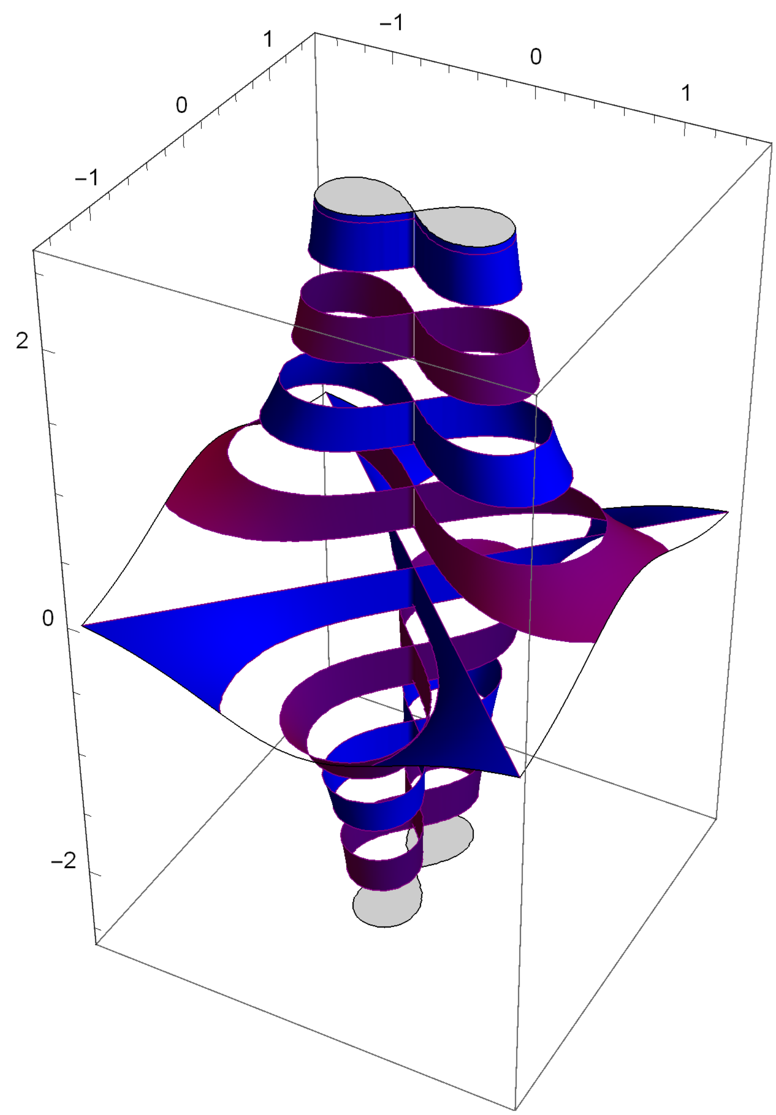
Université de Caen

# Mémo de Probabilités et Statistique

Christophe Chesneau

https://chesneau.users.lmno.cnrs.fr/




Caen, le 28 Septembre 2017

# Table des matières

# 1 Dénombrement

**Vocabulaires :**

| Notations | Vocabulaire | Notations | Vocabulaire |
|---|---|---|---|
| $\varnothing$ | ensemble vide | $A \cup B$ | réunion de $A$ et $B$ |
| $\Omega$ | ensemble plein | $A \cap B$ | intersection de $A$ et $B$ |
| $\{\omega\}$ | singleton de $\Omega$ | $A - B$ | intersection de $A$ et $\overline{B}$ |
| $A$ | partie de $\Omega$ | $A \cap B = \varnothing$ | $A$ et $B$ sont disjoints |
| $\omega \in A$ | $\omega$ appartient à $A$ | $A \subseteq B$ | $A$ est inclus dans $B$ |
| $\overline{A}$ | complémentaire de $A$ dans $\Omega$ | $A \times B$ | produit cartésien de $A$ et $B$ |

**Exemple :**

| Ensemble | Définition | $A = \{a, b\}$, $B = \{b, c\}$, $\Omega = \{a, b, c\}$ |
|---|---|---|
| $\overline{A}$ | $\{x \in \Omega;\ x \notin A\}$ | $\{c\}$ |
| $A \cup B$ | $\{x \in \Omega;\ x \in A$ ou $x \in B\}$ | $\{a, b, c\}$ |
| $A \cap B$ | $\{x \in \Omega;\ x \in A$ et $x \in B\}$ | $\{b\}$ |
| $A - B$ | $\{x \in \Omega;\ x \in A$ et $x \notin B\}$ | $\{a\}$ |
| $A \times B$ | $\{(x, y);\ x \in A$ et $y \in B\}$ | $\{(a, b), (a, c), (b, b), (b, c)\}$ |

**Opérations :** $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$, $(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$, $\left(\bigcup_{k=1}^{n} A_k\right) \cap B = \bigcup_{k=1}^{n} (A_k \cap B)$, $\left(\bigcap_{k=1}^{n} A_k\right) \cup B = \bigcap_{k=1}^{n} (A_k \cup B)$.

