Probabilités

Département Signal et Communications

ENST-Bretagne

Septembre 2002

samir.saoudi@enst-bretagne.fr

`http://www-sc.enst-bretagne.fr/~saoudi/transp_proba.ps`

`http://www-sc.enst-bretagne.fr/~saoudi/poly_proba.ps`

Contenu

Espace de probabilité

Définition 1 *Soit Ω un ensemble, $\mathcal{B}$ un ensemble de parties de Ω. On dit que $\mathcal{B}$ est une **tribu** de Ω si elle contient Ω et si elle est stable pour les opérations de complémentation et de réunion dénombrable, soit :*

1. $\Omega \in \mathcal{B}$
2. $A \in \mathcal{B} \Longrightarrow \overline{A} \in \mathcal{B}$
3. $(A_n)_{n \geq 1} \in \mathcal{B} \Longrightarrow \bigcup_{n=1}^{+\infty} A_n \in \mathcal{B}$

Le couple $(\Omega, \mathcal{B})$ définit un ***espace probabilisable*** (ou mesurable).

$\mathcal{B}$ est aussi stable pour l'opération d'intersection dénombrable.

Exemples de tribus :

- $\mathcal{B} = \mathcal{P}(\Omega)$ (ensemble des parties de Ω)
- $\mathcal{B} = \{\emptyset, \Omega\}$
- $\mathcal{B} = \{\emptyset, \Omega\, A, \overline{A}\}$ avec $(A \subset \Omega)$.

Définition 2 *Soit* $(\Omega_1, \mathcal{B}_1)$ *et* $(\Omega_2, \mathcal{B}_2)$ *deux espaces probabilisables (ou mesurables). Soit* f *une application de* Ω_1 *dans* Ω_2. f *est une* ***application mesurable*** *si :*

$$\forall B_2 \in \mathcal{B}_2 \text{ on a : } f^{-1}(B_2) \in \mathcal{B}_1$$

Rq : la composition de deux applications mesurables est une application mesurable.

Définition 3 *Soit* $(\Omega, \mathcal{B})$ *un espace probabilisable. Soit* μ *une application de la tribu* $\mathcal{B}$ *dans* $[0, +\infty]$. μ *est une* ***mesure positive*** *si pour toute suite dénombrable d'événements* $(A_n)_{n\geq 1}$ *mutuellement disjoints (ou incompatibles,* $A_i \cap A_j = \emptyset, \forall i \neq j$*), on a :*

$$\mu\left(\bigcup_{n=1}^{+\infty} A_n\right) = \sum_{n=1}^{+\infty} \mu(A_n)$$

Le triplet $(\Omega, \mathcal{B}, \mu)$ définit un ***espace mesuré***.

propriétés :

1. $\mu(\emptyset) = 0$
2. pour des événements mutuellements disjoints, on a :

$$\mu\left(\bigcup_1^n A_n\right) = \sum_1^n \mu\left(A_n\right)$$

3. $A \subset B \Longrightarrow \mu(A) \leq \mu(B)$
4. $\left(A_n\right)_{n\geq 1} \nearrow A$ (au sens de $\subset$) $(\bigcup A_n = A)$, $\Longrightarrow \mu\left(A_n\right)_{n\geq 1} \nearrow \mu(A)$
5. $\left(A_n\right)_{n\geq 1} \searrow A$ tel que $\mu(A_1) < +\infty$ $(\bigcap A_n = A)$, $\Longrightarrow$
   $\mu\left(A_n\right)_{n\geq 1} \searrow \mu(A)$

Définition 4 *Soit* $(\Omega, \mathcal{B}, \mu)$ *un espace mesuré.* μ *est une mesure borné si on a,* $\mu(\Omega) < +\infty$.

$\Longrightarrow \forall A \in \Omega, \mu(A) < +\infty$

Définition 5 *Soit* $(\Omega, \mathcal{B})$ *un espace probabilisable. Une* ***Probabilité*** P *est définie comme étant une mesure positive sur* $(\Omega, \mathcal{B})$ *telle que* $P(\Omega) = 1$

Le triplet $(\Omega, \mathcal{B}, P)$ définit un ***espace probabilisé***.

Définition 6 *Soit* $D = \{D_x = \left]-\infty, x\right[, x \in \mathbb{R}\}$ *(classe des demi-droites ouvertes). La tribu des boréliens de* $\mathbb{R}$*, notée* $\mathcal{B}(\mathbb{R})$*, est la tribu engendrée par* D *(*$\mathcal{B}(\mathbb{R}) = \sigma(D)$*).*

Remarque :

- $\mathcal{B}(\mathbb{R})$ contient tout les $\left[x, +\infty\right[$ $(= \overline{D}_x)$, tout les $\left[x, y\right[$ $(= \overline{D}_x \cap D_y)$, tous les $\left]x, +\infty\right[$ $(= \bigcap_{n=1}^{+\infty} \left[x + \frac{1}{n}, +\infty\right[)$.
- $\mathcal{B}(\mathbb{R})$ contient toutes les réunions dénombrables d'intervalles ouverts.
- $\mathcal{B}(\mathbb{R})$ est aussi la tribu engendrée par les ouverts (ou les fermés).

Rappels d'analyse combinatoire

Soit $S = \{s_1, s_2, \ldots, s_N\}$ une population formée de N individus distincts.
Un échantillon de taille p extrait de cette population est une suite **ordonnée** $\left(s_{i_1}, s_{i_2}, \ldots, s_{i_p}\right)$ de p éléments de S :

- **tirage avec remise** :
  $\left(s_{i_1}, s_{i_2}, \ldots, s_{i_p}\right)$ peut comporter plusieurs fois le même élément. $\Omega = S^p$ et $card(\Omega) = N^p$.
- **tirage sans remise** : $(p \leq N)$,
  $card(\Omega) = N(N-1)\ldots(N-p+1) = A_N^p = \frac{N!}{(N-p)!}$ (arrangement)

On peut définir les sous-populations de taille p de S $(p \leq N)$ pour lequel **l'ordre n'intervient plus** : $= A_N^p$ divisé par $p!$ (nombre de bijections (permutations) d'un ensemble à p éléments),

$$card(\Omega) = A_N^p / p! = \frac{N!}{p!(N-p)!} = C_N^p$$

C_N^p : **le nombre de combinaisons** de p éléments parmi N.

Probabilités conditionnelles

Définition 7 *Etant donné un espace probabilisé* $(\Omega, \mathcal{B}, P)$ *et un événement* B *de* $\mathcal{B}$ *de probabilité non nulle. La probabilité de l'événement* A *conditionnelllement à l'événement* B*, notée* $P(A/B)$*, est définie par :*

$$P(A/B) = \frac{P(A \cap B)}{P(B)}$$

Remarques :

1. $P(A \cap B) = P(B)\, P(A/B)$

   $P(A_1 \cap \ldots \cap A_n) =$

   $P(A_1)P(A_2/A_1)P(A_3/A_1 \cap A_2) \ldots P(A_n/A_1 \cap \ldots \cap A_{n-1})$

2. la règle de Bayes :

$$P(B/A) = \frac{P(B)P(A/B)}{P(A)}$$

Proposition 1 *L'application :*

$$\begin{array}{rccl} P_B : & \mathcal{B} & \longmapsto & [0,1] \\ & A & \longmapsto & P_B(A) = \frac{P(A\cap B)}{P(B)} \end{array}$$

est une probabilité (sur $(\Omega, \mathcal{B})$*).*

Formule de Bayes : Soit (A_i) une partition finie ou infinie de Ω telle que pour tout i, $P(A_i) > 0$. Une telle partition est caractérisée par :

$$A_i \cap A_j = \emptyset \qquad \text{et} \qquad \cup_i A_i = \Omega$$

$\Longrightarrow \forall B \in \mathcal{B}$:

$$P(B) = \sum_i P(B \cap A_i) = \sum_i P(B/A_i)P(A_i)$$

$\Longrightarrow$ la formule de Bayes :

$$P(A_i/B) = \frac{P(B/A_i)P(A_i)}{\sum_i P(B/A_i)P(A_i)}$$

Evénements indépendants

Définition 8 *Soit $(\Omega, \mathcal{B}, P)$ un espace probabilisé et A et B deux événements. On dit que A et B sont stochastiquement indépendants si et seulement si :*

$$P(A \cap B) = P(A)P(B)$$

Remarques :

1. Lorsque $P(A) \neq 0$, A et B sont stochastiquement indépendants si et seulement si $P(B/A) = P(B)$.
2. La notion d'indépendance dépend de la probabilité P.
3. Un événement B (tq $P(B) = 0$) est indépendant de n'importe quel événement A.
4. Si A et B sont indépendants alors $\overline{A}$ et B sont indépendants (idem pour A et $\overline{B}$ ou ($\overline{A}$ et $\overline{B}$)).
5. Si A et B sont indépendants alors :
   $P(A \cup B) = P(A) + P(B) - P(A)P(B)$

Variables Aléatoires

Définition 9 *Soit* $(\Omega, \mathcal{B})$ *un espace probabilisable et* $\mathbb{R}$ *muni de la tribu des boréliens* $\mathcal{B}(\mathbb{R})$. *Une variable aléatoire réelle est une application mesurable de* $(\Omega, \mathcal{B})$ *dans* $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ *:*

$$\begin{array}{rccc} X : & \Omega & \longmapsto & \mathbb{R} \\ & \omega & \longrightarrow & X(\omega). \end{array}$$

Remarques :

1. La définition d'une v.a. ne fait pas intervenir la probabilité P.
2. On utilise les lettres majuscule $X, Y, Z, \ldots$ pour noter ces v.a.
3. La mesurabilité de la v.a. X signifie que :

$$\forall B \in \mathcal{B}(\mathbb{R}), X^{-1}(B) = \{\omega \in \Omega / X(\omega) \in B\} \in \mathcal{B}$$

Loi d'une v.a.

$$(\Omega, \mathcal{B}, P) \quad \stackrel{X}{\longmapsto} \quad (\mathbb{R}, \mathcal{B}(\mathbb{R}), P_X)$$

P possède une mesure image sur $\mathcal{B}(\mathbb{R})$, notée P_X. Soit :

$$\begin{aligned} \forall B \in \mathcal{B}(\mathbb{R}), \quad P_X(B) &= P\left(X^{-1}(B)\right) \\ &= P\left(\{\omega \in \Omega / X(\omega) \in B\}\right) \\ &= P\left(X \in B\right) \end{aligned}$$

La probabilité induite de P par la v.a. X, est appelée ***la loi de*** X.

Fonction de répartition d'une v.a.r.

Définition 10 *La fonction de répartition F_X d'une variable aléatoire réelle X est définie par :*

$$\begin{array}{rccl} F_X : & \mathbb{R} & \longmapsto & [0,1] \\ & x & \longmapsto & F_X(x) = P_X\left(]-\infty, x[\right) = P\left(X^{-1}(]-\infty, x[)\right) = P(X < x) \end{array}$$

<u>Remarques</u> :

1. Toute fonction de répartition est :
   - croissante
   - continue à gauche
   - telle que :

$$\lim_{x \to +\infty} F_X(x) = 1 \qquad \text{et} \qquad \lim_{x \to -\infty} F_X(x) = 0.$$

   Réciproquement, toute fonction réelle d'une variable réelle satisfaisant à ces conditions est la fonction de répartition d'une loi de probabilité.

