  sont indépendantes.

## Loi du Khi-deux ( $\chi^2$ )

**Définition 24.** Soit  $X = (X_1, \dots, X_d)$  un vecteur aléatoire Gaussien centré de matrice de covariance identité (i.e. les composantes  $X_i$  sont indépendantes  $X_i \sim N(0, 1)$ ). On appelle la loi du Khi-deux de  $d$  degrés de liberté la loi de la v.a.

$$Y = X_1^2 + \dots + X_d^2$$

et on la note  $Y \sim \chi_d^2$ .

**Proposition 25.** Si  $Y \sim \chi_d^2$  alors  $Y \sim G(d/2, 1/2)$ ,  $E[Y] = d$ ,  $\text{Var}(Y) = 2d$ .

**Démonstration.** La loi du carré  $\chi^2$  de la Gaussienne standard  $X_1^2$  est  $G(1/2, 1/2)$ , en effet la méthode de la fonction muette donne

$$E[h(Q)] = \int_{\mathbb{R}} h(x^2) \frac{e^{-x^2/2}}{\sqrt{2\pi}} dx = \frac{\sqrt{2}}{\sqrt{\pi}} \int_0^\infty h(q) e^{-q/2} q^{-1/2} dq.$$

Donc la loi de la somme de  $d$  carrés de Gaussiennes standards est  $G(d/2, 1/2)$  par les propriétés des lois Gamma (voir la Proposition 23).

On a que  $E[Q] = 1$  et  $\text{Var}(Q) = 2$  par les propriétés des Gammas, et  $E[Y] = dE[Q]$  et  $\text{Var}(Y) = d\text{Var}(Q)$ .

## Loi de Student

**Définition 26.** On appelle loi de Student de paramètre  $d$ , notée  $T_d$ , la loi de la v.a.

$$T = \frac{X}{Y/d}$$

où  $X \sim N(0, 1)$ ,  $Y \sim \chi^2_d$  et  $X$  et  $Y$  sont indépendantes.  $T$  admet pour densité

$$f_T(t) = \frac{\Gamma((d+1)/2)}{\Gamma(d/2) \sqrt{d}} \left(1 + \frac{t^2}{d}\right)^{-(d+1)/2}.$$

**Remarque 27.** Lorsque  $d \rightarrow \infty$  on a que

$$\lim_d \frac{f_T(t)}{f_T(0)} = \lim_d \left(1 + \frac{t^2}{d}\right)^{-(d+1)/2} = e^{-t^2/2}$$

limite qui est proportionnelle à la densité d'une Gaussienne standard (centrée et réduite).