**Lois de Morgan :** $\overline{A \cup B} = \overline{A} \cap \overline{B}$, $\overline{A \cap B} = \overline{A} \cup \overline{B}$, $\overline{\bigcup_{k=1}^{n} A_k} = \bigcap_{k=1}^{n} \overline{A_k}$, $\overline{\bigcap_{k=1}^{n} A_k} = \bigcup_{k=1}^{n} \overline{A_k}$.

**Partition :** $(A_k)_{k \in \{1,\ldots,n\}}$ partition de $\Omega$ $\Leftrightarrow$ $(A_k)_{k \in \{1,\ldots,n\}}$ disjoints deux à deux et $\bigcup_{k=1}^{n} A_k = \Omega$.

**Cardinal :** Le nombre des éléments d'un ensemble fini $A$ est appelé cardinal de $A$. Il est noté $\mathrm{Card}(A)$.

**Formule du crible (ordre 2) :** $\mathrm{Card}(A \cup B) = \mathrm{Card}(A) + \mathrm{Card}(B) - \mathrm{Card}(A \cap B)$.

**Formule du crible (ordre $n$) :** $\mathrm{Card}\left(\bigcup_{k=1}^{n} A_k\right) = \sum_{k=1}^{n} (-1)^{k-1} \sum \ldots \sum_{(i_1,\ldots,i_k) \in U_k} \mathrm{Card}\left(\bigcap_{u=1}^{k} A_{i_u}\right)$,

où $U_k = \{(i_1, \ldots, i_k) \in \{1, \ldots, n\}^k;\ \ i_1 < \cdots < i_k\}$.

**Propriétés :** $\mathrm{Card}(\varnothing) = 0$, $A \cap B = \varnothing \Rightarrow \mathrm{Card}(A \cup B) = \mathrm{Card}(A) + \mathrm{Card}(B)$,

$\mathrm{Card}(\overline{A}) = \mathrm{Card}(\complement_\Omega A) = \mathrm{Card}(\Omega) - \mathrm{Card}(A)$, $\mathrm{Card}(A) = \mathrm{Card}(A \cap B) + \mathrm{Card}(A \cap \overline{B})$,

$\mathrm{Card}(A - B) = \mathrm{Card}(A) - \mathrm{Card}(A \cap B)$, $A \subseteq B \Rightarrow \mathrm{Card}(A) \leq \mathrm{Card}(B)$, $\mathrm{Card}(A \times B) = \mathrm{Card}(A)\,\mathrm{Card}(B)$.

**Principe additif :** On considère une situation qui nous amène à faire un choix parmi $n$ cas différents et exclusifs : le cas 1, ou le cas 2..., ou le cas $n$. Si, pour tout $k \in \{1, \ldots, n\}$, il y a $u_k$ possibilités pour le $k$-ème cas, alors le nombre total de possibilités est $\sum_{k=1}^{n} u_k$.

**Principe multiplicatif :** On considère une situation conjuguant $k$ étapes : une étape 1, et une étape 2..., et une étape $k$. Si, pour tout $i \in \{1, \ldots, k\}$, il y a $n_i$ possibilités pour la $k$-ème étape, alors le nombre total de possibilités est $\prod_{i=1}^{k} n_i$.

**Liste :** Liste ordonnée d'éléments avec répétitions.

**Arrangement :** Liste ordonnée d'éléments sans répétition.

**Permutation :** Arrangement de $n$ éléments parmi $n$.

**Combinaison :** Partie d'un ensemble ; l'ordre n'est pas pris en compte.

**Exemple :**

| Choix | avec répétition | sans répétition |
|---|---|---|
| avec ordre | Listes : $\left.\begin{matrix}(a,a) & (a,b) & (a,c)\\ (b,a) & (b,b) & (b,c)\\ (c,a) & (c,b) & (c,c)\end{matrix}\right\} 9$ | Arrangements : $\left.\begin{matrix}(a,b) & (a,c)\\ (b,a) & (b,c)\\ (c,a) & (c,b)\end{matrix}\right\} 6$ |
| sans ordre | Combinaisons avec répétitions : $\left.\begin{matrix}[a,a] & [a,b] & [a,c]\\ [b,b] & [b,c] & [c,c]\end{matrix}\right\} 6$ | Combinaisons : $\left.\begin{matrix}\{a,b\} & \{a,c\} & \{b,c\}\end{matrix}\right\} 3$ |