2. Voici quelques propriétés de F_X :

$$\begin{aligned}
P(X \leq x) &= F_X\left(x^+\right) \\
P(X = x) &= F_X\left(x^+\right) - F_X(x) \\
P(a \leq X < b) &= F_X(b) - F_X(a) \\
P(X \geq x) &= 1 - F_X(x) \\
\ldots etc
\end{aligned}$$

variables aléatoires discrètes

Définition 11 *Soit X une variable aléatoire définie sur un espace probabilisé $(\Omega, \mathcal{B}, P)$. X est une v.a. discrète si $X(\Omega)$ est un ensemble dénombrable.*

Remarques :

1. Exemples d'ensembles de $X(\Omega)$: $\mathbb{N}$, $\mathbb{Z}$, $\{0, 1, \ldots, n\}$, ...
2. La loi d'une variable aléatoire discrète est complètement déterminée par une famille dénombrable de nombre positifs $p_i = P(X = i)$ où $i \in X(\Omega)$ tels que $\sum_{i \in X(\Omega)} p_i = 1$.
3. Pour une variable aléatoire à valeurs entières (dans $\mathbb{N}$), la fonction de répartition est donnée par :

$$F_X(x) = \sum_{i<x} P(X = i)$$

Variables aléatoires absolument continues

Définition 12 *Soit X une variable aléatoire définie sur un espace probabilisé $(\Omega, \mathcal{B}, P)$. X est une v.a. absolument continue si et seulement si il existe une fonction $f_X(x)$, appelée densité de probabilité, telle que :*

$$\forall x \in \mathbb{R} \qquad F_X(x) = \int_{-\infty}^{x} f_X(t)\,dt$$

Remarques :

1. $P(X \in ([a,b])) = P(X \in (]a,b])) = P(X \in ([a,b[)) = P(X \in (]a,b[))$.
2. $\forall x \in \mathbb{R}$, on a : $P(X = x) = 0$.
3. La probabilité de tout borélien B est : $P(X \in B) = \int_B f_X(x)\,dx$
4. La densité de probabilité $f_X(x)$ est non nécessairement majorée.
5. f_X est une densité de probabilité si et seulement si :

$$f_X(x) \geq 0 \qquad \text{et} \qquad \int_{\mathbb{R}} f_X(x)\,dx = 1$$

Exemples de lois de probabilité

loi de Bernouilli : $0 < p < 1$, $X \sim \mathcal{B}(p)$, $P(X=1) = p = 1 - P(X=0)$

loi géométrique : $(0 < p < 1)$ $N \rightsquigarrow \mathbb{N}^*$, $\quad P(N=k) = (1-p)^{k-1}p$

Interprétation : si on répète une infinité de fois une épreuve de $\mathcal{B}(p)$ de manière indépendante. $N =$ indice de la première apparition d'un succés.

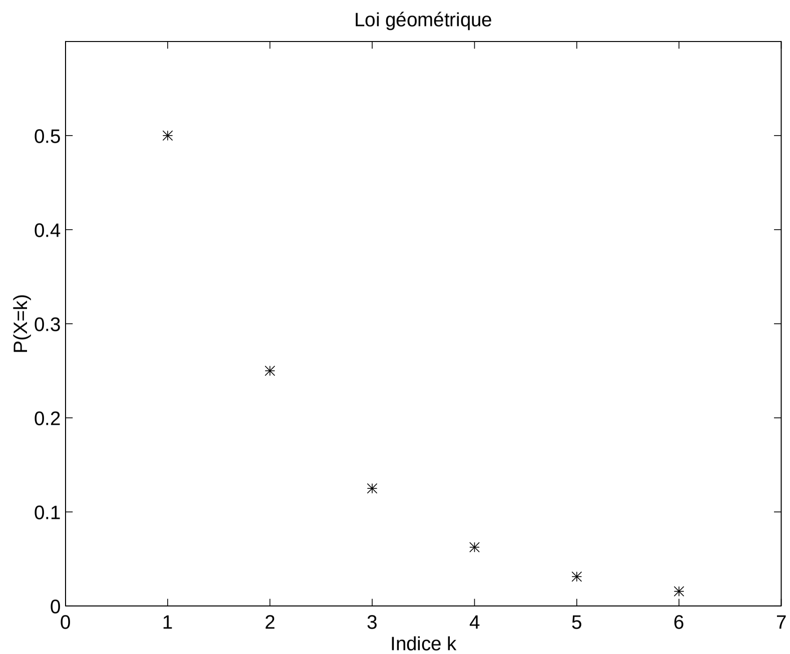


FIG. 1 – Distribution de la loi géométrique avec $p = 0.5$

loi binomiale :($X \sim \mathcal{B}\,(n,p)$,$n \in \mathbb{N}^*$ et $0 < p < 1.$)

$X \rightsquigarrow \{0, 1, \ldots, n\}$ et $P(X = k) = C_n^k p^k (1-p)^{n-k}$

Interprétation : lorsqu'on répète n fois une épreuve de Bernouilli de manière indépendante. $X =$ au nombre de succés.

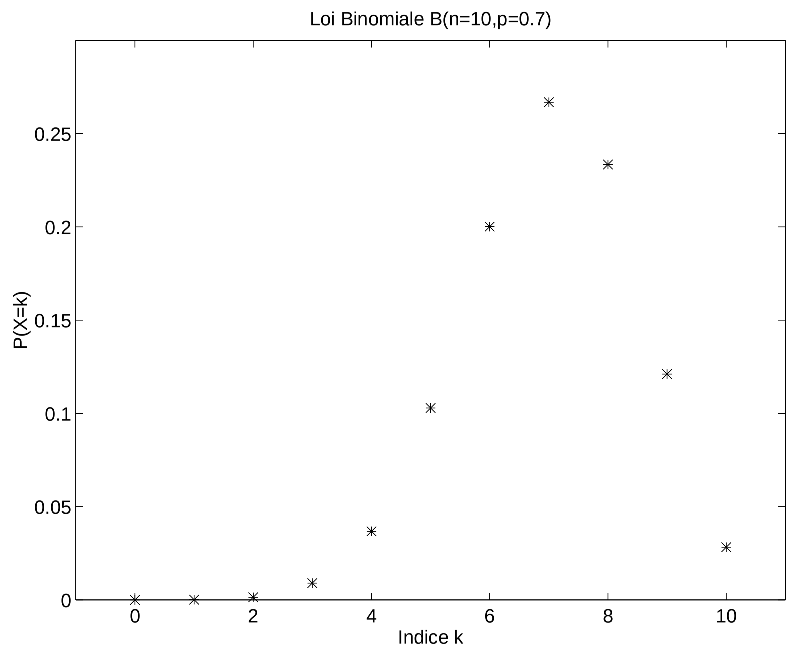


Fig. 2 – Distribution de la loi Binomiale avec $p = 0.7$ et $n = 10$

loi de Poisson : $(X \sim \mathcal{P}(\lambda),\ \lambda > 0)$

$X \rightsquigarrow \mathbb{N}$ et $P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}$

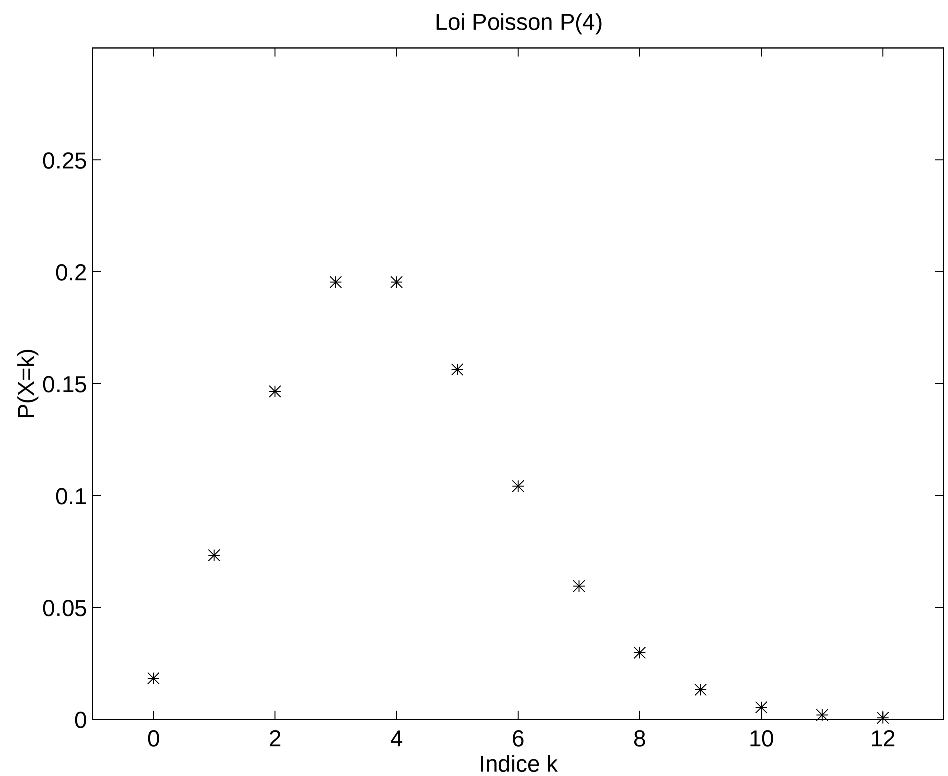


Fig. 3 – Distribution de la loi de Poisson $\mathcal{P}(\lambda = 4)$

loi uniforme : $X \sim \mathcal{U}_{[a,b]}$, $f_X(x) = \frac{1}{b-a} \mathbb{1}_{[a,b]}(x)$

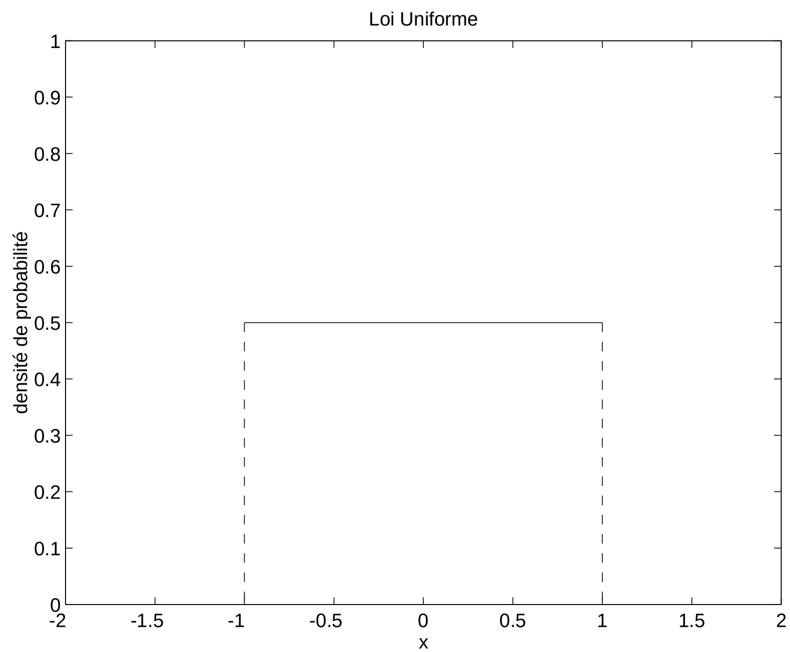


Fig. 4 – Distribution de la loi Uniforme $\mathcal{U}_{[-1,+1]}$

loi normale : (ou loi de Gauss) $X \sim \mathcal{N}\left(m, \sigma^2\right)$, $f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}}$

