

# Probabilités pour Statistique

Frédéric Bertrand<sup>1</sup>

<sup>1</sup>IRMA, Université de Strasbourg  
Strasbourg, France

ESIEA 4ème Année 01-03-2012

Ce chapitre s'appuie essentiellement sur :

le livre de D.Foata et A.Fuchs,  
professeurs à l'Université Louis Pasteur de Strasbourg,  
« Calcul des probabilités »,  
Masson, 1996.

# Sommaire

- 1 Variables aléatoires réelles
- 2 Variables aléatoires discrètes
- 3 Principales lois de probabilité discrètes
- 4 Variables aléatoires absolument continues
- 5 Loi normale de paramètres  $\mu$  et  $\sigma^2$
- 6 Vecteurs gaussiens, lois associées

## Définition et notation

Sur  $\mathbb{R}$  nous considérons généralement la tribu engendrée par les intervalles ouverts : nous l'appelons la tribu de Borel ou tribu borélienne. Nous la noterons  $\mathfrak{B}$ .

## Définition

*Soit  $(\Omega, \mathfrak{A}, \mathbb{P})$  un espace probabilisé. Une application*

$$X : \Omega \rightarrow \mathbb{R}$$

*est une variable aléatoire réelle, si pour tout  $A \in \mathfrak{B}$ , l'ensemble*

$$X^{-1}(A) = \{\omega \in \Omega \mid X(\omega) \in A\}$$

*est dans  $\mathfrak{A}$ .*

## Notation

Nous notons généralement

$$\{X \in A\}$$

l'ensemble  $X^{-1}(A)$ . De même nous noterons  $\{X \leq a\}$  l'événement  $X^{-1}(] - \infty; a])$  et plus généralement pour une relation  $R$  quelconque, nous noterons  $\{X R a\}$  l'événement  $X^{-1}(\{x \in \mathbb{R} \mid x R a\})$ .

Par ailleurs, pour les conjonctions nous utiliserons une notation analogue en séparant les conditions par des virgules. Par exemple  $\{X \in A\} \cap \{X \in B\}$  sera noté  $\{X \in A, X \in B\}$ .

## Exemple

Si  $A$  est un événement de  $\mathfrak{A}$ , l'application  $X : \Omega \rightarrow \mathbb{R}$  définie par

$$X(\omega) = \begin{cases} 1 & \text{si } \omega \in A \\ 0 & \text{si } \omega \notin A \end{cases}$$

est une variable aléatoire réelle, appelée la fonction indicatrice de l'événement  $A$  et nous la notons  $\mathbf{1}_A$ . En effet, pour tout  $B \in \mathfrak{B}$

$$\{X \in B\} = \begin{cases} \emptyset & \text{si } \{0; 1\} \cap B = \emptyset \\ A & \text{si } \{0; 1\} \cap B = \{1\} \\ \bar{A} & \text{si } \{0; 1\} \cap B = \{0\} \\ \Omega & \text{si } \{0; 1\} \cap B = \{0; 1\}. \end{cases}$$

## Théorème

*Soit  $(\Omega, \mathfrak{A}, \mathbb{P})$  un espace probabilisé et  $X : \Omega \rightarrow \mathbb{R}$  une variable aléatoire sur cet espace. Il existe sur l'espace probabilisable  $(\mathbb{R}, \mathfrak{B})$  une mesure de probabilité  $\mathbb{P}_X$  définie par*

$$\mathbb{P}_X[B] = \mathbb{P}[\{X \in B\}]$$

*pour tout  $B \in \mathfrak{B}$ . Nous l'appelons la loi de probabilité d'une variable aléatoire réelle  $X$ .*

## Exemple

Cherchons la loi de probabilité de l'indicatrice  $\mathbf{1}_A$  d'un événement  $A$ .

$$\mathbb{P}_{\mathbf{1}_A}[B] = \mathbb{P}[\{\mathbf{1}_A \in B\}] = \begin{cases} 0 & \text{si } \{0, 1\} \cap B = \emptyset \\ \mathbb{P}[A] & \text{si } \{0, 1\} \cap B = \{1\} \\ 1 - \mathbb{P}[A] & \text{si } \{0, 1\} \cap B = \{0\} \\ 1 & \text{si } \{0, 1\} \cap B = \{0, 1\}. \end{cases}$$



