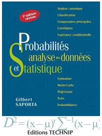
Probabilités et variables aléatoires

Université Hassiba Benbouali de Chlef

Références

- A. Monfort, Cours de statistique mathématique. Economica, 1997.

- G. Saporta, Probabilités, Analyse de Données et Statistique. Technip, 2006.



Rappels de probabilité

- Expérience aléatoire : expérience dont le résultat ne peut pas être prévu a priori
- Espace fondamental Ω : ensemble des résultats possibles
- Événement élémentaire ω : élément de l'espace fondamental
- Événement aléatoire ($\subset \Omega$) pouvant être VRAI ou FAUX suivant le résultat de l'expérience aléatoire

Liens entre ensembles et probabilités

ω	point de Ω	événement élémentaire
A	sous-ensemble de Ω	événement aléatoire
$\omega \in A$	ω appartient à A	ω réalise A
$A \subset B$	A est contenu dans B	A implique B
$A \cup B$	réunion de A et B	A ou B
$A \cap B$	intersection de A et B	A et B
$\bar{A}$	complémentaire de A	contraire de A
$\varnothing$	ensemble vide	événement impossible
Ω	ensemble plein	événement certain
$A \cap B = \phi$	A et B disjoints	A et B incompatibles

Rappels de probabilité

- $\mathcal{F}$ est une tribu sur Ω si c'est un ensemble de parties de Ω non-vide, stable par complémentaire et par union dénombrable
- $(\Omega, \mathcal{F})$ est un espace probabilisable, avec $\mathcal{F}$ est une tribu de parties de Ω

Définition 1.1

Une probabilité est une application $\mathbb{P} : \mathcal{F} \to [0,1]$ telle que

- $\mathbb{P}(\Omega) = 1$
- pour tous $A_1, \cdots, A_n$ disjoints ($A_i \cap A_j = \emptyset$),
  $$\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mathbb{P}\left(A_i\right)$$

- $(\Omega, \mathcal{F}, \mathbb{P})$ est un espace probabilisé

Propriétés

- $\mathbb{P}(\emptyset) = 0$
- $\mathbb{P}(\overline{A}) = 1 - \mathbb{P}(A)$
- $\mathbb{P}(A) \leq \mathbb{P}(B)$ si $A \subset B$
- $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$
- $\mathbb{P}\left(\bigcup_i A_i\right) \leq \sum_i \mathbb{P}\left(A_i\right)$
- pour tout évènement A de Ω, $0 \leq \mathbb{P}(A) \leq 1$

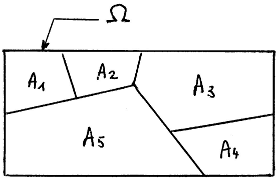
Rappels de probabilité

- Probabilité conditionnelle de l'événement A sachant B :
  $\mathbb{P}(A|B) = \dfrac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$
- Indépendance de 2 événements A et B :
  $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$
  $\mathbb{P}(A|B) = \mathbb{P}(A)$
  $\mathbb{P}(B|A) = \mathbb{P}(B)$
- Indépendance mutuelle des événements :
  $\mathbb{P}\left(\bigcap_i A_i\right) = \prod_i \mathbb{P}\left(A_i\right)$
- Théorème de Bayes pour 2 événements A et B quelconques :
  $\mathbb{P}(A|B) = \dfrac{\mathbb{P}(B|A)\mathbb{P}(A)}{\mathbb{P}(B)}$

Système complet d'événements

Définition 1.2

Un système complet d'événements, est une partition de Ω en événements $\{A_1, \ldots, A_n\}$ de probabilités strictement positives, $\mathbb{P}(A_i) > 0$ pour $1 \leqslant i \leqslant n$, et incompatibles deux à deux, i.e. avec $A_i \cap A_j = \emptyset$ pour $i \neq j$ et $\displaystyle\sum_{i=1}^{n} P(A_i) = 1$.