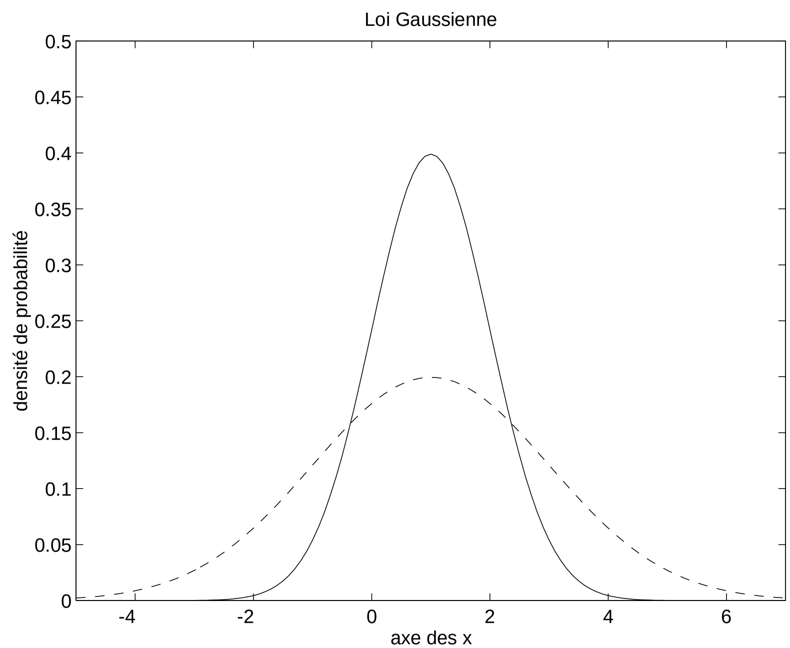


FIG. 5 – Distribution de la loi Normale : $\mathcal{N}(1, 1)$ et $\mathcal{N}\left(1, 2^2\right)$

loi Log Normale : $Z = z_0 + e^X$ où $X \sim \mathcal{N}\left(m, \sigma^2\right)$,

$$f_Z(z) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(\ln(z-z_0)-m)^2}{2\sigma^2}} \frac{1}{z-z_0} \mathbb{I}_{]z_0,+\infty}(z)$$

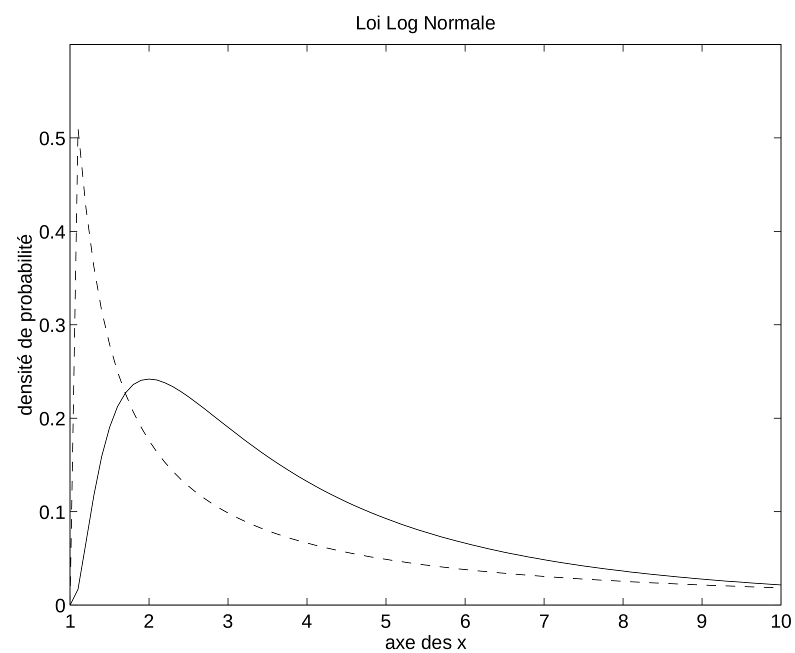


FIG. 6 – Distribution de la loi Log Normale : $z_0 = 1$, $m = 1$ et $\sigma = 1$ (trait continue) ou 2

loi Exponentielle : $X \sim \mathcal{E}(\lambda),\ f_X(x) = \lambda e^{-\lambda x}\,\mathbb{1}_{\mathbb{R}+}(x)$

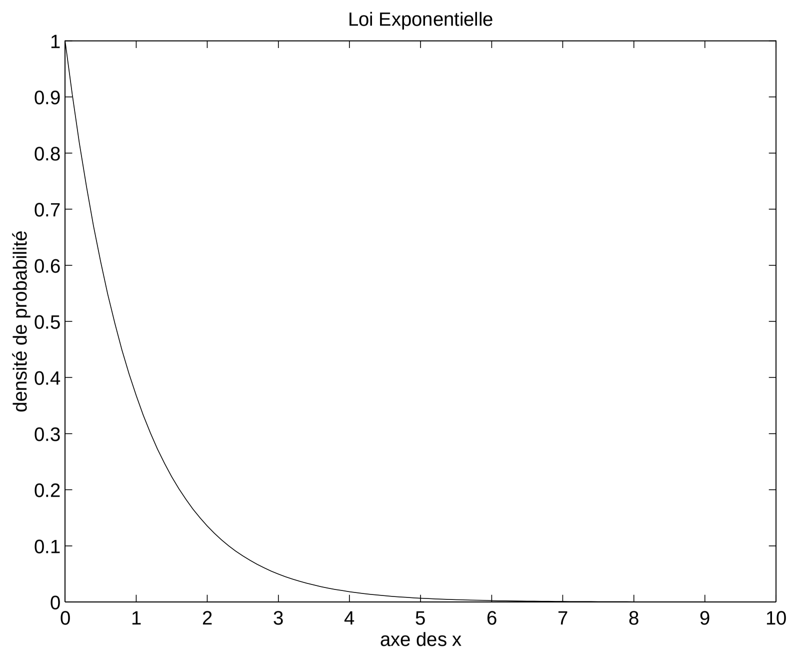


Fig. 7 – Distribution de la loi Exponentielle $\mathcal{E}(\lambda = 1)$

loi de Rayleigh :σ, $f_R(r) = \frac{r}{\sigma^2} e^{-\frac{r^2}{2\sigma^2}} \mathbb{1}_{\mathbb{R}^+}(r)$ Rq : $R = \sqrt{X^2 + Y^2}$ où X et Y sont deux v.a. indépendantes et $\sim \mathcal{N}\left(0, \sigma^2\right)$.

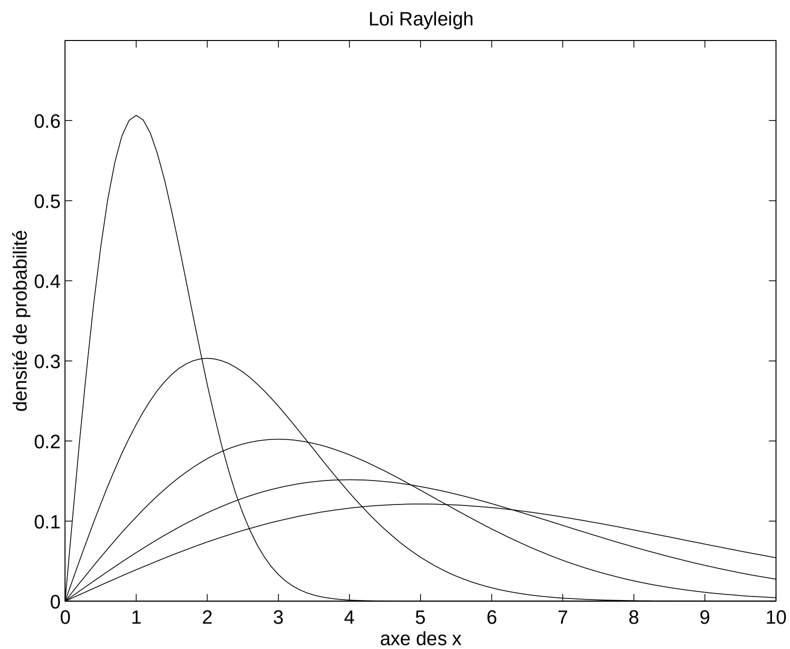


FIG. 8 – Distribution de la loi Rayleigh : $\sigma = 1, \ldots, 5$

loi Gamma : $X \sim \Gamma(\lambda, k)$, $f_X(x) = \frac{1}{\Gamma(k)} \lambda^k x^{k-1} e^{-\lambda x} \mathbb{1}_{\mathbb{R}+}(x)$ Rq : $\Gamma(k) = \int_{\mathbb{R}+} e^{-x} x^{k-1} dx$, si $k \in \mathbb{N}^*$, $\Gamma(k) = (k-1)!$.

Rq : $\mathcal{E}(\lambda) = \Gamma(\lambda, 1)$.

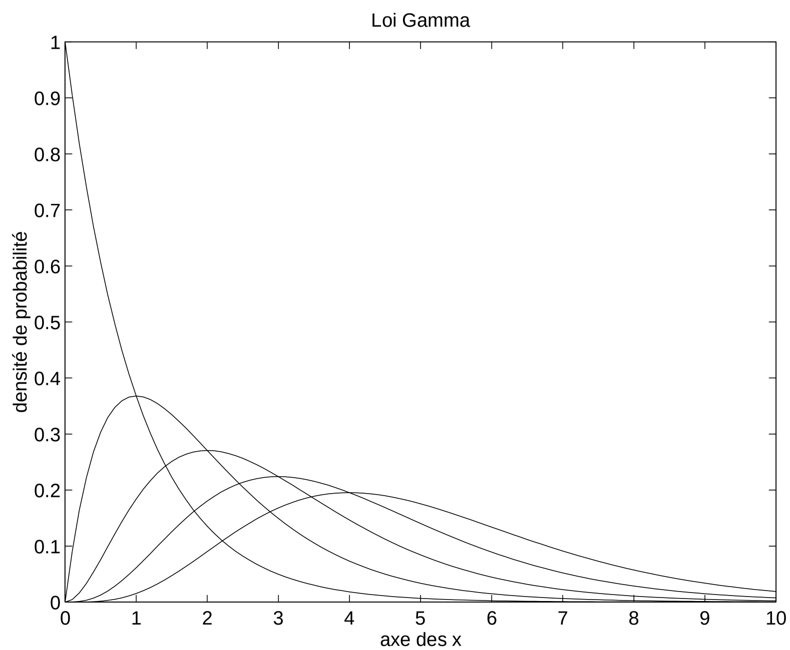


Fig. 9 – Distribution de la loi Gamma : $\Gamma(\lambda, k)$ avec $\lambda = 1$ et $k = 1 \ldots 5$

loi de Cauchy : $X \sim \mathcal{C}(\alpha, \beta)$, $f_X(x) = \frac{1}{\pi} \frac{\beta}{\beta^2 + (x-\alpha)^2}$ avec $x \in \mathbb{R}$

Rq :cette v.a. n'admet pas de moments d'ordre supérieur.