## Définition

Nous appelons **fonction de répartition d'une variable aléatoire réelle**  $X$  sur un espace probabilisé  $(\Omega, \mathfrak{A}, \mathbb{P})$ , l'application  $F_X$  de  $\mathbb{R}$  dans  $\mathbb{R}$  définie par

$$F_X(x) = \mathbb{P}_X([-\infty; x]) = \mathbb{P}\{X \leq x\}.$$

## Exemple

La fonction de répartition de l'indicatrice  $\mathbf{1}_A$  d'un événement  $A$  est l'application définie par

$$F_{\mathbf{1}_A}(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 - \mathbb{P}[A] & \text{si } 0 \leq x < 1 \\ 1 & \text{si } x \geq 1. \end{cases}$$

## Théorème

- 1  $0 \leq F_X(x) \leq 1$  ;
- 2  $F_X$  est une fonction croissante ;
- 3  $\lim_{x \rightarrow -\infty} F_X(x) = 0$  et  $\lim_{x \rightarrow +\infty} F_X(x) = 1$  ;
- 4  $F_X$  est continue à droite en chaque point.

## Définition

Deux variables aléatoires réelles  $X$  et  $Y$  sont **indépendantes** si pour tout  $A \in \mathfrak{B}$  et tout  $B \in \mathfrak{B}$  les événements  $\{X \in A\}$  et  $\{Y \in B\}$  sont indépendants, autrement dit, si

$$\mathbb{P}[\{X \in A, Y \in B\}] = \mathbb{P}[\{X \in A\}] \times \mathbb{P}[\{Y \in B\}].$$

## Théorème

Deux variables  $X$  et  $Y$  sont indépendantes si et seulement si pour tout  $x, y \in \mathbb{R}$

$$\mathbb{P}[\{X \leq x, Y \leq y\}] = F_X(x) \times F_Y(y).$$

# Sommaire

- 1 Variables aléatoires réelles
- 2 Variables aléatoires discrètes**
- 3 Principales lois de probabilité discrètes
- 4 Variables aléatoires absolument continues
- 5 Loi normale de paramètres  $\mu$  et  $\sigma^2$
- 6 Vecteurs gaussiens, lois associées

## Définition

Soit  $x$  un nombre réel. L'application

$$\varepsilon_x : \mathfrak{B} \rightarrow \mathbb{R}$$
$$A \mapsto \begin{cases} 1 & \text{si } x \in A \\ 0 & \text{si } x \notin A \end{cases}$$

est une loi de probabilité sur  $(\mathbb{R}, \mathfrak{B})$ , qui est entièrement portée par le singleton  $\{x\}$ . Nous l'appelons **la loi singulière en  $x$** .

## Définition

Soit  $(x_n)_{n \in I}$  une suite finie ou dénombrable d'éléments de  $\mathbb{R}$  et  $(\alpha_n)_{n \in I}$  une suite finie ou dénombrable de nombres réels strictement positifs, tels que

$$\sum_{n \in I} \alpha_n = 1.$$

Alors l'application de  $\mathfrak{B}$  dans  $\mathbb{R}$  définie par

$$\mathbb{P} = \sum_{n \in I} \alpha_n \varepsilon_{x_n}$$

est une mesure de probabilité sur  $(\mathbb{R}, \mathfrak{B})$  appelée loi de probabilité discrète de masses  $\alpha_n$  portées par les points  $x_n$ .

Nous avons

$$\mathbb{P}[B] = \sum_{n \in I} \alpha_n \varepsilon_{x_n}(B) = \sum_{x_n \in B} \alpha_n.$$

La probabilité de  $B$  est donc la somme des probabilités des points  $x_n$  contenus dans  $B$ .

Ainsi, pour une telle loi, la mesure de probabilité est entièrement concentrée en des masses ponctuelles et sa fonction de répartition est une fonction étagée.