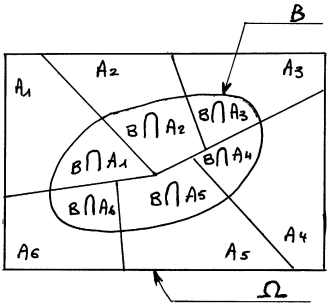


Formule de la probabilité totale

On suppose que les probabilités des événements inclus dans chacun des A_i sont connues et on va donc décomposer un événement quelconque B sur ce système

$$B = B \cap \Omega = B \cap \left(\bigcup_{i=1}^{n} A_i \right) = \bigcup_{i=1}^{n} (A_i \cap B)$$

Formule de la probabilité totale (suite)



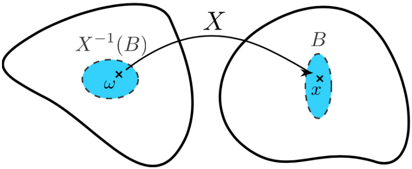
On aboutit ainsi à la formule de la probabilité totale :

$$\mathbb{P}(B) = \sum_{i=1}^{n} \mathbb{P}\left(A_i \cap B\right) = \sum_{i=1}^{n} \mathbb{P}\left(A_i\right) \mathbb{P}\left(B \mid A_i\right)$$

Variables aléatoires

Définition 2.1

Soient $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé et $(\mathcal{X}, \mathcal{A})$ un espace mesurable. On appelle variable aléatoire de Ω vers $\mathcal{X}$, toute fonction mesurable X de Ω vers $\mathcal{X}$.



Cette condition de mesurabilité de X assure que l'image réciproque par X de tout élément B de la tribu $\mathcal{A}$ possède une probabilité et permet ainsi de définir, sur $(\mathcal{X}, \mathcal{A})$, une mesure de probabilité, notée $\mathbb{P}_X$, par

$$\mathbb{P}_X(B) = \mathbb{P}\left(X^{-1}(B)\right) = \mathbb{P}\left(X \in B\right).$$

La mesure $\mathbb{P}_X$ est l'image, par l'application X, de la probabilité $\mathbb{P}$ définie sur $(\Omega, \mathcal{F})$.

Variables aléatoires

- Lorsque $(\mathcal{X}, \mathcal{A}) := (\mathbb{R}, \mathcal{B}(\mathbb{R}))$, on dit que X est une variable aléatoire réelle.
- Lorsque, pour un entier $d \geq 1$, $(\mathcal{X}, \mathcal{A}) := (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$, on dit que X est un vecteur aléatoire.

Variable aléatoire discrète

- Loi d'une variable aléatoire discrète X : $\mathbb{P}(X = x)$ pour $x \in \mathcal{X}$
- Propriétés :
  - $0 \leq \mathbb{P}(X = x) \leq 1$ pour tout $x \in \mathcal{X}$
  - $\sum_{x \in \mathcal{X}} \mathbb{P}(X = x) = 1$
  - Valeur de la probabilité de toute partie de $\mathcal{X}$:
    $\mathbb{P}(X \in A) = \sum_{x \in A} \mathbb{P}(X = x)$

Remarque

- Ne pas confondre une variable aléatoire, notée X, avec la valeur prise par cette variable, notée x.

Variable aléatoire continue

- Densité de probabilité d'une variable aléatoire discrète X : $f_X(x)$
- Propriétés :
  - $f_X(x) \geq 0$ pour tout x
  - $\int_{\mathbb{R}} f_X(x) \, \mathrm{d}x = 1$
  - $\mathbb{P}(X \in I) = \int_I f_X(x) \, \mathrm{d}x$ pour tout intervalle I de $\mathbb{R}$

Variable aléatoire continue (suite)

Remarque

- La description d'une loi continue diffère de celles des lois discrètes puisque pour une variable aléatoire continue X, la probabilité que X prenne une valeur bien précise x est nulle, $\mathbb{P}(X = x) = 0$.