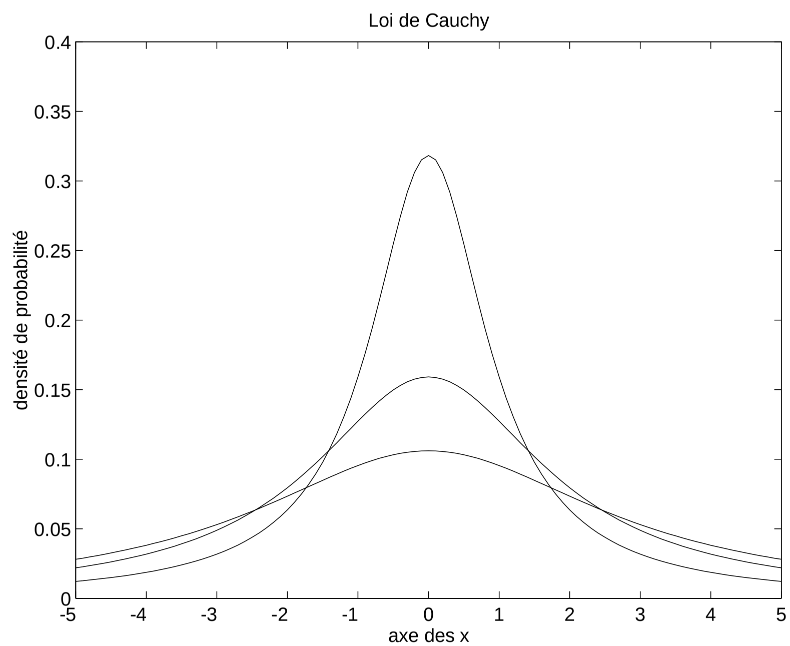


Fig. 10 – Distribution de la loi de Cauchy : $\mathcal{C}(\alpha, \beta)$ avec $\alpha = 0$ et $\beta = 1, 2$ ou 3

loi Beta : $X \sim \mathcal{B}(p,q)$ ($p > -1$ et $q > -1$), $f_X(x) = \frac{x^p(1-x)^q}{B(p,q)}$ avec $x \in [0,1]$, $B(p,q) = \int_0^1 x^p(1-x)^q dx$

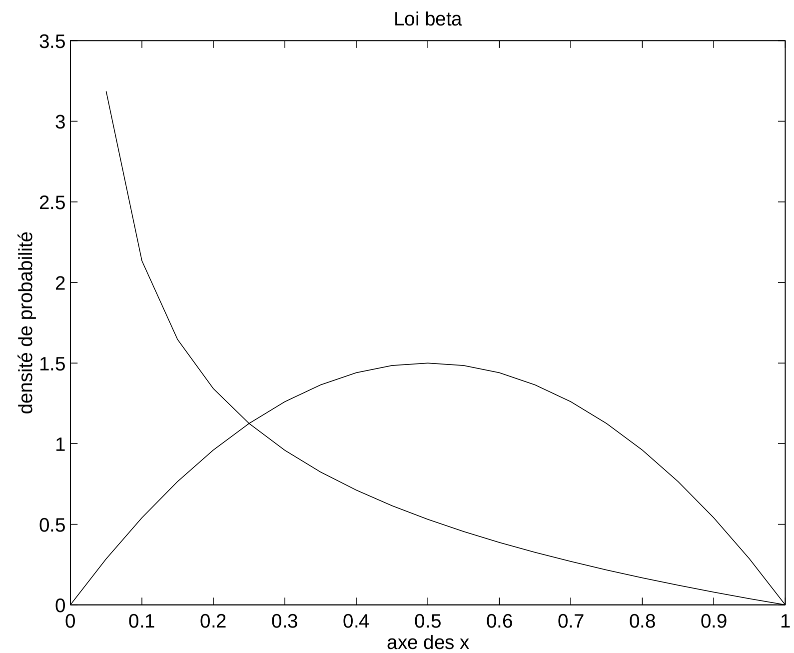


Fig. 11 – Distribution de la loi Beta : $\mathcal{B}(-1/2,1)$ et $\mathcal{B}(1,1)$

Exercice

Une variable aléatoire absolument continue X est telle que son domaine de définition $D_X = [-1, +1]$ et le graphe de sa densité de probabilité $f_X(x)$ forme avec l'axe des x un triangle isocèle.

1. Donner l'équation de $f_X(x)$ et de la fonction de répartition $F_X(x)$.
2. Calculer la probabilité $P(-0.4 \leq X \leq +0.5)$.
3. Calculer la probabilité $P(X > 0)$.
4. Donner la densité de probabilité ainsi que la fonction de répartition de la variable aléatoire $Y = |X|$.
5. Donner la densité de probabilité ainsi que la fonction de répartition de la variable aléatoire $Y = X^2$.

Fonctions et paramètres d'une variable aléatoire

Espérance mathématique

Définition 13 *Soit X une variable aléatoire définie sur un espace probabilisé $(\Omega, \mathcal{B}, P)$. On appelle espérance mathématique (ou valeur moyenne) de X, l'integrale, si elle existe (i.e. si $X \in L^1(\Omega, \mathcal{B}, P)$) :*

$$\mathbb{E}[X] = \int_\Omega X(\omega)\, dP(\omega)$$

<u>Propriétés importantes</u> :

1. $X \in L^1(\Omega, \mathcal{B}, P) \Longleftrightarrow \int_\Omega |X(\omega)|\, dP(\omega) < +\infty$
2. $\forall(\alpha, \beta) \in \mathbb{R}^2 \qquad \mathbb{E}[\alpha X + \beta Y] = \alpha\mathbb{E}[X] + \beta\mathbb{E}[Y]$
3. $\mathbb{E}[\alpha] = \alpha$
4. $X \geq 0 \quad \Longrightarrow \quad \mathbb{E}[X] \geq 0.$
5. $X \geq Y \quad \Longrightarrow \quad \mathbb{E}[X] \geq \mathbb{E}[Y].$
6. $\mathbb{E}[|X|] = 0 \quad \Longrightarrow \quad X = 0 \quad P-$ presque partout (c'est à dire que $X = 0$ partout sauf sur un ensemble de mesure nulle).

Théorème de transfert

Théorème 1 *Soit X une variable aléatoire définie sur un espace probabilisé $(\Omega, \mathcal{B}, P)$ et soit g une fonction déterministe mesurable :*

$$(\Omega, \mathcal{B}, P) \xrightarrow{X} (\mathbb{R}, \mathcal{B}(\mathbb{R}), P_X) \xrightarrow{g} (\mathbb{R}, \mathcal{B}(\mathbb{R}), P_X).$$

Si $g(X)$ est intégrable, alors :

$$\mathbb{E}[g(X)] = \int_\Omega g(X(\omega))\, dP(\omega) = \int_\mathbb{R} g(x)\, dP_X(x).$$

– Pour une v.a. discrète :

$$\mathbb{E}[g(X)] = \sum_{i \in X(\Omega)} g(i) P(X = i).$$

– Pour une v.a. absolument continue :

$$\mathbb{E}[g(X)] = \int_\mathbb{R} g(x) f_X(x)\, dx.$$

Moments d'ordre supérieur

Définition 14 *Soit $k \in \mathbb{N}^*$, si $X \in L^k(\Omega, \mathcal{B}, P)$ (c'est à dire que X^k est absolument intégrable), le moment d'ordre k de X est défini par :*

$$\mathbb{E}\left[X^k\right] = \int_\Omega X^k(\omega)\,dP(\omega) = \int_\mathbb{R} x^k\,dP_X(x)$$

Définition 15 *Le moment centré d'ordre k de X est défini par :*

$$\mathbb{E}\left[(X - E[X])^k\right]$$

Définition 16 *La variance d'une variable aléatoire est définie comme étant le moment centré d'ordre 2, soit :*

$$Var(X) = \sigma_X^2 = \mathbb{E}\left[(X - E[X])^2\right].$$

Rermarque : $\sigma_X^2 = \mathbb{E}\left[X^2\right] - (\mathbb{E}[X])^2$.

L'écart-type : $\sigma_X = \sqrt{Var(X)} = \sqrt{\mathbb{E}\left[X^2\right] - (\mathbb{E}[X])^2}$.

Inégalité de Markov et de Bienaymé-Tchebychev

Inégalité de Markov

Théorème 2 *SoitX une v.a. de moment d'ordre k fini, pour tout ε réel strictement positif, on a :*

$$P\left[|X| \geq \varepsilon\right] \leq \frac{\mathbb{E}\left[|X|^k\right]}{\varepsilon^k}.$$

Inégalité de Bienaymé-Tchebychev

Théorème 3 *Soit X une v.a. de moment d'ordre 2 fini, pour tout ε réel strictement positif, on a :*

$$P\left[|X - \mathbb{E}\left[X\right]| \geq \varepsilon\right] \leq \frac{\sigma_X^2}{\varepsilon^2}.$$

Fonction caractéristique

Définition 17 *On appelle fonction caractéristique d'une variable aléatoire X la fonction à valeurs complexes Φ_X définie sur $\mathbb{R}$ par :*

$$\Phi_X(u) = \mathbb{E}\left[e^{iuX}\right].$$

<u>Propriétés</u> :

1. Au coefficient -2π près, Φ_X est la transformée de Fourier de la mesure positive P_X sur $\mathbb{R}$.
2. L'existence de $\Phi_X(u)$ pour tout u résulte du fait que :

$$\forall u \in \mathbb{R}, \forall x \in \mathbb{R}, |e^{iux}| < 1.$$

3. $\Phi_X(0) = 1$.
4. $\forall u \in \mathbb{R}, |\Phi_X(u)| \leq 1$.
5. $\Phi_X(u)$ est continue.

6. $\forall u \in \mathbb{R}, [\Phi_X(u)]^* = \Phi_X(-u)$.
7. $\forall (a,b) \in \mathbb{R}^2, \forall u \in \mathbb{R}$, on a : $\Phi_{aX+b}(u) = e^{iub}\Phi_X(au)$.
8. La fonction $\Phi_X(u)$ caractérise la loi entièrement, soit :

$$(\Phi_X(u) = \Phi_Y(u), \forall u \in \mathbb{R}) \Longleftrightarrow (X \text{ et } Y \text{ ont la même loi}).$$

9. Si X est absolument continue et si Φ_X est absolument integrable, alors :

$$f_X(x) = \frac{1}{2\pi}\int_{\mathbb{R}} \Phi_X(u)e^{-iux}du.$$

10. Si X est une v.a. discrète et si Φ_X est absolument intégrable, alors :

$$P(X = k) = \frac{1}{2\pi}\int_0^{2\pi} \Phi_X(u)e^{-iuk}du.$$

11. Si les moments d'ordre k existent et si Φ_X est de classe $\mathcal{C}^k$ (k fois continuement dérivable) , alors :

$$\mathbb{E}\left[X^k\right] = \frac{1}{i^k}\Phi_X^{(k)}(0).$$

Fonction génératrice :

Définition 18 *Soit X une variable aléatoire entière. La fonction génératrice est définie par :*

$$
\begin{array}{rccl}
g_X : & \{z \in \mathbb{C} / |z| \leq 1\} & \longrightarrow & \mathbb{C} \\
& z & \longmapsto & g_X(z) = \mathbb{E}\left(z^X\right)
\end{array}
$$

<u>Remarques</u> :

- Lorsque $X \in L^k(\Omega, \mathcal{B}, P)$, alors g_X est k fois différentiable et on a :

$$g_X^{(k)}(1) = \sum_{n \geq k} n(n-1)\dots(n-k+1)P(X=n)$$

- $g_X(1) = 1$.
- $g_X^{'}(1) = \mathbb{E}(X)$.
- $g_X^{(2)}(1) = \mathbb{E}(X(X-1))$.
- ….