## Définition

*Une variable aléatoire réelle est discrète si sa loi de probabilité est discrète.*

## Définition

L'espérance mathématique d'une variable aléatoire discrète  $X$  de loi  $\mathbb{P}_X = \sum_{n \in I} \alpha_n \varepsilon_{x_n}$  ( $I = \{0; 1; \dots; N\}$  ou  $I = \mathbb{N}$ ) est le nombre réel

$$\mathbb{E}[X] = \sum_{n \in I} x_n \mathbb{P}_X[\{x_n\}] = \sum_{n \in I} x_n \alpha_n.$$

## Définition

*Si  $X$  est une variable aléatoire discrète, le moment d'ordre  $k$  de  $X$  centré en  $a$  est le nombre réel (s'il existe)*

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

*Si  $a = 0$  nous l'appelons le moment d'ordre  $k$  de  $X$ .*

*Si  $a = \mathbb{E}[X]$  nous l'appelons le moment centré d'ordre  $k$  de  $X$ .*

## Définition

*Si  $X$  est une variable aléatoire discrète, nous appelons variance de  $X$  et nous la notons  $\text{Var}[X]$ , le moment centré d'ordre 2 de  $X$ .*

## Théorème (Théorème de Huygens.)

*Nous avons*

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

## Théorème

*Si  $X$  et  $Y$  sont deux variables aléatoires discrètes indépendantes, la variance de  $X + Y$  est égale à la somme des variances de  $X$  et de  $Y$ .*

## Remarques

1. Ce théorème se généralise pour  $n$  variables aléatoires discrètes indépendantes. En effet, nous avons que la variance de  $X_1 + \dots + X_n$  est égale à la somme des variances de  $X_1, \dots, X_n$ .
2. La variance d'une somme de variables aléatoires n'est en général pas égale à la somme des variances des variables aléatoires, sauf cas particulier comme nous venons de le voir. En effet, nous avons le résultat fondamental suivant.

## Proposition

$$\text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y] + 2\text{Cov}(X, Y).$$

## Définition

*Nous appelons la covariance de  $X$  et de  $Y$ , la quantité suivante*

$$\text{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y].$$

## Remarques

1. Nous avons

$$\text{Cov}(X, X) = \text{Var}[X].$$

2. De plus, nous avons

$$(\text{Cov}(X, Y))^2 \leq \text{Var}[X]\text{Var}[Y].$$

Cette inégalité permet de définir le coefficient de corrélation  $\rho(X, Y)$  qui est toujours compris entre  $-1$  et  $+1$ .



## Définition

*Le coefficient de corrélation linéaire entre les variables aléatoires  $X$  et  $Y$  se définit par*

$$\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}[X]\text{Var}[Y]}}.$$

## Remarque

Pour deux variables indépendantes, le coefficient de corrélation linéaire est égal à 0. Attention, la réciproque est en général fautive. Un coefficient de corrélation linéaire nul n'entraîne pas que les variables sont indépendantes. Deux exceptions notables où la non-corrélation et l'indépendance sont équivalentes :

- 1 les couples  $(X, Y)$  gaussiens (voir la section de ce chapitre s'y rapportant)
- 2 les couples de variables de Bernoulli (voir la section suivante).

Variables aléatoires réelles

Variables aléatoires discrètes

**Principales lois de probabilité discrètes**

Variables aléatoires absolument continues

Loi normale de paramètres  $\mu$  et  $\sigma^2$

Vecteurs gaussiens, lois associées

Loi de Bernoulli de paramètre  $p$

Loi binomiale de paramètres  $N$  et  $p$

Loi de Poisson de paramètre  $\lambda$

# Sommaire

- 1 Variables aléatoires réelles
- 2 Variables aléatoires discrètes
- 3 Principales lois de probabilité discrètes**
- 4 Variables aléatoires absolument continues
- 5 Loi normale de paramètres  $\mu$  et  $\sigma^2$
- 6 Vecteurs gaussiens, lois associées

Avec  $p = 0,5$ , cette loi modélise le jeu de pile ou face.  
Plus généralement, elle intervient lorsque la variable aléatoire est l'indicatrice d'un événement.