Fonction de répartition

- Fonction de répartition d'une variable aléatoire réelle X :

$$\mathbb{F}_X \ : \ \mathbb{R} \to [0,1] \text{ définie par } \forall x \in \mathbb{R},\ \mathbb{F}_X(x) = \mathbb{P}(X \leq x)$$

- Propriétés :
  - $\mathbb{F}_X$ est une fonction croissante, continue à gauche
  - $\lim\limits_{x \longrightarrow -\infty} \mathbb{F}_X(x) = 0$ et $\lim\limits_{x \longrightarrow +\infty} \mathbb{F}_X(x) = 1$
  - Pour tous $a, b \in \mathbb{R}$ et $a < b$, $\mathbb{P}(a < X \leq b) = \mathbb{F}_X(b) - \mathbb{F}_X(a)$
  - Si X est discrète $\mathbb{F}_X(x) = \sum\limits_{x_i \leq x} \mathbb{P}(X = x_i)$
  - Si X est continue $\mathbb{F}_X(x) = \int_{-\infty}^{x} f_X(t)\ \mathrm{d}t$

Moments des variables aléatoires

- Espérance d'une variable aléatoire :
$$\mathbb{E}[X] = \begin{cases} \displaystyle\sum_{x \in \mathcal{X}} x\mathbb{P}(X = x) & \text{si } X \text{ est discrète} \\ \displaystyle\int_{\mathbb{R}} x f_X(x)\, \mathrm{d}x & \text{si } X \text{ est continue} \end{cases}$$
- Espérance d'une fonction d'une variable aléatoire :
$$\mathbb{E}[\varphi(X)] = \begin{cases} \displaystyle\sum_{x \in \mathcal{X}} \varphi(x)\mathbb{P}(X = x) & \text{si } X \text{ est discrète} \\ \displaystyle\int_{\mathbb{R}} \varphi(x) f_X(x)\, \mathrm{d}x & \text{si } X \text{ est continue} \end{cases}$$
- Le moment non centré d'ordre $r \in \mathbb{N}^*$, est $m_r(X) = \mathbb{E}[X^r]$.

Remarque

- L'espérance peut ne pas exister .

Propriétés

- $\mathbb{E}[a] = a$
- $\mathbb{E}[aX] = a\mathbb{E}[X]$
- $\mathbb{E}[aX + bY] = a\mathbb{E}[X] + b\mathbb{E}[Y]$
- $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$ si X et Y sont indépendantes

Variance d'une variable aléatoire

- La variance mesure la déviation moyenne autour de la moyenne espérée $\mathbb{E}[X]$, et est définie par

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

- Pour mesurer la dispersion d'une variable aléatoire X, on considère souvent en statistiques l'écart-type, lié à la variance par :

$$\sigma_X = \sqrt{\mathrm{Var}[X]}$$

- Propriétés
  - $\mathrm{Var}[X + a] = \mathrm{Var}[X]$
  - $\mathrm{Var}[aX] = a^2\mathrm{Var}[X]$
  - $\mathrm{Var}[X + Y] = \mathrm{Var}[X] + \mathrm{Var}[Y]$ si X et Y sont indépendantes

Loi Uniforme discrète $\mathcal{U}(n)$

- Loi d'une v. a. X prenant les valeurs $1, 2, \ldots, n$ avec la même probabilité

$$\mathbb{P}(X = x) = \frac{1}{n} \qquad \forall x \in \mathcal{X} = \{1, 2, \ldots, n\}$$

- Moments :

$$\mathbb{E}[X] = \frac{n+1}{2} \qquad \mathrm{Var}[X] = \frac{n^2 - 1}{12}$$

- Exemple : Réalisation d'un nombre (entre 1 et 6) après avoir jeté un dé

Loi de Bernoulli $\mathcal{B}(p)$

- Loi d'une v. a. X ne pouvant prendre que 2 valeurs 1 et 0 avec les probabilités p et $1-p$

$$\mathbb{P}(X = x) = p^x (1-p)^{1-x}$$

- Moments :

$$\mathbb{E}[X] = p \qquad \mathrm{Var}[X] = p(1-p)$$

- Exemple : Réalisation de pile (ou face) après avoir jeté une pièce

Loi Binomiale $\mathcal{B}(n, p)$

- Répétition de l'expérience de Bernoulli n fois, X est la somme des résultats des expériences

$$\mathbb{P}(X = x) = C_n^x p^x (1-p)^{n-x} \qquad \forall x \in \mathcal{X} = \{0, 1, \ldots, n\}$$