Cas d'une v.a. discrète

Soit $X \sim \mathcal{P}(\lambda)$, (Loi de poisson), $P(X=k) = e^{-\lambda} \frac{\lambda^k}{k!}$

Espérance mathématique : $\mathbb{E}[X] = \lambda$.

Variance : $\sigma_X^2 = \mathbb{E}\left[X^2\right] - (\mathbb{E}[X])^2 = \lambda$.

Fonction génératrice : $g_X(z) = \mathbb{E}\left(z^X\right) = e^{-\lambda} e^{+\lambda z}$

Fonction caractéristique : $\Phi_X(u) = \mathbb{E}\left[e^{iuX}\right] = e^{-\lambda} e^{+\lambda(\cos u + i \sin u)}$

et inversement : $P(X=k) = \frac{1}{2\pi} \int_0^{2\pi} \Phi_X(u) e^{-iuk} du = e^{-\lambda} \frac{\lambda^k}{k!}$.

Rappel (th des Résidus) : $f(z)$ complexe de pôle a d'ordre m :

$$Res(f,a) = \frac{1}{(m-1)!} \left(\left[(z-a)^m f(z) \right]^{(m-1)} \right)_{z=a}$$

Fonction de répartition : $F_X(x) = \sum_{k<x} e^{-\lambda} \frac{\lambda^k}{k!}$

Cas d'une v.a. absolument continue

Soit $X \sim \mathcal{E}(\lambda))$, $f_X(x) = \lambda e^{-\lambda x} \mathbb{1}_{\mathbb{R}+}(x)$

Espérance mathématique : $\mathbb{E}[X] = \int_0^{+\infty} x f_X(x) dx = \frac{1}{\lambda}$

Variance : $\sigma_X^2 = \frac{1}{\lambda^2}$

Fonction caractéristique : $\Phi_X(u) = \mathbb{E}\left[e^{iuX}\right] = \int_0^{+\infty} e^{iux} f_X(x) dx = \frac{\lambda}{\lambda - iu}$

Fonction de répartition : $F_X(x) = P(X < x) = (1 - e^{-\lambda x}) \mathbb{1}_{[0,+\infty[}(x)$

Exercice :

Calculer la moyenne et la variance d'une v.a. qui suit la loi Log Normale.

Vecteurs aléatoires

Définition 19 *Soit* $(\Omega, \mathcal{B})$ *un espace probabilisable donné et* $\mathbb{R}^p$ *muni de la tribu des boréliens* $\mathcal{B}(\mathbb{R}^p)$*. Un vecteur aléatoire (de dimension p) est une application mesurable de* $(\Omega, \mathcal{B})$ *dans* $(\mathbb{R}^p, \mathcal{B}(\mathbb{R}^p))$ *:*

$$\begin{array}{rccl} \vec{X}: & \Omega & \longrightarrow & \mathbb{R}^p \\ & \omega & \longmapsto & \vec{X}(\omega) = (X_1(\omega), \ldots, X_p(\omega)) \end{array}$$

Soit : $\forall B \in \mathcal{B}(\mathbb{R}^p), \vec{X}^{-1}(B) = \{\omega \in \Omega / \vec{X}(\omega) \in B\} \in \mathcal{B}$

loi de probabilité : $\forall B \in \mathcal{B}(\mathbb{R}^p), P_{\vec{X}}(B) = P\{\vec{X} \in B\} = P\{\vec{X}^{-1}(B)\}$

Espérance mathématique : $\mathbb{E}\left[\vec{X}\right] =^t (\mathbb{E}[X_1], \ldots, \mathbb{E}[X_p])$.

Remarques :

1. $\forall (\alpha, \beta) \in \mathbb{R}^2 \qquad \mathbb{E}\left[\alpha\vec{X} + \beta\vec{Y}\right] = \alpha\mathbb{E}\left[\vec{X}\right] + \beta\mathbb{E}\left[\vec{Y}\right]$
2. $\forall \vec{a} \in \mathbb{R}^p \qquad \mathbb{E}\left[\vec{a} \cdot \vec{X}\right] = \vec{a} \cdot \mathbb{E}\left[\vec{X}\right]$

Fonction de répartition

Définition 20 *La fonction de répartition $F_{\vec{X}}$ d'un vecteur aléatoire $\vec{X}$ est définie par :*

$$
\begin{array}{llll}
F_{\vec{X}} : & \mathbb{R}^p & \longmapsto & [0,1] \\
 & \vec{x} & \longmapsto & F_{\vec{X}}(\vec{x}) = F_{\vec{X}}(x_1, \ldots, x_p) = \\
 & & & P_{\vec{X}}\left(]-\infty, x_1[\times \ldots \times]-\infty, x_p[\right) = \\
 & & & P\left(\vec{X}^{-1}(]-\infty, x_1[\times \ldots \times]-\infty, x_p[)\right) = \\
 & & & P\left(\{X_1 < x_1\} \cap \ldots \cap \{X_p < x_p\}\right)
\end{array}
$$

loi marginale : $\forall B \in \mathcal{B}(\mathbb{R}), P_{X_k}(B) = P_{\vec{X}}\left(\underbrace{\mathbb{R} \times \ldots \mathbb{R}}_{k-1} \times B \times \underbrace{\mathbb{R} \times \ldots \mathbb{R}}_{p-k}\right)$

Rq : La loi conjointe permet de déterminer toutes les lois marginales mais la réciproque est fausse.

Soit (X, Y) un couple aléatoire tel que : ($0 \leq \alpha \leq 1/2$)

$P_{(X,Y)}$	$x = 0$	$x = 1$	$P_Y \downarrow$
$y = 0$	$P\{X = 0, Y = 0\} = \alpha$	$P\{X = 1, Y = 0\} = \frac{1}{2} - \alpha$	$P\{Y = 0\} = \frac{1}{2}$
$y = 1$	$P\{X = 0, Y = 1\} = \frac{1}{2} - \alpha$	$P\{X = 1, Y = 1\} = \alpha$	$P\{Y = 1\} = \frac{1}{2}$
$P_X \rightarrow$	$P\{X = 0\} = \frac{1}{2}$	$P\{X = 1\} = \frac{1}{2}$	

TAB. 1 – Exemple de loi dont les lois marginales ne permettent pas de retrouver la loi conjointe

La loi conjointe du couple (X, Y) dépend du paramètre α alors que chacune des deux lois marginales est indépendante de ce paramètre.

Définition 21 *Soit $\vec{X}$ un vecteur aléatoire de dimension p défini sur un espace probabilisé $(\Omega, \mathcal{B}, P)$. On dit que $\vec{X}$ admet une densité de probabilité sur $\mathbb{R}^p$ si et seulement si il existe une fonction $f_{\vec{X}}(x)$ de $\mathbb{R}^p$ dans $\mathbb{R}$, appelée densité de probabilité, telle que :*

$$\forall B \in \mathcal{B}(\mathbb{R}^p) \qquad P_{\vec{X}}(B) = \int_B f_{\vec{X}}(x_1, \ldots, x_p)\, dx_1 \ldots dx_p.$$

<u>Remarques</u> :

1. $\int_{\mathbb{R}} f_{\vec{X}}(x_1, \ldots, x_p)\, dx_1 \ldots dx_p = 1$.
2. Si $g(\vec{X}) \in L^1(\Omega, \mathcal{B}, P)$, $g : \mathbb{R}^p \to \mathbb{R}$, (théorème de transfert) :
$$\mathbb{E}\left[g(\vec{X})\right] = \int_{\mathbb{R}^p} g(\vec{x}) f_{\vec{X}}(\vec{x})\, d\vec{x}.$$
3. $f_{\vec{X}}(x_1, \ldots, x_p) = \frac{\partial^p F_{\vec{X}}}{\partial x_1 \ldots \partial x_p}(x_1, \ldots, x_p)$.
4. $f_{X_k}(x_k) = \int_{\mathbb{R}^{p-1}} f_{\vec{X}}(x_1, \ldots, x_p)\, dx_1 \ldots dx_{k-1} dx_{k+1} dx_p$.
5. Fonction caractéristique : $\forall \vec{u} \in \mathbb{R}^p$, $\Phi_{\vec{X}}(\vec{u}) = \mathbb{E}\left[e^{i\vec{u}\cdot\vec{X}}\right]$.

Changement de variables

$$\begin{array}{rccc} \phi : & \mathbb{R}^p & \longmapsto & \mathbb{R}^p \\ & \vec{x} & \longmapsto & \vec{y} \end{array}$$

Hypothèses :

1. ϕ est une bijection tel que ϕ et ϕ^{-1} soient de classe $\mathcal{C}^1$ et que :

$$\forall \vec{x} \in \mathbb{R}^p, \det J_{\phi^{-1}}(\vec{y}) = \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \dots & \frac{\partial x_1}{\partial y_p} \\ \vdots & \ddots & \vdots \\ \frac{\partial x_p}{\partial y_1} & \dots & \frac{\partial x_p}{\partial y_p} \end{vmatrix} \neq 0.$$

2. $\vec{X}$ un vecteur aléatoire de dimesion p de d.p. $f_{\vec{X}}(\vec{x})$
3. $\vec{Y} = \phi(\vec{X})$, de d.p. $f_{\vec{Y}}(\vec{y})$:

$$f_{\vec{Y}}(\vec{y}) = f_{\vec{X}}\left(\phi^{-1}(\vec{y})\right) \left|\det J_{\phi^{-1}}(\vec{y})\right|.$$

Ex : Soit (X, Y) un couple de v.a. définie sur $\mathbb{R}$ de d.p. $f_{(X,Y)}(x, y)$. Quelle est la d.p. de $V = X + Y$?

Variables aléatoires indépendantes

Définition 22 *Soit $X_1, \dots, X_p$ p variables aléatoires réelles définies sur l'espace probabilisé $(\Omega, \mathcal{B}, P)$. $X_1, \dots, X_p$ sont indépendantes si et seulement si :*

$$\forall B_1, \dots, B_p \in \mathcal{B}(\mathbb{R}), P\left(X_1 \in B_1, \dots, X_p \in B_p\right) = P_{X1}(B_1) \dots P_{Xp}(B_p)$$

Remarques :

1. $P_{\vec{X}}$ est alors une loi produit : $P_{\vec{X}} = P_{X1} \times \dots \times P_{Xp}$
2. Cas d'une d.p. sur $\mathbb{R}^p$: $f_{\vec{X}}(x_1, \dots, x_p) = f_{X_1}(x_1) \dots f_{X_p}(x_p)$.
3. Proposition : $X_1, \dots, X_p$ sont indépendantes si et seulement si $\forall g_1, \dots, g_p \in \mathcal{C}_b(\mathbb{R} \to \mathbb{R})$ (continues et bornées) $\mathbb{E}\left[g_1(X_1) \dots g_p(X_p)\right] = \mathbb{E}\left[g_1(X_1)\right] \dots \mathbb{E}\left[g_p(X_p)\right]$.
4. X et Y sont indépendantes :$f_{X+Y}(v) = \int_{\mathbb{R}} f_{(X,Y)}(u, v-u) du = \int_{\mathbb{R}} f_X(u) f_Y(v-u) du = f_X * f_Y(v)$

Lois conditionnelles

Soit (X,Y) un couple aléatoire de loi $P_{X,Y}$. On peut définir la loi conditionnelle de Y sachant $\{X=x\}$: $P_Y^{X=x}$.