**Exemple :** « la boule tirée de l'urne est blanche ».

## Définition

*La loi de Bernoulli de paramètre  $p$  est définie par*

$$B(1; p) = (1 - p)\varepsilon_0 + p\varepsilon_1.$$

## Proposition

*L'espérance et la variance d'une variable aléatoire suivant une loi de Bernoulli de paramètre  $p$  sont égales respectivement à*

$$\mathbb{E}[X] = p$$

$$\text{Var}[X] = p(1 - p).$$

Cette loi intervient dans la sommation de  $N$  variables aléatoires de Bernoulli indépendantes.

**Exemple** : le nombre de boules blanches tirées, *avec remise*, d'une urne.

**Exemple** : si  $p = 0,5$ , c'est la loi de la variable aléatoire donnant le nombre de faces lorsque nous lançons  $N$  fois une pièce équilibrée.

Soit  $(X_i)_{i=1, \dots, N}$  une suite finie de variables de Bernoulli de paramètre  $p$  indépendantes et

$$X = \sum_{i=1}^N X_i.$$

Pour que  $X$  soit égale à  $k$  il faut que  $k$  exactement des  $N$  variables  $X_i$  soient égales à 1. L'ensemble de toutes ces possibilités correspond donc à l'ensemble des sous-ensembles à  $k$  éléments de  $\{1; \dots; N\}$ . Soit  $A$  un tel sous ensemble.

Comme les  $X_i$  sont indépendantes, en posant  $q = 1 - p$ , nous avons

$$\mathbb{P} \left[ \left( \bigcap_{i \in A} \{X_i = 1\} \right) \cap \left( \bigcap_{i \in \bar{A}} \{X_i = 0\} \right) \right] = p^k q^{N-k}.$$

En sommant sur tous les sous-ensembles à  $k$  éléments de  $\{1; \dots; N\}$ , puisque nous venons de voir que toutes les possibilités correspondantes sont équiprobables, nous obtenons, pour tout entier  $k$  entre 0 et  $N$

$$B(N; p)(\{X = k\}) = \binom{N}{k} q^{N-k} p^k.$$



## Définition

La loi binomiale de paramètres  $N$  et  $p$  est définie par

$$B(N; p) = \sum_{k=0}^N \binom{N}{k} q^{N-k} p^k \varepsilon_k.$$

## Remarque

Pour  $N = 1$ , ce n'est rien d'autre que la loi de Bernoulli  $B(1; p)$ , ce qui justifie a posteriori la notation utilisée pour cette dernière.

## Proposition

*L'espérance et la variance d'une variable aléatoire suivant une loi binomiale de paramètres  $N, p$  sont égales respectivement à*

$$\mathbb{E}[X] = Np$$

$$\text{Var}[X] = Np(1 - p).$$

C'est la loi des événements rares. Soit  $(X_n)_{n \in \mathbb{N}}$  une suite de variables binomiales  $\mathfrak{B}(n; p_n)$  avec  $\lim_{n \rightarrow +\infty} np_n = \lambda$ . En posant  $\lambda_n = np_n$  nous obtenons

$$\begin{aligned} \mathbb{P}[\{X_n = k\}] &= \binom{n}{k} p_n^k (1 - p_n)^{n-k} \\ &= \frac{n!}{k!(n-k)!} \frac{\lambda_n^k}{n^k} \left(1 - \frac{\lambda_n}{n}\right)^{n-k} \\ &= \frac{\lambda_n^k}{k!} \frac{n(n-1)\dots(n-k+1)}{n^k} \frac{\left(1 - \frac{\lambda_n}{n}\right)^n}{\left(1 - \frac{\lambda_n}{n}\right)^k}. \end{aligned}$$

Nous en déduisons que

$$\lim_{n \rightarrow +\infty} \mathbb{P}[\{X_n = k\}] = \frac{\lambda^k}{k!} e^{-\lambda}.$$

La limite ainsi obtenue est la loi de Poisson de paramètre  $\lambda$ .  
Elle s'écrit donc

$$\mathfrak{P}_\lambda = \sum_{k=0}^{+\infty} \frac{\lambda^k}{k!} e^{-\lambda} \varepsilon_k.$$

Comme  $\sum_{k=0}^{+\infty} \frac{\lambda^k}{k!} e^{-\lambda} = 1$  c'est bien une loi de probabilité.