- Moments :

$$\mathbb{E}[X] = np \qquad \mathrm{Var}[X] = np(1-p)$$

- Exemple : Nombre de réalisations de pile après n essais

Loi de Poisson $\mathcal{P}(\lambda)$

- $\mathbb{P}(X = x) = \dfrac{\lambda^x e^{-\lambda}}{x!} \qquad \forall x \in \mathbb{N}$
- Moments :

$$\mathbb{E}[X] = \lambda \qquad \mathrm{Var}[X] = \lambda$$

- Exemple :
  - Nombre de personnes à la queue du bus après un intervalle de temps
  - Nombre d'appels téléphoniques pendant un intervalle de temps

Loi Uniforme $\mathcal{U}[a,b]$

- Cette loi modélise un phénomène uniforme sur un intervalle donné. Sa densité est alors,

$$f_X(x) = \frac{1}{b-a}\mathbb{1}_{[a,b]}(x)$$

- Fonction de répartition :

$$\mathbb{F}_X(x) = \frac{x-a}{b-a} \text{ pour } x \in [a,b]$$

- Moments :

$$\mathbb{E}[X] = \frac{a+b}{2} \qquad \mathrm{Var}[X] = \frac{(b-a)^2}{12}$$

Loi Exponentielle $\mathcal{E}(\lambda)$

- Densité
$$f_X(x) = \lambda e^{-\lambda x} \mathbb{1}_{[0,+\infty[}(x)$$
- Fonction de répartition :
$$\mathbb{F}_X(x) = 1 - e^{-\lambda x}$$
- Moments :
$$\mathbb{E}[X] = \frac{1}{\lambda} \qquad \mathrm{Var}[X] = \frac{1}{\lambda^2}$$
- Exemple :
  - Temps d'attente à la queue du bus
  - Durée de vie d'un composant électrique

Loi Normale $\mathcal{N}(m, \sigma^2)$

- Densité

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2\sigma^2}(x-m)^2}$$

- Moments :

$$\mathbb{E}[X] = m \qquad \mathrm{Var}[X] = \sigma^2$$

- Exemple :
  - Répartition des erreurs de mesure autour de la "vraie valeur"
- Propriétés : Soit $X_1 \sim \mathcal{N}(m_1, \sigma_1^2)$ et $X_2 \sim \mathcal{N}(m_2, \sigma_2^2)$ deux v.a. indépendantes, alors $X_1 + X_2 \sim \mathcal{N}\left(m_1 + m_2, \sigma_1^2 + \sigma_2^2\right)$.

Loi du Khi-deux $\chi^2_{(n)}$

- La loi du Khi-deux à n degrés de liberté, $\chi^2_{(n)}$ est la loi de la somme des carrés de n v. a. indépendantes de loi $\mathcal{N}(0,1)$.

$$f_Z(z) = \frac{\left(\frac{1}{2}\right)^{\frac{k}{2}}}{\Gamma\left(\frac{k}{2}\right)} z^{\frac{k}{2}-1} \mathrm{e}^{-\frac{z}{2}} \mathbb{1}_{z \geq 0}$$

  où $\Gamma(a) = \displaystyle\int_0^{+\infty} \mathrm{e}^{-x} x^{a-1} \, \mathrm{d}x$ est la "fonction gamma". Avec $\Gamma(z+1) = z\Gamma(z)$ et $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$.
- Moments :

$$\mathbb{E}[Z] = n \qquad \mathrm{Var}[Z] = 2n$$

- On l'utilise pour les variances empiriques d'échantillons gaussiens.