Cas X discrète $(P\{X=x\}>0)$: $P\{Y \in B/X=n\} = \frac{P\{Y\in B \text{ et } X=n\}}{P\{X=n\}}$.

Cas continue $(f_X(x)>0)$: $f_Y^{X=x}(y) = \frac{f_{X,Y}(x,y)}{f_X(x)}$

Remarques :

1. $f_{X,Y}(x,y) \Rightarrow f_X(x) \Rightarrow f_Y^{X=x}(y)$.
2. $f_Y^{X=x}(y)$ et $f_X(x) \Rightarrow f_{X,Y}(x,y)$.
3. Espérance conditionnelle (est une v.a.) d'une fonction $g(X,Y)$ sachant $\{X=x\}$ par :
$$\mathbb{E}^{X=x}\left[g(X,Y)\right] = \int_{\mathbb{R}} g(x,y) f_Y^{X=x}(y)dy.$$
4. On a :
$$\mathbb{E}\left[g(X,Y)\right] = \mathbb{E}\left[\mathbb{E}^{X=x}\left[g(X,Y)\right]\right].$$

Covariance

Définition 23 *Soit $\vec{X} \in L^2(\Omega, \mathcal{B}, P)$, la covariance entre deux composantes X_i et X_j du vecteur aléatoire $\vec{X}$ est défine par :*

$$Cov\,(X_i, X_j) = \mathbb{E}\,[(X_i - \mathbb{E}\,[X_i])(X_j - \mathbb{E}\,[X_j])]$$

<u>Remarques</u> :

1. $Cov\,(X_i, X_i) = Var(X_i) = \sigma^2_{X_i}$.
2. $Cov\,(X_i, X_j) = \mathbb{E}\,[X_i X_j] - \mathbb{E}\,[X_i]\,\mathbb{E}\,[X_j]$
3. Si X_i et X_j sont indépendantes $\Rightarrow Cov\,(X_i, X_j) = 0$ (X_i et X_j sont non corrélées). La réciproque est en général fausse.
4. Le coefficient de corrélation entre X_i et X_j : $\rho_{X_i X_j} = \frac{Cov(X_i, X_j)}{\sigma_{X_i}\sigma_{X_j}}$
5. $|\rho_{X_i X_j}| \leq 1$.
6. $\rho_{X_i X_j} = \pm 1$ si et seulement si X_i et X_j sont proportionnelles.

Matrice de covariance

Définition 24 *Soit $\vec{X} \in L^2(\Omega, \mathcal{B}, P)$, la matrice de covariance du vecteur aléatoire $\vec{X}$ est défine par :*

$$K_{\vec{X}} = \left[Cov\left(X_i, X_j\right)\right]_{1 \leq i,j \leq p}$$

<u>Remarques</u> :

1. $K_{\vec{X}}$ est une matrice symétrique (${}^tK_{\vec{X}} = K_{\vec{X}}$).
2. $K_{\vec{X}}$ est une matrice positive ($\forall \vec{u} \in \mathbb{R}^p$, on a : ${}^t\vec{u}K_{\vec{X}}\vec{u} \geq 0$).
3. En notation matricielle, si $\vec{X}$ est un vecteur colonne, on peut écrire :
$$K_{\vec{X}} = \mathbb{E}\left[\left(\vec{X} - \mathbb{E}\left[\vec{X}\right]\right)^t \left(\vec{X} - \mathbb{E}\left[\vec{X}\right]\right)\right]$$
4. $\vec{Z} = A\vec{X} + b \Rightarrow K_{\vec{Z}} = AK_{\vec{X}}{}^tA$

Droite et courbe de régression

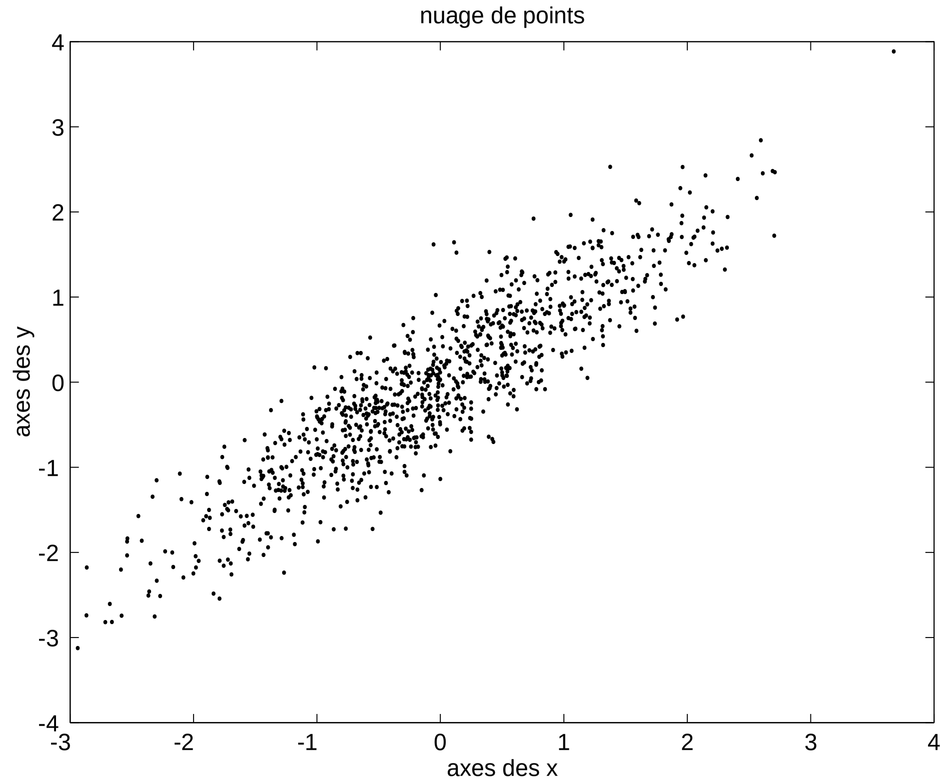


Fig. 12 – Nuage de points représentant 1000 réalisations d'un couple gaussien corrélées ($\sigma_X = \sigma_Y = 1$ et $\rho = 0.9$)

Définition 25 *On appelle droite de régression de Y sur X, la droite d'équation : "$y = ax + b$" où a et b minimisent la quantité* $\mathbb{E}\left[(Y - aX - b)^2\right]$.

(estimation linéaire de Y en fonction de X au sens des "moindres carrées")

$$y - \mathbb{E}[Y] = \frac{Cov(X,Y)}{\sigma_X^2}(x - \mathbb{E}[X])$$

Définition 26 *On appelle courbe de régression de Y sur X, la fonction d'équation : "$y = \phi(x)$" où ϕ minimise la quantité* $\mathbb{E}\left[(Y - \phi(X))^2\right]$.

L'équation de la courbe de régression de Y sur X est :

$$y = \phi(x) = \mathbb{E}^{X=x}[Y]$$

Exercice 1 :

Soit (X, Y) un couple de variables aléatoires indépendantes, telles que X est de loi uniforme sur $[0, 1]$ et Y est de loi exponentielle de paramètre 1 sur $\mathbb{R}^+$.

1. On pose $U = X + Y$ et $V = X - Y$, quelle est la densité de probabilité conjointe du couple (U, V) ?
2. Calculer les lois marginales de (U, V).
3. Calculer les espérances et les variances et covariances de (X, Y) et de (U, V).

Exercice 2 :

Deux personnes ont rendez-vous entre 19h et 20h. Les heures d'arrivées sont aléatoires de distribution uniforme sur $[0, 1]$ (en heure) et indépendantes. Soit Z la variable égale au temps d'attente de la première personne arrivée. Calculer la loi et l'espérance de Z (le temps moyen d'attente).

Exercice 3 : Soit $X = (X_1, X_2)$ un couple aléatoire dont les composantes sont indépendantes et uniforméments réparties sur l'intervalle $[-1, +1]$.
Soit $Y = (Y_1, Y_2)^T$ le couple déduit de X par la transformation linéaire $Y = AX$ où A est la matrice donnée par :

$$A = \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix}$$

1. Donner une interprétation géométrique de l'action de la matrice A sur le couple X.
2. Donner la densité de probabilité conjointe du couple $X = (X_1, X_2)$.
3. Donner la densité de probabilité conjointe du couple $Y = (Y_1, Y_2)^T$.
4. Calculer la densité de probabilité marginale $f_{Y_1}(y_1)$ de Y_1.
5. En déduire la densité de probabilité conditionnelle $f_{Y_2}^{Y_1=y_1}(y_2)$ de Y_2 sachant $Y_1 = y_1$.

Exercice 4 : Soit $(X_n)_{n\geq 1}$, une suite formée de v.a. indépendantes et identiquement distribuées, d'espérance μ, et N une v.a. de Poisson d'intensité λ, indépendante du processus X_n. Calculer :

$$\mathbb{E}\{\prod_{n=1}^{N} X_n\}$$

Vecteurs gaussiens

Définition 27 *Soit $\vec{X} \in L^2(\Omega, \mathcal{B}, P)$ un vecteur aléatoire. On dit que $\vec{X}$ suit une loi gaussienne ($\mathcal{N}(\vec{m}, K)$) si sa fonction caractéristique $\Phi_{\vec{X}}(\vec{u})$ est de la forme :*

$$\Phi_{\vec{X}}(\vec{u}) = e^{i\vec{u}\cdot\vec{m} - \frac{1}{2}{}^t\vec{u}K\vec{u}}$$

où $\vec{m} \in \mathbb{R}^p$ et K une matrice $p \times p$ symétrique positive.

Remarques :

1. $\vec{X}$ est gaussien $\Leftrightarrow$ toute combinaison linéaire des X_i est gaussienne.
2. K est diagonale $\Leftrightarrow$ les X_i sont indépendantes.
3. Pour un couple gaussien : l'indépendance $\Leftrightarrow$ la non corrélation.
4. Si $\det K \neq 0$, $\vec{X}$ est absolument continue sur $\mathbb{R}^p$ $\Rightarrow$:

$$f_{\vec{X}}(\vec{x}) = \frac{1}{(2\pi)^{p/2}\sqrt{\det K}} \exp -\frac{1}{2}{}^t(\vec{x}-\vec{m})K^{-1}(\vec{x}-\vec{m})$$

5. Soit A une matrice $q \times p$ et $\vec{b} \in \mathbb{R}^q$, on a :
$\vec{X} \sim \mathcal{N}(\vec{m}, K) \Longrightarrow \vec{Y} = A\vec{X} + \vec{b} \sim \mathcal{N}\left(A\vec{m} + \vec{b}, AK^tA\right)$

Exemple de couple de v.a. gaussiens dépendants et non corrélées

Soit $X \sim \mathcal{N}(0,1)$, ϵ une v.a. de Bernouilli ($\epsilon = \pm 1$), avec $P(\epsilon = 1) = 1/2$.
On définit, une v.a. $Y = \epsilon X$. X et Y sont dépendantes.