**Nombre de listes :** Le nombre de listes possibles de $k$ éléments parmi $n$ est $n^k$.

**Factorielle :** $n! = 1 \times 2 \times \ldots \times (n-1) \times n = \prod_{i=1}^{n} k$. On pose $0! = 1$.

**Nombre de permutations :** Le nombre de permutations de $n$ éléments est $n!$.

**Nombre d'arrangements :** On appelle nombre d'arrangements "$k$ parmi $n$" l'entier : $A_n^k = \dfrac{n!}{(n-k)!}$.

C'est le nombre d'arrangements possibles de $k$ éléments parmi $n$.

**Coefficient binomial :** On appelle coefficient binomial "$k$ parmi $n$" l'entier : $\dbinom{n}{k} = \dfrac{n!}{k!(n-k)!}$.

Si $k \notin \{0, \ldots, n\}$, on pose $\dbinom{n}{k} = 0$.

**Nombre de combinaisons :** Le nombre de combinaisons possibles de $k$ éléments parmi $n$ est $\dbinom{n}{k}$.

## 2 Espaces probabilisés et probabilités

**Expérience aléatoire :** On appelle expérience aléatoire toute expérience dont on connaît parfaitement les conditions mais dont on ne peut pas prévoir l'issue. Elle est notée $\mathcal{E}$.

**Univers :** On appelle univers d'une expérience aléatoire l'ensemble des issues possibles. On le note $\Omega$.

**Événement :** On appelle événement toute partie de l'univers $\Omega$. Les événements sont notés par des lettres majuscules : $A$, $B$, $C$... On dit qu'un événement $A$ se réalise (ou est réalisé) lors d'une expérience aléatoire si et seulement si l'issue de cette expérience aléatoire appartient à $A$.

**Vocabulaires :**

| Notations | Lire | Vocabulaires |
|---|---|---|
| $\varnothing$ | . | événement impossible |
| $\Omega$ | grand omega | univers (ou événement certain) |
| $\omega$ | omega | issue |
| $\{\omega\}$ | singleton omega | événement élémentaire |
| $A$ | . | événement |
| $\omega \in A$ | omega dans $A$ | $\omega$ est une réalisation possible de $A$ |

| | | |
|---|---|---|
| $\overline{A}$ | $A$ barre | événement contraire de $A$ |
| $A\cup B$ | $A$ union $B$ | réalisation de $A$, ou $B$, ou les deux |
| $A\cap B$ | $A$ inter $B$ | réalisation de $A$ et $B$ |
| $A-B$ | $A$ privé de $B$ | réalisation de $A$ et $\overline{B}$ |
| $A\cap B=\varnothing$ | . | $A$ et $B$ sont incompatibles |
| $A\subset B$ | $A$ inclus dans $B$ | $A$ implique $B$ |

On a : $\overline{A\cup B}=\overline{A}\cap\overline{B}=$ "ni $A$, ni $B$ se réalise" et $\overline{A\cap B}=\overline{A}\cup\overline{B}$.

**Système complet d'événements :** $(A_k)_{k\in\{1,\ldots,n\}}$ forme un système complet d'événements $\Leftrightarrow$ $(A_k)_{k\in\{1,\ldots,n\}}$ sont incompatibles deux à deux et $\bigcup\limits_{k=1}^{n} A_k=\Omega$.

**Tribu :** On dit que $\mathcal{A}$ est une tribu de l'univers $\Omega$ si, et seulement si, $\Omega\in\mathcal{A}$, pour tout $A\in\mathcal{A}$, $\overline{A}\in\mathcal{A}$ et, pour toute famille d'événements $(A_k)_{k\in\mathbb{N}^*}$ tels que, pour tout $k\in\mathbb{N}^*$, $A_k\in\mathcal{A}$, on a $\bigcup\limits_{k=1}^{\infty} A_k\in\mathcal{A}$.

Si $\Omega$ est fini ou infini dénombrable, on prend $\mathcal{A}=\mathcal{P}(\Omega)$ (l'ensemble des parties de $\Omega$).