Loi de Student $\mathcal{T}(n)$

- La loi de Student à n degrés de liberté, $\mathcal{T}(n)$ est la loi du rapport $\dfrac{X}{\sqrt{Y/n}}$, où les variables aléatoires X et Y sont indépendantes , X de loi $\mathcal{N}(0,1)$, Y de loi $\chi^2_{(n)}$. Elle a pour densité :
$$f_T(t) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)}\left(1+\frac{t^2}{n}\right)^{-\frac{n+1}{2}}$$
- On l'utilise pour étudier la moyenne empirique d'un échantillon gaussien.

Loi Gamma $\Gamma(a, \lambda)$

- La loi gamma de paramètres $a > 0$ et $\lambda > 0$, notée $\Gamma(a, \lambda)$ a pour densité :

$$f_X(x) = \frac{\lambda^a}{\Gamma(a)} x^{a-1} \mathrm{e}^{-\lambda x} \mathbb{1}_{[0,+\infty[}(x)$$

- Moments :

$$\mathbb{E}[X] = \frac{a}{\lambda} \qquad \mathrm{Var}[X] = \frac{a}{\lambda^2}$$

- Propriétés : Soit $X_1 \sim \Gamma(a, \lambda)$ et $X_2 \sim \Gamma(b, \lambda)$ deux v.a. indépendantes, alors $X_1 + X_2 \sim \Gamma(a + b, \lambda)$.
- Pour $a = 1$, la loi $\Gamma(1, \lambda)$ est la loi exponentielle $\mathcal{E}(1)$.
- Pour n entier, $a = \frac{n}{2}$ et $\lambda = \frac{1}{2}$, la loi $\Gamma\left(\frac{n}{2}, \frac{1}{2}\right)$ est loi de Khi-deux à n degrés de liberté.

Loi Béta $\mathcal{B}(a,b)$

- La loi béta de paramètres $a > 0$ et $b > 0$, notée $\mathcal{B}(a,b)$ a pour densité :

$$f_X(x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1}\mathbb{1}_{[0,1]}(x)$$

- Moments :

$$\mathbb{E}[X] = \frac{a}{a+b} \qquad \mathrm{Var}[X] = \frac{ab}{(a+b)^2(a+b+1)}$$

Indicateurs de forme

- On appelle moment centré d'ordre $r \in \mathbb{N}^*$ la quantité, lorsqu'elle existe :

$$\mu_r(X) = \mathbb{E}\left[(X - \mathbb{E}[X])^r\right]$$

- Coefficient d'asymétrie de Fisher (skewness) $\gamma_1 = \frac{\mu_3}{\sigma^3}$
  - si $\gamma_1 = 0$, il y a symétrie ;
  - si $\gamma_1 > 0$, il y a étalement à droite ;
  - si $\gamma_1 < 0$, il y a étalement à gauche.
- Coefficient d'aplatissement de Fisher (kurtosis) $\gamma_2 = \frac{\mu_4}{\sigma^4} - 3$
  - si $\gamma_2 = 0$, la distribution est comparable à celle de la loi normale. On dit qu'elle est *mésokurtique*.

Indicateurs de forme (suite)

- si $\gamma_2 > 0$, la distribution est plus pointue que celle de la loi normale. On dit qu'elle est *leptokurtique* ;
- si $\gamma_2 < 0$, la distribution est plus aplatie que celle de la loinormale. On dit qu'elle est *platykurtique*.

Changement de variable

Soit X une v.a de densité $f_X > 0$ sur $\mathbb{R}$ et φ une fonction continue et bijective. On cherche à déterminer la loi de probabilité de la v.a. $Y = \varphi(X)$:

$$\mathbb{G}_Y(y) = \mathbb{P}(Y \leq y) = \mathbb{P}(\varphi(X) \leq y)$$

La densité de Y peut s'écrire sous la forme :

$$f_Y(y) = \frac{f_X(\varphi^{-1}(y))}{|\varphi'(\varphi^{-1}(y))|}$$

Loi du couple

- Si X et Y sont deux v.a. réelles continues, la loi de probabilité du couple (X,Y) est déterminée par sa fonction de répartition $\mathbb{F}_{X,Y}$, définie sur $\mathbb{R}^2$ par :

$$\mathbb{F}_{X,Y}(x,y) = \mathbb{P}\,(X \leq x, Y \leq y)$$

- Si $\mathbb{F}_{X,Y}$ est deux fois dérivable par rapport aux deux variables, alors la loi de (X,Y) est dite absolument continue, de densité $f_{X,Y}$ définie par :

$$f_{X,Y}(x,y) = \frac{\partial^2 \mathbb{F}_{X,Y}(x,y)}{\partial x \partial y}$$