1. Calculer la d.p. de Y.
2. Vérifier que X et Y sont non corrélées.
3. Vérifier que (X, Y) est non gaussien. Conclure.

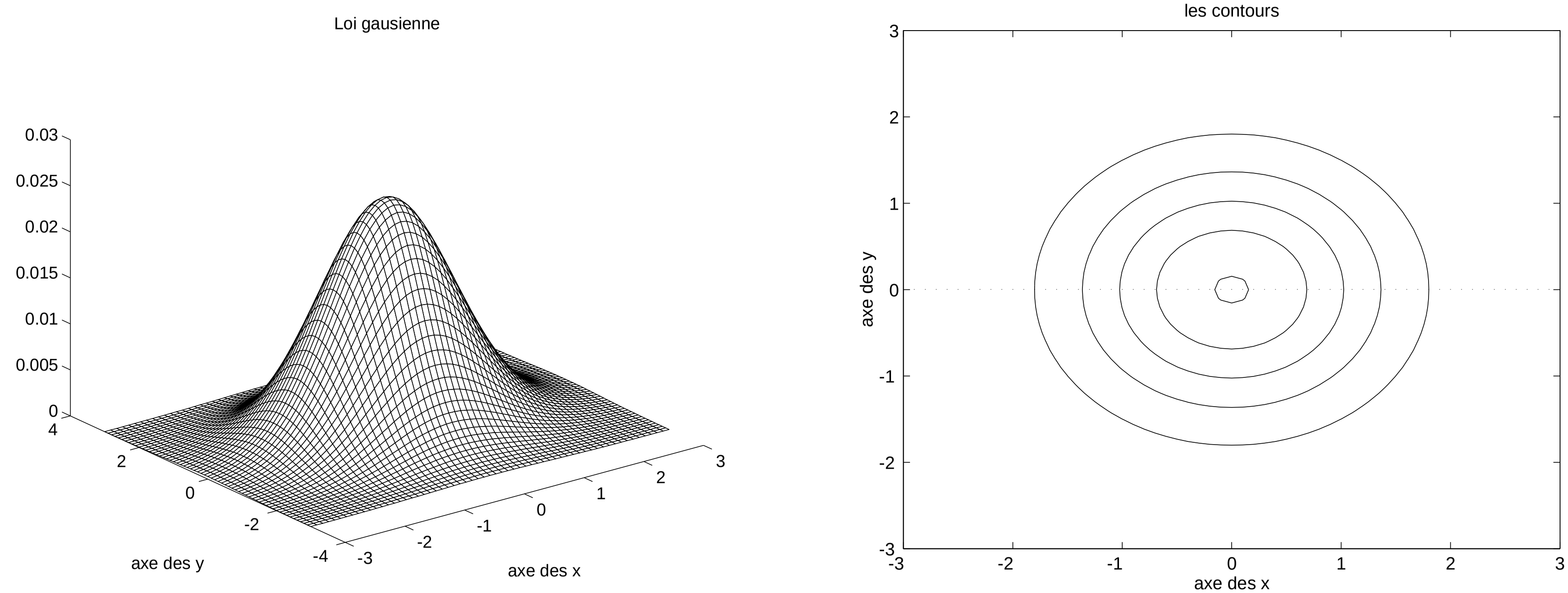
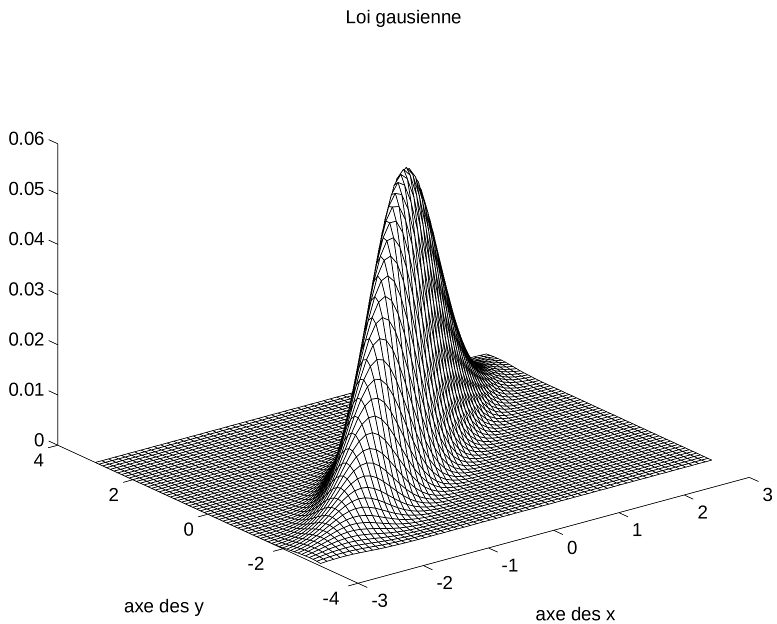


Fig. 13 – Loi gaussienne bidimensionnelle avec $\vec{m} = \vec{0}$, $\sigma_X = 1$, $\sigma_Y = 1$ et $\rho = 0$, ainsi que les contours correspondants



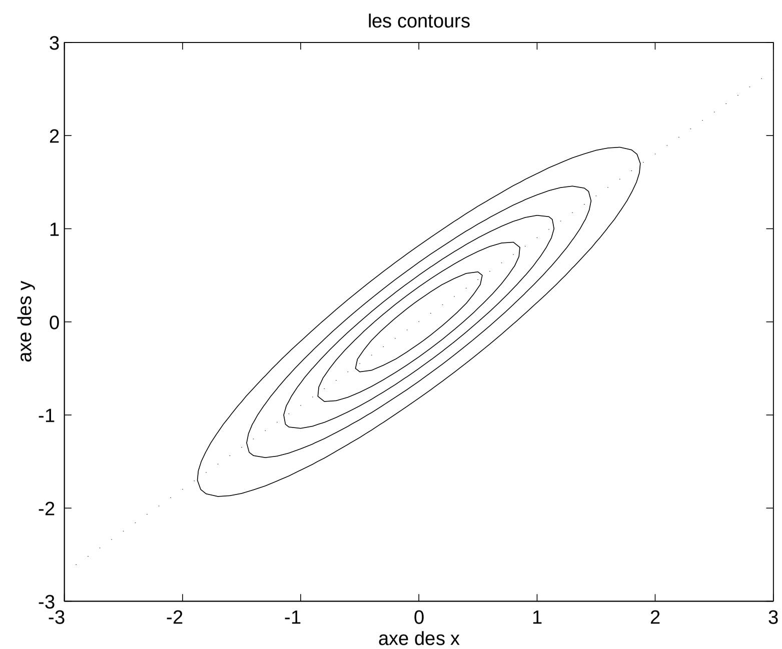


Fig. 14 – Loi gaussienne bidimensionnelle avec $\vec{m} = \vec{0}$, $\sigma_X = 1$, $\sigma_Y = 1$ et $\rho = 0.9$, ainsi que les contours correspondants

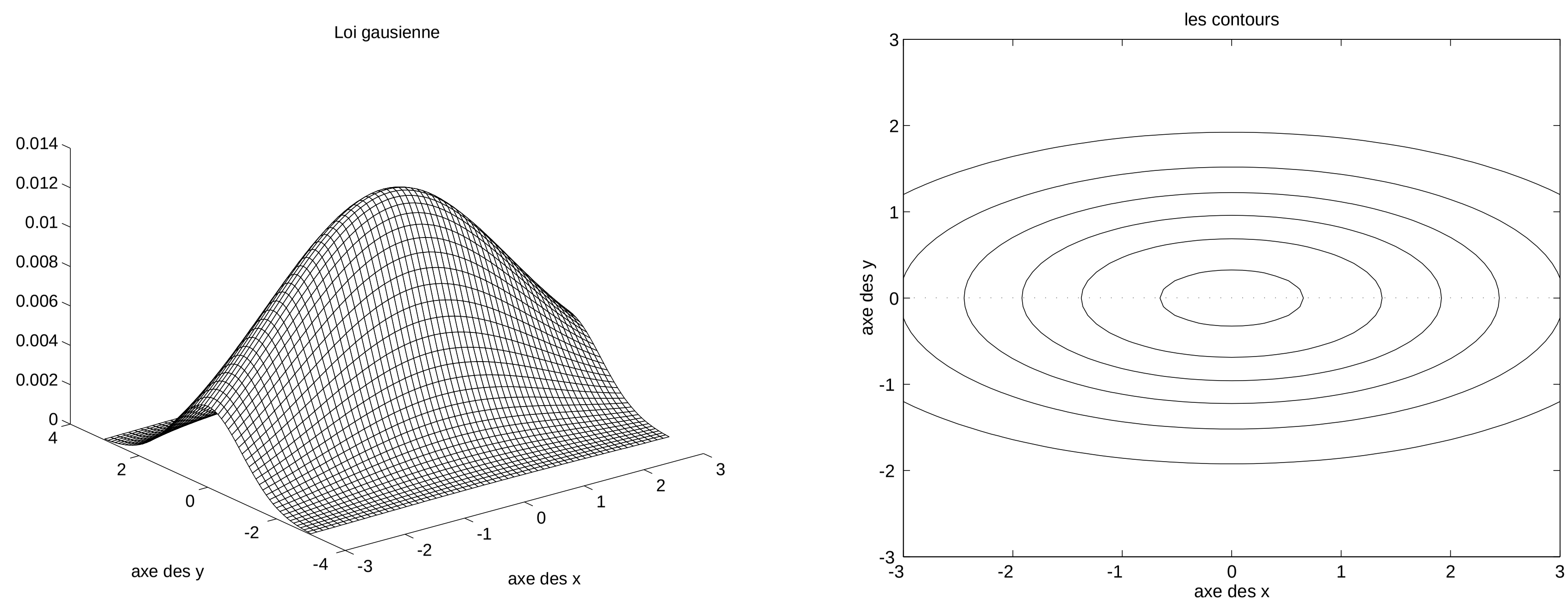


Fig. 15 – Loi gaussienne bidimensionnelle avec $\vec{m} = \vec{0}$, $\sigma_X = 2$, $\sigma_Y = 1$ et $\rho = 0$, ainsi que les contours correspondants