## Proposition

*L'espérance et la variance d'une variable aléatoire suivant une loi de Poisson de paramètre  $\lambda$  sont égales respectivement à*

$$\mathbb{E}[X] = \lambda$$

$$\text{Var}[X] = \lambda.$$

# Sommaire

- 1 Variables aléatoires réelles
- 2 Variables aléatoires discrètes
- 3 Principales lois de probabilité discrètes
- 4 Variables aléatoires absolument continues**
- 5 Loi normale de paramètres  $\mu$  et  $\sigma^2$
- 6 Vecteurs gaussiens, lois associées

## Définition

Une variable aléatoire  $X$  définie sur un espace probabilisé  $(\Omega, \mathfrak{A}, P)$  est absolument continue s'il existe une fonction  $f_X$  définie sur  $\mathbb{R}$  telle que

- 1  $f_X(t) \geq 0$  pour tout  $t \in \mathbb{R}$  ;
- 2 l'ensemble des points de discontinuités de  $f_X$  est fini et ces discontinuités sont de 1<sup>ère</sup> espèce (i.e. la limite à gauche et à droite en chaque point existe) ;
- 3 pour tout  $x$  réel la fonction de répartition  $F_X$  de  $X$  est donnée par

$$F_X(x) = \int_{-\infty}^x f_X(t) dt.$$

La fonction  $f_X$  est appelée densité de la loi de probabilité de  $X$ .

## Définition

*L'espérance mathématique d'une variable aléatoire absolument continue  $X$  de densité  $f_X$  est le nombre réel (s'il existe)*

$$\mathbb{E}[X] = \int_{-\infty}^{+\infty} tf_X(t) dt.$$

*La variable aléatoire  $X$  a une espérance si cette intégrale converge absolument.*



## Définition

*Le moment d'ordre  $k$  (s'il existe) est le nombre réel*

$$m_k(X) = \int_{-\infty}^{+\infty} t^k f_X(t) dt.$$

*Le moment d'ordre  $k$  centré en  $a$  est défini par*

$${}_a m_k(X) = m_k(X - a).$$

*Le moment centré d'ordre  $k$  est le moment d'ordre  $k$  centré en  $a = \mathbb{E}[X]$ .*

## Définition

*La variance de  $X$  est le moment centré d'ordre 2.*

$$\text{Var}[X] = m_2(X - \mathbb{E}[X]).$$

## Remarque

Le **théorème de Huygens** est encore vérifié pour les variables aléatoires absolument continues :

$$\text{Var}[X] = m_2(X) - \mathbb{E}^2[X].$$

## Définition

*Lorsqu'une variable aléatoire a une espérance égale à 0, alors elle est dite centrée. Lorsqu'elle a une variance égale à 1, alors elle est dite réduite.*

## Théorème

*Si  $X$  et  $Y$  sont deux variables aléatoires absolument continues indépendantes, la variance de  $X + Y$  est égale à la somme des variances de  $X$  et de  $Y$ .*

# Sommaire

- 1 Variables aléatoires réelles
- 2 Variables aléatoires discrètes
- 3 Principales lois de probabilité discrètes
- 4 Variables aléatoires absolument continues
- 5 Loi normale de paramètres  $\mu$  et  $\sigma^2$**
- 6 Vecteurs gaussiens, lois associées

## Définition

*Une variable aléatoire  $X$  suit une loi normale de paramètres  $\mu$  et  $\sigma^2$ , notée  $\mathcal{N}(\mu; \sigma^2)$ , si la densité est la fonction définie par*

$$f_{\mu, \sigma^2}(t) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(t-\mu)^2}{2\sigma^2}}.$$

## Remarques

1 Rappel :

$$\int_{-\infty}^{+\infty} e^{-t^2} dt = \sqrt{\pi}.$$

2  $f_{\mu, \sigma^2}(\mu + u) = f_{\mu, \sigma^2}(\mu - u)$ .

3 En posant  $u = \frac{t - \mu}{\sigma\sqrt{2}}$ ,  $du = \frac{dt}{\sigma\sqrt{2}}$  nous obtenons

$$\begin{aligned} \int_{-\infty}^{+\infty} f_{\mu, \sigma^2}(t) dt &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{+\infty} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt \\ &= \frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-u^2} du \\ &= 1. \end{aligned}$$

## Proposition

*La fonction de répartition de la loi normale de paramètres  $\mu$  et  $\sigma^2$  vérifie*

$$F_X(\mu - x) = 1 - F_X(\mu + x).$$

Cette propriété est mise à profit dans les tables de la loi normale centrée et réduite où elle permet de ne mentionner que les valeurs de  $F_X$  correspondant aux valeurs positives de  $x$ .