Loi du couple (suite)

- La fonction de répartition se calcule alors par intégration :

$$\mathbb{F}_{X,Y}(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(u,v)\ \mathrm{d}u\mathrm{d}v$$

- Les fonctions de répartition marginales de X et Y sont définies à partir de la f.d.r. du couple en faisant tendre x, respectivement y, vers plus l'infini :

$$\mathbb{F}_X(x) = \lim_{y \longrightarrow +\infty} \mathbb{F}_{X,Y}(x,y) \text{ et } \mathbb{F}_Y(y) = \lim_{x \longrightarrow +\infty} \mathbb{F}_{X,Y}(x,y)$$

- Les densités marginales sont obtenues par :

$$f_X(x) = \int_{-\infty}^{+\infty} f_{X,Y}(x,y)\ \mathrm{d}y \text{ et } f_Y(y) = \int_{-\infty}^{+\infty} f_{X,Y}(x,y)\ \mathrm{d}x$$

Loi du couple (suite)

- Les lois conditionnelles sont définies par les densités conditionnelles :

$$f_{X|Y=y}(x) = \frac{f_{X,Y}(x,y)}{f_Y(y)} \text{ et } f_{Y|X=x}(y) = \frac{f_{X,Y}(x,y)}{f_X(x)}$$

- L'indépendance des v. a. X et Y se définit alors par :

$$f_{X,Y}(x,y) = f_X(x) \times f_Y(y) \quad \forall x \in \mathbb{R},\ \forall y \in \mathbb{R}$$

  On a bien entendu dans ce cas :

$$f_{X|Y=y}(x) = f_X(x) \text{ et } f_{Y|X=x}(y) = f_Y(y) \quad \forall x \in \mathbb{R},\ \forall y \in \mathbb{R}$$

Moments associés à un couple

- Si $\varphi : \mathbb{R}^2 \to \mathbb{R}$ est une application continue, l'espérance de $\varphi(X, Y)$ se calcule pour une loi de densité $f_{X,Y}$ par l'intégrale :

$$\mathbb{E}\left[\varphi(X, Y)\right] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \varphi(x, y) f_{X,Y}(x, y) \, \mathrm{d}x\mathrm{d}y$$

La covariance

- On définit la covariance par :

$$\mathrm{Cov}(X,Y) = \mathbb{E}\left[(X-\mathbb{E}[X])\,(Y-\mathbb{E}[Y])\right] = \mathbb{E}[XY]-\mathbb{E}[X]\mathbb{E}[Y]$$

- On a les propriétés suivantes :
  - $\mathrm{Var}[X+Y] = \mathrm{Var}[X] + \mathrm{Var}[Y] + 2\mathrm{Cov}(X,Y)$
  - $\mathrm{Cov}(X,Y) = \mathrm{Cov}(Y,X)$
  - $\mathrm{Cov}(aX+b, cY+d) = ac\mathrm{Cov}(X,Y)$
  - $\mathrm{Cov}(aX+bY, Z) = a\mathrm{Cov}(X,Z) + b\mathrm{Cov}(Y,Z)$
- On peut définir également le coefficient de corrélation linéaire par :

$$\rho_{X,Y} = \frac{\mathrm{Cov}(X,Y)}{\sigma_X \sigma_Y}$$

Remarque

- Dans le cas particulier où les v.a. X et Y sont indépendantes :