**Probabilité $\mathbb{P}$ :** On appelle probabilité définie sur $(\Omega,\mathcal{A})$ toute application $\mathbb{P}:\mathcal{A}\to\mathbb{R}$ telle que :

- pour tout $A\in\mathcal{A}$, $\mathbb{P}(A)\in[0,1]$,
- $\mathbb{P}(\Omega)=1$,
- pour toute suite d'événements $(A_k)_{k\in\mathbb{N}^*}$ incompatibles deux à deux, $\mathbb{P}\left(\bigcup\limits_{k=1}^{\infty} A_k\right)=\sum\limits_{k=1}^{\infty}\mathbb{P}(A_k)$.

Si $\Omega$ est fini, on peut remplacer le dernier point par : $A\cap B=\varnothing\Rightarrow\mathbb{P}(A\cup B)=\mathbb{P}(A)+\mathbb{P}(B)$. Le réel $\mathbb{P}(A)$, prononcer "P de A", est la probabilité que l'événement $A$ se réalise.

**Probabilité sur un univers dénombrable :** Soient $\Omega$ un univers fini ou infini dénombrable (comme, par exemple, $\Omega=\mathbb{N}$), et $\mathcal{A}=\mathcal{P}(\Omega)$ l'ensemble des parties de $\Omega$. Si l'application $\mathbb{Q}:\mathcal{A}\to\mathbb{R}$ vérifie :

- pour tout $\omega\in\Omega$, $\mathbb{Q}(\{\omega\})\geq 0$,
- $\sum\limits_{\omega\in\Omega}\mathbb{Q}(\{\omega\})=1$,
- pour tout $A\in\mathcal{A}$, $\mathbb{Q}(A)=\sum\limits_{\omega\in A}\mathbb{Q}(\{\omega\})$,

alors $\mathbb{Q}$ est une probabilité définie sur $(\Omega,\mathcal{A})$.

**Espace probabilisé :** Le triplet $(\Omega,\mathcal{A},\mathbb{P})$ est appelé espace probabilisé.

**Propriétés :** $\mathbb{P}(\varnothing)=0$, $A\cap B=\varnothing\Rightarrow\mathbb{P}(A\cup B)=\mathbb{P}(A)+\mathbb{P}(B)$, $\mathbb{P}(\overline{A})=1-\mathbb{P}(A)$,

$\mathbb{P}(A)=\mathbb{P}(A\cap B)+\mathbb{P}(A\cap\overline{B})$, $A\subseteq B\Rightarrow\mathbb{P}(A)\leq\mathbb{P}(B)$.

**Formule d'inclusion-exclusion :** $\mathbb{P}(A\cup B)=\mathbb{P}(A)+\mathbb{P}(B)-\mathbb{P}(A\cap B)$.

**Formule d'inclusion-exclusion (à 3 termes) :**

$\mathbb{P}(A\cup B\cup C)=\mathbb{P}(A)+\mathbb{P}(B)+\mathbb{P}(C)-\mathbb{P}(A\cap B)-\mathbb{P}(A\cap C)-\mathbb{P}(B\cap C)+\mathbb{P}(A\cap B\cap C)$.

**Formule d'inclusion-exclusion (à $n$ termes) :** $\mathbb{P}\left(\bigcup\limits_{k=1}^{n} A_k\right)=\sum\limits_{k=1}^{n}(-1)^{k-1}\sum\limits_{(i_1,\ldots,i_k)\in U_k}\!\!\ldots\sum\mathbb{P}\left(\bigcap\limits_{u=1}^{k} A_{i_u}\right)$,

où $U_k=\{(i_1,\ldots,i_k)\in\{1,\ldots,n\}^k;\ \ i_1<\ldots<i_k\}$.

**Cas particulier :** $(A_k)_{k\in\{1,\ldots,n\}}$ incompatibles deux à deux $\Rightarrow$ $\mathbb{P}\left(\bigcup\limits_{k=1}^{n} A_k\right)=\sum\limits_{k=1}^{n}\mathbb{P}(A_k)$.

**Inégalité :** On a toujours $\mathbb{P}\left(\bigcup\limits_{k=1}^{n} A_k\right)\leq\sum\limits_{k=1}^{n}\mathbb{P}(A_k)$.

**Propriétés de limite monotone :** $A_{n+1} \subset A_n \Rightarrow \lim_{n\to\infty} \mathbb{P}(A_n) = \mathbb{P}\left(\bigcap_{k=1}^{\infty} A_k\right)$,

$A_n \subset A_{n+1} \Rightarrow \lim_{n\to\infty} \mathbb{P}(A_n) = \mathbb{P}\left(\bigcup_{k=1}^{\infty} A_k\right)$.