## Proposition

*L'espérance et la variance d'une variable aléatoire suivant une loi normale de paramètres  $\mu$  et  $\sigma^2$  sont égales respectivement à*

$$\begin{aligned}\mathbb{E}[X] &= \mu \\ \text{Var}[X] &= \sigma^2.\end{aligned}$$



## Proposition

*Tous les moments d'ordre impair de  $X$ , où  $X$  suit une loi normale centrée et réduite, sont nuls. Tous les moments d'ordre pair de  $X$ , où  $X$  suit une loi normale centrée et réduite, sont égaux à*

$$m_{2k}(X) = \frac{(2k)!}{2^k k!}.$$

## Remarques

1. Introduisons les coefficients d'asymétrie  $\gamma_1$  et d'aplatissement  $\gamma_2$  (en anglais skewness et kurtosis) que nous détaillerons d'avantage par la suite

$$\gamma_1 = \frac{\mathbb{E}[X]m_3(X)}{\sigma^3} \quad \text{et} \quad \gamma_2 = \frac{\mathbb{E}[X]m_4(X)}{\sigma^4}.$$

2. Nous avons, pour une loi normale centrée et réduite

$$\gamma_1 = 0 \quad \text{et} \quad \gamma_2 = 3.$$

## Proposition

*Si  $X_1$  et  $X_2$  sont des variables indépendantes suivant respectivement des lois normales de paramètres respectivement  $(\mu_1, \sigma_1^2)$  et  $(\mu_2, \sigma_2^2)$ , alors  $X_1 + X_2$  est une variable aléatoire normale de paramètres  $\mu_1 + \mu_2$  et  $\sigma_1^2 + \sigma_2^2$ .*

## Remarque

Nous ne pouvons pas dire que toute combinaison linéaire de  $p$  variables gaussiennes non indépendantes soit encore gaussienne. Il faut pour cela que le  $p$ -uplet de variables suive une loi normale à  $p$ -dimensions (voir la section suivante).

# Sommaire

- 1 Variables aléatoires réelles
- 2 Variables aléatoires discrètes
- 3 Principales lois de probabilité discrètes
- 4 Variables aléatoires absolument continues
- 5 Loi normale de paramètres  $\mu$  et  $\sigma^2$
- 6 Vecteurs gaussiens, lois associées**

## Définition

*Un vecteur aléatoire  $\mathbf{X}$  est une application de  $(\Omega, \mathfrak{A}, \mathbb{P})$  dans un espace vectoriel réel, en général  $\mathbb{R}^p$  muni de sa tribu borélienne.*

En pratique,  $\mathbb{R}^p$  est muni de sa base canonique et on identifiera  $\mathbf{X}$  au  $p$ -uplet de variables aléatoires formé par ses composantes sur cette base  $\mathbf{X} = (X_1, X_2, \dots, X_p)$ .

## Définition

La fonction de répartition  $F_{\mathbf{X}}$  est une application de  $\mathbb{R}^p$  dans  $\mathbb{R}^p$  définie par

$$F_{\mathbf{X}}(x_1, \dots, x_p) = \mathbb{P}[X_1 < x_1, \dots, X_p < x_p].$$

## Définition

La densité  $f$  si elle existe est définie par

$$f_{\mathbf{X}}(x_1, \dots, x_p) = \frac{\partial^p F_{\mathbf{X}}}{\partial x_1 \dots \partial x_p}.$$

Le résultat suivant fondamental permet de définir des lois de probabilité à  $p$ -dimensions à partir des lois unidimensionnelles.