$$
\begin{aligned}
\mathbb{E}[XY] &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xy f_{X,Y}(x,y)\,\mathrm{d}x\mathrm{d}y \\
&= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xy f_X(x) f_Y(y)\,\mathrm{d}x\mathrm{d}y \\
&= \int_{-\infty}^{+\infty} x f_X(x)\,\mathrm{d}x \int_{-\infty}^{+\infty} y f_Y(y)\,\mathrm{d}y = \mathbb{E}[X]\mathbb{E}[Y]
\end{aligned}
$$

  et par conséquent : $\mathrm{Cov}(X,Y) = 0$.
- Il faut faire attention à la réciproque, généralement fausse, c'est-à-dire que si deux v.a. ont une covariance nulle elles ne sont pas forcément indépendantes, sauf dans le cas particulier où (X,Y) est un couple gaussien.

Changements de variables

- Le couple $(U, V) = \varphi(X, Y)$ admet la densité de probabilité :

$$f_{U,V}(u, v) = f_{X,Y}\left(\varphi^{-1}(u, v)\right)\left|\mathrm{J}_{\varphi^{-1}}\right|$$

  On rappelle que $\mathrm{J}_{\varphi^{-1}}$ est le jacobien de φ^{-1}, c'est-à-dire le déterminant de la matrice jacobienne. Si $\varphi^{-1}(u, v) = (\psi_1(u, v), \psi_2(u, v))$, alors :

$$\mathrm{J}_{\varphi^{-1}} = \det \begin{pmatrix} \dfrac{\partial \psi_1(u, v)}{\partial u} & \dfrac{\partial \psi_1(u, v)}{\partial v} \\ \dfrac{\partial \psi_2(u, v)}{\partial u} & \dfrac{\partial \psi_2(u, v)}{\partial v} \end{pmatrix}$$

Exemple

Supposons que X et Y suivent la même loi normale $\mathcal{N}(0, 1)$ et que X et Y soient indépendantes.

- Déterminer la densité du couple (R, Θ) obtenu par le passage en coordonnées polaires.

Loi d'une somme

La loi de la v.a. $Z = X + Y$ avec X et Y indépendantes est donnée par :

$$f_Z(z) = \int_{-\infty}^{+\infty} f_X(x) f_Y(z-x) \, \mathrm{d}x = \int_{-\infty}^{+\infty} f_X(z-y) f_Y(y) \, \mathrm{d}y$$

et $f_Z(z)$ s'appelle alors le produit de convolution de f_X et f_Y.

Vecteurs aléatoires

Soit $X = \begin{pmatrix} X_1 \\ \vdots \\ X_d \end{pmatrix} \in \mathbb{R}^d$ un vecteur aléatoire dont les composantes sont de carré intégrable.

- Le vecteur moyenne de X est défini par : $\mathbb{E}[X] = \begin{pmatrix} \mathbb{E}[X_1] \\ \vdots \\ \mathbb{E}[X_d] \end{pmatrix}$
- et sa matrice de covariance par :

$$\begin{aligned}
\mathrm{Var}[X] &= \mathbb{E}\left[(X - \mathbb{E}[X])(X - \mathbb{E}[X])^\top\right] \\
&= \begin{pmatrix} \mathrm{Var}\,[X_1] & \mathrm{Cov}\,(X_1, X_2) & \cdots & \mathrm{Cov}\,(X_1, X_d) \\ \mathrm{Cov}\,(X_2, X_1) & \mathrm{Var}\,[X_2] & \cdots & \mathrm{Cov}\,(X_2, X_d) \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{Cov}\,(X_d, X_1) & \mathrm{Cov}\,(X_d, X_2) & \cdots & \mathrm{Var}\,[X_d] \end{pmatrix}
\end{aligned}$$

Remarque

Remarque

- Pour tout $1 \leq i, j \leq d$, la covariance entre X_i et X_j est donnée par

$$\begin{aligned} \operatorname{Cov}(X_i, X_j) &= \mathbb{E}\left[(X_i - \mathbb{E}[X_i])(X_j - \mathbb{E}[X_j])\right] \\ &= \mathbb{E}[X_i X_j] - \mathbb{E}[X_i]\,\mathbb{E}[X_j] \end{aligned}$$

- Il est facile de voir que l'on a $\operatorname{Cov}(X_i, X_i) = \operatorname{Var}[X_i]$.
- De plus, si les variables X_i et X_j sont indépendantes on a $\operatorname{Cov}(X_i, X_j) = 0$.
- En général, la réciproque est fausse, sauf pour des vecteurs gaussiens.