**Équiprobabilité :** Si aucun événement élémentaire n'a plus de chance de se réaliser que les autres, on dit qu'il y a équiprobabilité.

**Probabilité uniforme :** Si $\Omega$ est fini et qu'il y a équiprobabilité, alors on peut considérer l'espace probabilisé $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$, où $\mathbb{P}$ désigne la probabilité uniforme définie par $\mathbb{P}(A) = \frac{\mathrm{Card}(A)}{\mathrm{Card}(\Omega)}$, $A \in \mathcal{A}$.

# 3 Probabilités conditionnelles et indépendance

**Probabilité (conditionnelle) de $A$ sachant $B$ :** L'application $\mathbb{P}_B : \mathcal{A} \to [0,1]$ définie par $\mathbb{P}_B(A) = \frac{\mathbb{P}(A\cap B)}{\mathbb{P}(B)}$, $A \in \mathcal{A}$, est une probabilité. Le réel $\mathbb{P}_B(A)$, prononcer "P de A sachant B", est la probabilité que l'événement $A$ se réalise sachant que l'événement $B$ est réalisé (avec $\mathbb{P}(B) \neq 0$).

**Formule des probabilités composées (ordre 3) :** $\mathbb{P}(A\cap B\cap C) = \mathbb{P}(A)\mathbb{P}_A(B)\mathbb{P}_{A\cap B}(C)$.

**Formule des probabilités totales :** $\mathbb{P}(A) = \mathbb{P}_B(A)\mathbb{P}(B) + \mathbb{P}_{\overline{B}}(A)\mathbb{P}(\overline{B})$.

**Formule de Bayes :** $\mathbb{P}_B(A) = \frac{\mathbb{P}_A(B)\mathbb{P}(A)}{\mathbb{P}_A(B)\mathbb{P}(A) + \mathbb{P}_{\overline{A}}(B)\mathbb{P}(\overline{A})}$.

**Formule des probabilités composées (ordre $n$) :** $\mathbb{P}\left(\bigcap_{k=1}^{n} A_k\right) = \mathbb{P}(A_1)\prod_{i=2}^{n} \mathbb{P}_{\bigcap_{k=1}^{i-1} A_k}(A_i)$.

**Formule des probabilités totales (ordre $n$) :** $(A_k)_{k\in\{1,\ldots,n\}}$ système complet d'événements $\Rightarrow$ $\mathbb{P}(B) = \sum_{k=1}^{n} \mathbb{P}_{A_k}(B)\mathbb{P}(A_k)$.

**Formule de Bayes (ordre $n$) :** $(A_k)_{k\in\{1,\ldots,n\}}$ système complet d'événements $\Rightarrow$ $\mathbb{P}_B(A_i) = \frac{\mathbb{P}_{A_i}(B)\mathbb{P}(A_i)}{\sum_{k=1}^{n} \mathbb{P}_{A_k}(B)\mathbb{P}(A_k)}$.

**Indépendance :**

- $A$ et $B$ indépendants $\Leftrightarrow$ $\mathbb{P}(A\cap B) = \mathbb{P}(A)\mathbb{P}(B)$ $\Leftrightarrow$ $\mathbb{P}_B(A) = \mathbb{P}(A)$ $\Leftrightarrow$ $\mathbb{P}_A(B) = \mathbb{P}(B)$.
- $A$ et $B$ indépendants $\Leftrightarrow$ $\overline{A}$ et $B$ indépendants $\Leftrightarrow$ $A$ et $\overline{B}$ indépendants $\Leftrightarrow$ $\overline{A}$ et $\overline{B}$ indépendants
- $(A_k)_{k\in\{1,\ldots,n\}}$ indépendants deux à deux $\Leftrightarrow$ pour tout $(k,l) \in \{1,\ldots,n\}^2$ avec $k \neq l$, $\mathbb{P}(A_k \cap A_l) = \mathbb{P}(A_k)\mathbb{P}(A_l)$.
- $(A_k)_{k\in\{1,\ldots,n\}}$ indépendants $\Leftrightarrow$ pour tout $I \subseteq \{1,\ldots,n\}$, $\mathbb{P}\left(\bigcap_{k\in I} A_k\right) = \prod_{k\in I} \mathbb{P}(A_k)$.

# 4 Variables aléatoires réelles (*var*) discrètes

**Variable aléatoire réelle (*var*) :** Une *var* $X$ est une application $X : \Omega \to \mathbb{R}$ telle que, pour tout $x \in \mathbb{R}$, $\{X \leq x\} = \{\omega \in \Omega;\ X(\omega) \leq x\} \in \mathcal{A}$.