### Théorème (Théorème de Cramer-Wold.)

*La loi du vecteur aléatoire  $\mathbf{X}$  est entièrement déterminée par celles de toutes les combinaisons linéaires de ses composantes.*

## Définition

Nous appelons espérance de  $\mathbf{X} = (X_1, X_2, \dots, X_p)$  le vecteur

$$\mathbb{E}[\mathbf{X}] = \boldsymbol{\mu} = \begin{bmatrix} \mathbb{E}[X_1] \\ \mathbb{E}[X_2] \\ \cdot \\ \vdots \\ \cdot \\ \mathbb{E}[X_p] \end{bmatrix} .$$



## Définition

La matrice de variance-covariance, notée  $\Sigma$ , de  $\mathbf{X}$  est définie par

$$\Sigma = \begin{bmatrix} \text{Var}[X_1] & \text{Cov}[X_1, X_2] & \dots & \text{Cov}[X_1, X_p] \\ \cdot & \text{Var}[X_2] & \dots & \cdot \\ \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \dots & \text{Var}[X_p] \end{bmatrix} = \mathbb{E}[\mathbf{X}'\mathbf{X}] - \boldsymbol{\mu}\boldsymbol{\mu}'.$$

## Remarques

1.  $\Sigma$  est une matrice carrée symétrique réelle d'ordre  $p$ .
2. Si les variables  $X_i$  sont réduites,  $\Sigma$  s'identifie avec la matrice des corrélations linéaires

$$\begin{bmatrix} 1 & \rho(X_1, X_2) & \dots & \rho(X_1, X_p) \\ \cdot & 1 & \dots & \cdot \\ \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & 1 & \cdot \\ \cdot & \cdot & \dots & 1 \end{bmatrix}.$$

## Définition

*$\mathbf{X}$  est un vecteur gaussien à  $p$  dimensions si toute combinaison linéaire de ses composantes suit une loi de Laplace-Gauss à une dimension.*

Le **théorème de Cramer-Wold** permet d'établir que la loi de  $\mathbf{X}$  est ainsi parfaitement déterminée. Nous remarquerons que la normalité de chaque composante ne suffit nullement à définir un vecteur gaussien.

## Théorème

*Les composantes d'un vecteur gaussien  $\mathbf{X}$  sont indépendantes si et seulement si  $\Sigma$ , qui représente la matrice de variance-covariance de  $\mathbf{X}$ , est diagonale, c'est-à-dire si elles sont non corrélées.*

## Notation

Nous noterons  $\mathcal{N}(\boldsymbol{\mu}; \boldsymbol{\Sigma})$  la loi normale à  $p$  dimensions d'espérance  $\boldsymbol{\mu}$  et de matrice de variance-covariance  $\boldsymbol{\Sigma}$ .

## Théorème

Si  $\boldsymbol{\Sigma}$  est régulière, alors  $\mathbf{X}$  admet pour densité

$$f_{\mathbf{X}}(x_1, x_2, \dots, x_p) = \frac{1}{(2\pi)^{p/2} (\det \boldsymbol{\Sigma})^{p/2}} \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})\right).$$

## Définition

*Soit  $p$  un entier positif. Une variable aléatoire  $X$  suit une loi de Pearson ou loi du khi-deux à  $p$  degrés de liberté si sa densité est définie pour chaque  $t > 0$  par*

$$f_X(t) = \frac{1}{2^{\frac{p}{2}} \Gamma\left(\frac{p}{2}\right)} \exp\left(-\frac{t}{2}\right) t^{\frac{p}{2}-1}.$$

## Remarque

La somme de deux variables aléatoires de loi du  $\chi^2$  indépendantes à  $p$  et  $q$  degrés de liberté respectivement est encore une variable aléatoire du  $\chi^2$  à  $p + q$  degrés de liberté.

## Proposition

*L'espérance et la variance d'une variable aléatoire suivant une loi du  $\chi^2$  à  $p$  degrés de liberté sont égales respectivement à*

$$\begin{aligned}\mathbb{E}[X] &= p \\ \text{Var}[X] &= 2p.\end{aligned}$$

## Théorème

*Si  $U_1, U_2, \dots, U_p$  sont  $p$  variables gaussiennes centrées-réduites indépendantes, alors la variable aléatoire*

*$\sum_{i=1}^p U_i^2$  suit une loi de Pearson ou loi du khi-deux à  $p$  degrés de liberté.*

## Remarque

Dans certains ouvrages, ce théorème est une définition. En effet, au lieu d'introduire la fonction de densité pour caractériser la loi du khi-deux, certains auteurs donnent la définition suivante :

## Définition

*Soient  $U_1, U_2, \dots, U_p$ ,  $p$  variables gaussiennes centrées-réduites indépendantes. La variable aléatoire*

*$X = \sum_{i=1}^p U_i^2$  suit alors une loi du khi-deux à  $p$  degrés de liberté.*



## Remarque

Suite à cette définition, la loi du khi-deux est donc la loi de la somme des carrés des composantes d'un vecteur gaussien centré et de matrice de variance-covariance  $\mathbf{I}$ .