Théorème 5.1

Si X est un vecteur (colonne) aléatoire de $\mathbb{R}^d$ de vecteur moyenne $\boldsymbol{m}$ et de matrice de covariance $\boldsymbol{\Sigma}$ Alors si A est une matrice réelle $k \times d$, le vecteur aléatoire AX de $\mathbb{R}^k$ a pour vecteur moyenne $A\boldsymbol{m}$ et pour matrice de covariance $A\boldsymbol{\Sigma}A^\top$.

Preuve

C'est une simple conséquence de la linéarité de l'espérance. Pour la moyenne on a :

$$\mathbb{E}[AX] = A\mathbb{E}[X] = A\boldsymbol{m}$$

et pour la matrice de covariance

$$\begin{aligned}
\mathrm{Var}[AX] &= \mathbb{E}\left[(AX - \mathbb{E}[AX])(AX - \mathbb{E}[AX])^\top\right] \\
&= \mathbb{E}\left[A(X - \mathbb{E}[X])\left(A(X - \mathbb{E}[X])\right)^\top\right] \\
&= A\mathbb{E}\left[(X - \mathbb{E}[X])(X - \mathbb{E}[X])^\top\right] A^\top = A\boldsymbol{\Sigma}A^\top
\end{aligned}$$

Densité gaussienne en dimension d

Définition 5.2

Un vecteur aléatoire de $\mathbb{R}^d$ est un vecteur aléatoire gaussien si et seulement si toute combinaison linéaire de ses composantes est une variable aléatoire réelle gaussienne, i.e. :

$$\forall a \in \mathbb{R}^d \qquad a^\top X \sim \mathcal{N}(m, \sigma^2).$$

Soit $X \sim \mathcal{N}_d(\boldsymbol{m}, \boldsymbol{\Sigma})$ un vecteur gaussien en dimension d. Si $\boldsymbol{\Sigma}$ est inversible, la densité de X est donc donnée par

$$f(\mathbf{x}) = \frac{1}{|\boldsymbol{\Sigma}|^{\frac{1}{2}}(2\pi)^{\frac{d}{2}}} \exp\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{m})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{m})\right),$$

Pour tout vecteur gaussien X de $\mathbb{R}^d$, les trois propriétés suivantes sont équivalentes :

1. Les composantes $X_1, \ldots, X_d$ sont mutuellement indépendantes.
2. Les composantes $X_1, \ldots, X_d$ sont deux à deux indépendantes.
3. La matrice de covariance $\mathbf{\Sigma}$ de X est diagonale.

Fonctions associées aux lois

- Soit X une variable aléatoire entière et positive, la fonction génératrice de X est la série entière

$$G_X(s) = \mathbb{E}[s^X] = \sum_{k=0}^{\infty} \mathbb{P}(X = k) s^k$$

On a

$$\mathbb{E}[X] = G'_X(1) \text{ et } \mathrm{Var}[X] = G''_X(1) + G'_X(1) - G'_X(1)^2.$$

- La fonction génératrice des moments d'une variable aléatoire X est la fonction M_X définie par $M_X(t) = \mathbb{E}\left[\mathrm{e}^{tX}\right]$. On a

$$\mathbb{E}\left[X^r\right] = M_X^{(r)}(0) = \left.\frac{\mathrm{d}^r M_X(t)}{\mathrm{d}t^r}\right|_{t=0}.$$

Fonctions associées aux lois (suite)

- La fonction caractéristique d'une variable aléatoire X est la fonction ϕ_X définie par $\phi_X(t) = \mathbb{E}\left[\mathrm{e}^{\mathrm{i}tX}\right]$. On a

$$\phi_X^{(r)}(0) = \mathrm{i}^r m_k(X).$$