**Support d'une *var* :** Le support d'une *var* $X$ est l'ensemble des valeurs atteintes par $X$. Il est noté $X(\Omega)$.

**Loi d'une *var* :** La loi d'une *var* $X$ est la probabilité $\mathbb{P}_X$ définie par $\mathbb{P}_X(A) = \mathbb{P}(X \in A) = \mathbb{P}(\{\omega \in \Omega;\ X(\omega) \in A\})$, $A \subseteq \mathbb{R}$.

***Var* discrète :**

- On appelle *var* discrète toute *var* prenant un ensemble de valeurs fini, ou infini dénombrable.
- La loi d'une *var* discrète $X$ est donnée par $\mathbb{P}(X=k)$, $k \in X(\Omega)$, où $\mathbb{P}(X=k)$ désigne la probabilité que l'événement $\{X=k\} = \{\omega \in \Omega;\ X(\omega)=k\}$ se réalise.

**Propriétés :** $\sum_{k\in X(\Omega)} \mathbb{P}(X=k) = 1$, $\mathbb{P}(X\in\mathcal{D}) = \sum_{k\in X(\Omega)\cap\mathcal{D}} \mathbb{P}(X=k)$.

**Exemples :**

| Loi | uniforme | de Bernoulli | binomiale | géométrique | de Poisson |
|---|---|---|---|---|---|
| paramètres | $n\in\mathbb{N}^*$ | $p\in]0,1[$ | $n\in\mathbb{N}^*$, $p\in]0,1[$ | $p\in]0,1[$ | $\lambda>0$ |
| $X\sim$ | $\mathcal{U}(\{1,\ldots,n\})$ | $\mathcal{B}(p)$ | $\mathcal{B}(n,p)$ | $\mathcal{G}(p)$ | $\mathcal{P}(\lambda)$ |
| $X(\Omega)$ $(k\in)$ | $\{1,\ldots,n\}$ | $\{0,1\}$ | $\{0,\ldots,n\}$ | $\mathbb{N}^*$ | $\mathbb{N}$ |
| $\mathbb{P}(X=k)$ | $\frac{1}{n}$ | $p$ si $k=1$ | $\binom{n}{k}p^k(1-p)^{n-k}$ | $p(1-p)^{k-1}$ | $e^{-\lambda}\frac{\lambda^k}{k!}$ |

**Caractérisation :** $q_k$, $k\in\mathcal{D}$, caractérise la loi d'une *var* discrète $\Leftrightarrow$ $q_k\geq 0$, et $\sum_{k\in\mathcal{D}} q_k = 1$.

**Loi et probabilité uniforme :** $\mathbb{P}$ probabilité uniforme $\Rightarrow$ $\mathbb{P}(X=k) = \frac{\mathrm{Card}(\{X=k\})}{\mathrm{Card}(\Omega)}$, $k\in X(\Omega)$.

**Égalité en loi :** $X$ et $Y$ suivent la même loi $\Leftrightarrow$ $X(\Omega)=Y(\Omega)$ et $\mathbb{P}(X=k)=\mathbb{P}(Y=k)$, $k\in X(\Omega)$.

**Fonction de répartition :** $F(x) = \mathbb{P}(X\leq x) = \sum_{k\in]-\infty,x]\cap X(\Omega)} \mathbb{P}(X=k)$, $x\in\mathbb{R}$.

**Fonction de répartition et loi :** Pour tout $k\in X(\Omega)$, on a $\mathbb{P}(X=k) = F(k)-F(k_*)$, où $k_*$ est la plus grande valeur de $X(\Omega)$ vérifiant $k_*<k$.

**Espérance :** $\mathbb{E}(X) = \sum_{k\in X(\Omega)} k\mathbb{P}(X=k)$.

**Formule du transfert :** $\mathbb{E}(g(X)) = \sum_{k\in X(\Omega)} g(k)\mathbb{P}(X=k)$.

**Propriétés :** $\mathbb{E}(a)=a$, $X(\Omega)\subseteq[0,\infty[$ $\Rightarrow$ $\mathbb{E}(X)\geq 0$, $\mathbb{E}(aX+b) = a\mathbb{E}(X)+b$, $\mathbb{E}(aX^2+bX+c) = a\mathbb{E}(X^2)+b\mathbb{E}(X)+c$.