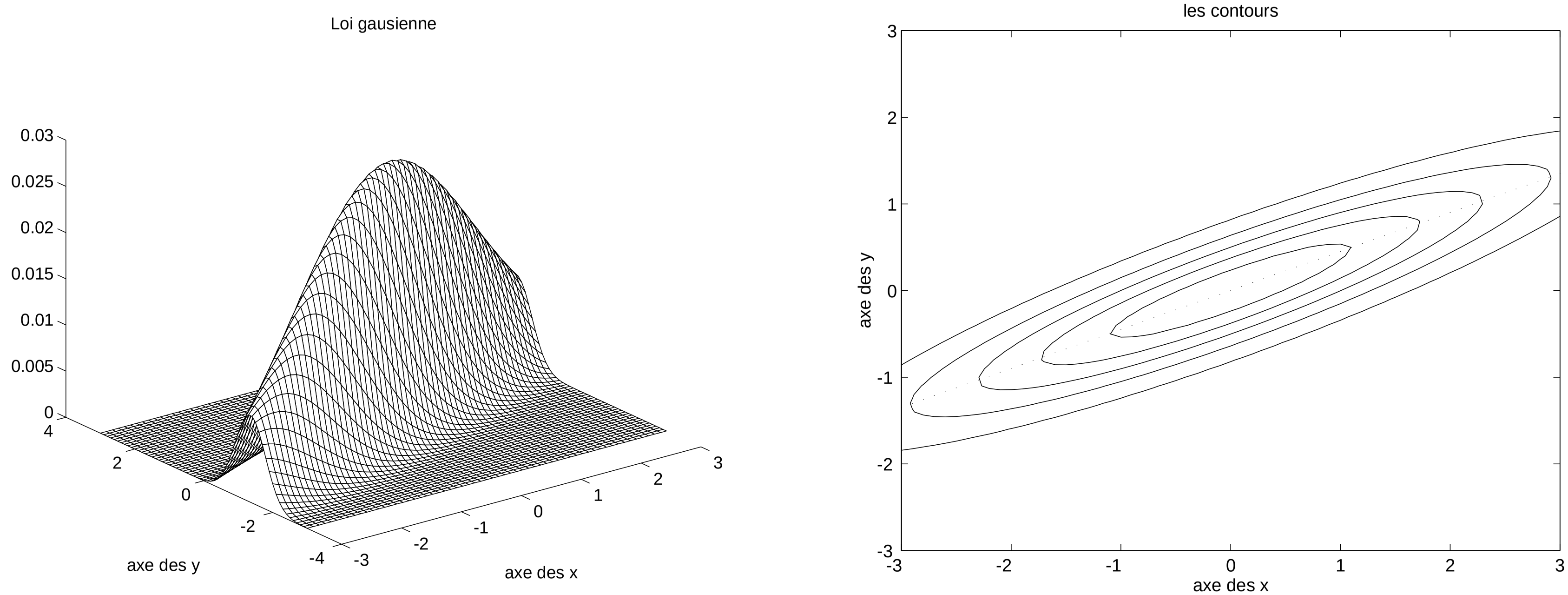


Fig. 16 – Loi gaussienne bidimensionnelle avec $\vec{m} = \vec{0}$, $\sigma_X = 2$, $\sigma_Y = 1$ et $\rho = 0.9$, ainsi que les contours correspondants

Exercice :

Soit $X = (X_1, X_2)^T$ un vecteur aléatoire réel centré, à deux dimensions, de loi gaussienne de matrice de covariance Λ donnée par $(X \sim N(\overrightarrow{0}, \Lambda))$:

$$\Lambda = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$$

avec $|\rho| < 1$,

On définit un nouveau couple aléatoire :

$$\begin{cases} Y_1 = X_1 - X_2 \\ Y_2 = X_1 + X_2 \end{cases}$$

1. Montrer que Y_1 et Y_2 sont indépendantes.
2. Donner l'expression de la matrice de covariance du couple (Y_1, Y_2).
3. En déduire la densité de probabilité du couple (Y_1, Y_2).

4. Calculer la densité de probabilité conditionnelle $f_{X_2}^{X_1=x_1}(x_2)$ de X_2 sachant $X_1 = x_1$. Quelle est cette loi ?
5. En déduire, sans calculs, les équations de la droite et de la courbe de régression de X_2 sur X_1.

Convergence en calcul de probabilité

Convergence en probabilité :

Définition 28 *Soit* $(X_n)_{n\geq 1}$ *une suite de variables aléatoires. La suite* $(X_n)_{n\geq 1}$ *converge stochastiquement (ou en probabilité) vers la variable aléatoire* X *(*$X_n \xrightarrow{st} X$*) si et seulement si :*

$$\forall \varepsilon > 0, \lim_{n\to+\infty} P\left[|X_n - X| > \varepsilon\right] = 0.$$

Convergence en moyenne quadratique :

Définition 29 *Soit* $(X_n)_{n\geq 1}$ *une suite de variables aléatoires. La suite* $(X_n)_{n\geq 1}$ *converge en moyenne quadratique vers la variable aléatoire* X *(*$X_n \xrightarrow{m.q} X$*) si et seulement si :*

$$\lim_{n\to+\infty} \mathbb{E}\left[|X_n - X|^2\right] = 0.$$

Théorème 4 *La convergence en moyenne quadratique implique la convergence en probabilité. Soit :*

$$(X_n \xrightarrow{m.q} X) \Longrightarrow (X_n \xrightarrow{st} X).$$

Loi faible des grands nombres :

Proposition 2 *Soit* $(X_n)_{n\geq 1}$ *une suite de variables aléatoires indépendantes, de même loi et de variance finie, alors :*

$$Y_n = \frac{X_1 + X_2 + \ldots + X_n}{n} \xrightarrow{m.q.} \mathbb{E}\left[X_1\right].$$

Convergence presque sûre :

Définition 30 *Soit* $(\Omega, \mathcal{B}, P)$ *un espace probabilisé. On dit que la suite* $(X_n)_{n\geq 1}$ *converge presque sûrement vers* X *(*$X_n \xrightarrow{p.s.} X$*) ssi :*

$$\exists \Omega_0 \subset \Omega / P\left(\Omega_0\right) = 1 \text{ et } \forall \omega \in \Omega_0, \lim_{n \to +\infty} X_n(\omega) = X(\omega).$$

Lemme de Borel-Cantelli :

Lemme 1 *Une condition suffisante de convergence presque sûre est :*

$$\forall \varepsilon > 0 \sum_{n=1}^{+\infty} P\left[|X_n - X| > \varepsilon\right] < \infty \Longrightarrow (X_n \xrightarrow{p.s.} X).$$

Loi forte des grands nombres :

Proposition 3 *Soit* $(X_n)_{n\geq 1}$ *une suite de variables aléatoires indépendantes, de même loi et de moyenne finie, alors :*

$$Y_n = \frac{X_1 + X_2 + \ldots + X_n}{n} \xrightarrow{p.s.} \mathbb{E}\left[X_1\right].$$

Convergence en Loi :

Définition 31 *Soit* $(X_n)_{n\geq 1}$ *une suite de variables aléatoires. La suite* (X_n) *converge en loi vers* X *(*$X_n \xrightarrow{\mathcal{L}} X$*) ssi les fonctions de répartition convergent simplement en tout point de continuité, soit :*

$$\forall x \in \mathbb{R}, \lim_{n\to+\infty} F_{X_n}(x) = F_X(x), \text{ si } F_X \text{ est continue en } x.$$

<u>Remarques</u> : Les trois conditions suivantes sont équivalentes :

1. $\forall x \in \mathbb{R}, \lim_{n\to+\infty} F_{X_n}(x) = F_X(x)$, si F_X est continue en x.
2. $\forall u \in \mathbb{R}, \lim_{n\to+\infty} \Phi_{X_n}(u) = \Phi_X(u)$
3. $\forall g \in \mathcal{C}$ (continue, borné) $\lim_{n\to+\infty} \mathbb{E}\left[g(X_n)\right] = \mathbb{E}\left[g(X)\right]$

Schéma mnémothechnique :

$$\begin{array}{ccccc} (X_n \xrightarrow{m.q} X) & \Longrightarrow & (X_n \xrightarrow{st} X) & \Longrightarrow & (X_n \xrightarrow{\mathcal{L}} X) \\ & & \Uparrow & & \\ & & (X_n \xrightarrow{p.s.} X) & & \end{array}$$

Thèorème de la limite centrale :

Théorème 5 *Soit* $(X_n)_{n\geq 1}$ *une suite de variables aléatoires indépendantes, de même loi et de variance finie* $(\sigma^2 = var(X_1^2))$, $(X_n \in L^2(\Omega, \mathcal{B}, P))$, *alors :*

$$\frac{S_n - \mathbb{E}[S_n]}{\sigma(S_n)} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$$

où $S_n = X_1 + \ldots + X_n$ *et donc* $\sigma(S_n) = \sqrt{n}\sigma$.

D'où $(m = \mathbb{E}[X_i])$:

$$P\left\{\frac{X_1 + \ldots + X_n - nm}{\sqrt{n}\sigma} \leq a\right\} \xrightarrow{n \to +\infty} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{-\frac{x^2}{2}} dx$$

Exemple d'utilisation du théorème central limite

On s'intéresse aux résultats d'un référendum au sein d'une large population, de taille N. Chaque individu peut voter **oui** ou **non** (on néglige les abstentions). Nous cherchons à déterminer le pourcentage de oui, qu'on notera p. Nous disposons pour cela d'une sous-population de taille n tirée au hasard. Si S_n désigne le nombre de oui, S_n suit la loi binomiale $\mathcal{B}(n,p)$.

$$P(S_n = k) = C_n^k p^k (1-p)^{n-k}$$

où p désigne la proportion de oui (= Probabilité qu'un individu pris au hasard vote oui). On a $S_n = X_1 + \ldots + X_n$, où les $X_k = 1$ pour oui et 0 pour non et forment une suite i.i.d. (Bernouilli $\mathcal{B}(p)$). $\hat{p} = S_n/n$. Trouver n tel que $p \in \left[\frac{S_n}{n} - \varepsilon, \frac{S_n}{n} + \varepsilon\right]$ avec une probabilité de se tromper au plus égale à α. Soit :

$$P\left\{\left|\frac{S_n}{n} - p\right| \leq \varepsilon\right\} \geq 1 - \alpha$$

D'après le théorème de la limite centrale, on a :

$$\frac{S_n - np}{\sqrt{n}\sigma} \sim \mathcal{N}(0,1)$$

avec $\sigma = \sqrt{p(1-p)}$. D'où :

$$\begin{aligned}
P\left\{\left|\frac{S_n}{n} - p\right| \leq \varepsilon\right\} &= P\left\{\left|\frac{S_n-np}{\sqrt{n}\sigma}\right| \leq \frac{\varepsilon\sqrt{n}}{\sigma}\right\} \\
&\approx \int_{-\frac{\varepsilon\sqrt{n}}{\sigma}}^{+\frac{\varepsilon\sqrt{n}}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \\
&= 2\int_0^{+\frac{\varepsilon\sqrt{n}}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \\
&= erf\left(\frac{\varepsilon\sqrt{n}}{\sqrt{2}\sigma}\right)
\end{aligned}$$

où :

$$erf(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$$

On veut donc que :

$$erf\left(\frac{\varepsilon\sqrt{n}}{\sqrt{2}\sigma}\right) \geq 1-\alpha$$

soit,

$$n \geq \frac{2\sigma^2}{\varepsilon^2}\left(erf^{-1}(1-\alpha)\right)^2$$

Notons que σ n'est pas connu mais que $\sigma \leq 1/2$. Le nombre d'échantillon minimale est donc :

$$n_{min} = \frac{1}{2\varepsilon^2}\left(erf^{-1}(1-\alpha)\right)^2$$

Application numérique :

$\alpha = 0.05$, $\varepsilon = 0.03$, on trouve que le nombre d'échantillon est 1067. On a dans ce cas le pourcentage de oui est $p = S_n/n$ à 3% près avec une probabilité de 0.95.

Exercice : Soit X une variable aléatoire qui suit la loi de Cauchy dont la densité de probabilité est donnée par :

$$f_X(x) = \frac{1}{\pi}\frac{1}{1+x^2}$$

Soit $(X_i)_{i\geq 1}$ une suite de variables aléatoires indépendantes et de même loi que celle de X. On définit $(S_n)_{n\geq 1}$ une suite de variables aléatoires définie par :

$$S_n = \sum_{i=1}^{n} X_i$$

1. Calculer la fonction caractéristique de la variable aléatoire $\frac{S_n}{n}$.
2. Montrer que $\frac{S_n}{n}$ converge en loi vers X.
3. Montrer que $\frac{S_n}{n^2}$ converge en probabilité (ou stochastiquement) vers 0.
4. Montrer que $\frac{S_n}{n^3}$ converge presque-sûrement vers 0 (Utiliser le lemme de Borel-Cantelli).