## Définition

Soient  $n$  et  $p$  deux entiers positifs. La variable aléatoire  $X$  suit une loi de Fisher-Snedecor à  $n$  et  $p$  degrés de liberté si sa densité est définie pour chaque  $t > 0$  par

$$f_X(t) = \frac{1}{B\left(\frac{n}{2}, \frac{p}{2}\right)} \frac{\left(\frac{n}{p}\right)^{\frac{n}{2}} t^{\frac{n}{2}-1}}{\left(1 + \frac{n}{p}t\right)^{\frac{n+p}{2}}}.$$

## Proposition

*L'espérance et la variance d'une variable aléatoire suivant une loi du F de Fisher-Snedecor à  $n$  et  $p$  degrés de liberté sont égales respectivement à*

$$\mathbb{E}[X] = \frac{p}{p-2} \quad \text{si } p > 2$$

$$\text{Var}[X] = 2 \frac{p^2}{n} \frac{n+p-2}{(p-2)^2(p-4)} \quad \text{si } p > 4.$$

## Théorème

*Soient  $X$  et  $Y$  étant deux variables aléatoires suivant indépendamment des lois du  $\chi^2$  à  $n$  et  $p$  degrés de liberté respectivement. Alors la variable aléatoire*

$$\frac{X/n}{Y/p}$$

*suit une loi de Fisher-Snedecor à  $n$  et  $p$  degrés de liberté.*

## Définition

*Une variable aléatoire  $X$  suit une loi de Student à  $n$  degrés de liberté si sa densité est égale à :*

$$f_X(t) = \frac{1}{\sqrt{n}B\left(\frac{1}{2}, \frac{n}{2}\right) \left(1 + \frac{t^2}{n}\right)^{\frac{n+1}{2}}}.$$

## Proposition

*L'espérance et la variance d'une variable aléatoire suivant une loi de Student à  $n$  degrés de liberté sont égales respectivement à*

$$\begin{aligned}\mathbb{E}[X] &= 0 \quad \text{si } n > 1 \\ \text{Var}[X] &= \frac{n}{n-2} \quad \text{si } n > 2.\end{aligned}$$

## Remarques

1. Si  $n \rightarrow \infty$ , alors  $T_n \rightarrow_{\mathcal{L}} \mathcal{N}(0; 1)$ . Ici la seconde flèche est une flèche qui signifie « converge en loi » qui est une convergence qui va être définie par la suite.
2. Nous avons la relation fondamentale entre les variables de Student et de Fisher-Snedecor

$$(T_n)^2 = F(1, n).$$

## Remarque

Une autre définition de la loi de Student peut être la suivante :

## Proposition

*Soient une variable aléatoire  $U$  suivant une loi normale centrée et réduite et  $X$  une variable suivant indépendamment de  $U$  une loi  $\chi^2(n)$ . Alors la variable aléatoire*

$$T_n = \frac{U}{\sqrt{X/n}}.$$

*suit la loi de Student à  $n$  degrés de liberté.*



## Remarque

De plus, si  $(X_1, \dots, X_n)$  est une suite de  $n$  variables aléatoires indépendantes et identiquement distribuées d'une loi normale d'espérance  $\mu$  et de variance  $\sigma^2$ , alors la variable aléatoire

$$\frac{\hat{\mu}_n - \mu}{S_n / \sqrt{n-1}}$$

suit une loi de Student à  $n-1$  degrés de liberté, où

$$\hat{\mu}_n = \frac{1}{n} \sum_{i=1}^n X_i \text{ et } S_n^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \hat{\mu}_n)^2.$$