**Moment d'ordre $p$ :** $m_p = \mathbb{E}(X^p) = \sum_{k\in X(\Omega)} k^p\mathbb{P}(X=k)$.

**Variance :** $\mathbb{V}(X) = \mathbb{E}((X-\mathbb{E}(X))^2)$.

**Propriétés :** $\mathbb{V}(X)\geq 0$, $\mathbb{V}(X)=0$ $\Leftrightarrow$ $X$ est une *var* constante, $\mathbb{V}(aX+b) = a^2\mathbb{V}(X)$.

**Formule de König-Huyghens :** $\mathbb{V}(X) = \mathbb{E}(X^2)-(\mathbb{E}(X))^2$.

**Exemples :**

| $X\sim$ | $\mathcal{U}(\{1,\ldots,n\})$ | $\mathcal{B}(p)$ | $\mathcal{B}(n,p)$ | $\mathcal{G}(p)$ | $\mathcal{P}(\lambda)$ |
|---|---|---|---|---|---|
| $\mathbb{E}(X)$ | $\frac{n+1}{2}$ | $p$ | $np$ | $\frac{1}{p}$ | $\lambda$ |
| $\mathbb{V}(X)$ | $\frac{n^2-1}{12}$ | $p(1-p)$ | $np(1-p)$ | $\frac{1-p}{p^2}$ | $\lambda$ |

**Écart-type :** $\sigma(X) = \sqrt{\mathbb{V}(X)}$.

**Inégalité de Markov :** Soit $h:\mathbb{R}\to]0,\infty[$ une fonction croissante. Pour tout $\epsilon>0$, $\mathbb{P}(|X|\geq\epsilon)\leq\frac{\mathbb{E}(h(|X|))}{h(\epsilon)}$.

**Inégalité de Bienaymé-Tchebychev :** Pour tout $\epsilon>0$, $\mathbb{P}(|X-\mathbb{E}(X)|\geq\epsilon)\leq\frac{\mathbb{V}(X)}{\epsilon^2}$.

**Fonction génératrice :** $G(s) = \mathbb{E}(s^X) = \sum_{k\in X(\Omega)} s^k\mathbb{P}(X=k)$, $s\in[-1,1]$.

**Fonction génératrice et ses dérivées :** $G'(s) = \mathbb{E}(Xs^{X-1})$, $G''(s) = \mathbb{E}(X(X-1)s^{X-2})$,

$$G^{(k)}(s) = \mathbb{E}\left(\prod_{i=0}^{k-1}(X-i)s^{X-k}\right).$$

**Propriétés :** $\mathbb{E}(X) = G'(1)$, $\mathbb{V}(X) = G''(1) + G'(1) - (G'(1))^2$.

**Fonction génératrice et égalité en loi :** $X$ et $Y$ suivent la même loi $\Leftrightarrow$ $G_X(s) = G_Y(s)$ pour tout $s \in [-1,1]$.

**Exemples :**

| $X \sim$ | $\mathcal{U}(\{1,\ldots,n\})$ | $\mathcal{B}(p)$ | $\mathcal{B}(n,p)$ | $\mathcal{G}(p)$ | $\mathcal{P}(\lambda)$ |
|---|---|---|---|---|---|
| $G(s)$ | $\frac{s}{n}\left(\frac{1-s^n}{1-s}\right)$ | $1+p(s-1)$ | $(1+p(s-1))^n$ | $\frac{ps}{1-(1-p)s}$ | $e^{\lambda(s-1)}$ |

# 5 Modélisation : loi binomiale et loi géométrique

**Schéma d'urne :** Une urne contient $N$ boules dont $N_1$ sont blanches et $N_2 = N - N_1$ sont noires. On effectue des tirages au hasard et avec remise.

- La *var* $X$ égale au nombre de boules blanches obtenues en $n$ tirages suit la loi binomiale $\mathcal{B}(n,p)$ avec $p = \frac{N_1}{N}$.
- La *var* $X$ égale au nombre de tirages nécessaires pour obtenir pour la première fois une boule blanche suit la loi géométrique $\mathcal{G}(p)$ avec $p = \frac{N_1}{N}$.

**Répétition d'expériences :** On répète, dans les mêmes conditions, une même expérience aléatoire au cours de laquelle un événement $A$ a une probabilité $p$ d'être réalisé.

- La *var* $X$ égale au nombre de réalisations de l'événement $A$ en $n$ expériences (indépendantes) suit la loi binomiale $\mathcal{B}(n,p)$.
- La *var* $X$ égale au nombre d'expériences (indépendantes) nécessaires pour obtenir une réalisation de $A$ suit la loi géométrique $\mathcal{G}(p)$.

# 6 Couples de *var* discrètes

**Couple de *var* :** Soient $X$ et $Y$ deux *var*. On appelle couple de *var* l'application $(X,Y) : \Omega \to \mathbb{R}^2$ vérifiant $(X,Y)(\omega) = (X(\omega), Y(\omega))$, $\omega \in \Omega$.

**Support d'un couple de *var* :** Le support d'un couple de *var* $(X,Y)$ est l'ensemble des valeurs atteintes par $(X,Y)$. Il est noté $(X,Y)(\Omega)$.

**Loi d'un couple de *var* :** La loi d'un couple de *var* $(X,Y)$ est la probabilité $\mathbb{P}_{(X,Y)}$ définie par

$$\mathbb{P}_{(X,Y)}(A) = \mathbb{P}((X,Y) \in A) = \mathbb{P}(\{\omega \in \Omega;\ (X,Y)(\omega) \in A\}),\ A \subseteq \mathbb{R}^2.$$

**Couple de *var* discrètes :**

- On appelle couple de *var* discrètes tout couple de *var* $(X,Y)$ prenant un ensemble de valeurs fini, ou infini dénombrable. Dès lors, $X$ est une *var* discrète et $Y$ aussi.
- La loi d'une couple de *var* discrètes $(X,Y)$ est donnée par $\mathbb{P}(\{X=i\} \cap \{Y=j\})$, $(i,j) \in (X,Y)(\Omega)$. On la présente parfois sous la forme d'un tableau quand cela est possible.

**Propriétés :** $\sum\sum_{(i,j)\in(X,Y)(\Omega)} \mathbb{P}(\{X=i\} \cap \{Y=j\}) = 1$, $\mathbb{P}((X,Y) \in \mathcal{D}) = \sum\sum_{(i,j)\in(X,Y)(\Omega)\cap\mathcal{D}} \mathbb{P}(\{X=i\} \cap \{Y=j\})$.

**Caractérisation :** $q_{i,j}$, $(i,j) \in \mathcal{D}$, caractérise la loi d'un couple de *var* discrètes $\Leftrightarrow$ $q_{i,j} \geq 0$ et $\sum\sum_{(i,j)\in\mathcal{D}} q_{i,j} = 1$.

**Loi et probabilité uniforme :** $\mathbb{P}$ probabilité uniforme $\Rightarrow$ $\mathbb{P}(\{X=i\}\cap\{Y=j\}) = \dfrac{\mathrm{Card}(\{X=i\}\cap\{Y=j\})}{\mathrm{Card}(\Omega)}$, $(i,j) \in (X,Y)(\Omega)$.

**Loi et probabilité conditionnelle :** $\mathbb{P}(\{X=i\}\cap\{Y=j\}) = \begin{cases} \mathbb{P}_{\{Y=j\}}(X=i)\mathbb{P}(Y=j), \\ \mathbb{P}_{\{X=i\}}(Y=j)\mathbb{P}(X=i). \end{cases}$

**Fonction de répartition :** $F(x,y) = \mathbb{P}(\{X \leq x\}\cap\{Y \leq y\})$, $(x,y) \in \mathbb{R}^2$.

**Lois marginales :**

- On a $X(\Omega) = \bigcup_{j\in\mathbb{R}} \{i \in \mathbb{R};\ \mathbb{P}(\{X=i\}\cap\{Y=j\}) \neq 0\}$,
  la loi de $X$ est donnée par $\mathbb{P}(X=i) = \sum_{j\in Y(\Omega)} \mathbb{P}(\{X=i\}\cap\{Y=j\})$, $i \in X(\Omega)$.
- On a $Y(\Omega) = \bigcup_{i\in\mathbb{R}} \{j \in \mathbb{R};\ \mathbb{P}(\{X=i\}\cap\{Y=j\}) \neq 0\}$,
  la loi de $Y$ est donnée par $\mathbb{P}(Y=j) = \sum_{i\in X(\Omega)} \mathbb{P}(\{X=i\}\cap\{Y=j\})$, $j \in Y(\Omega)$.

**Indépendance de 2 *var* discrètes :** $X$ et $Y$ indépendantes $\Leftrightarrow$ $(X,Y)(\Omega) = X(\Omega)\times Y(\Omega)$ et, pour tout $(i,j) \in (X,Y)(\Omega)$, $\mathbb{P}(\{X=i\}\cap\{Y=j\}) = \mathbb{P}(X=i)\mathbb{P}(Y=j)$.

**Sur l'indépendance :**

- S'il existe $(i_0,j_0) \in (X,Y)(\Omega)$ tel que $\mathbb{P}(\{X=i_0\}\cap\{Y=j_0\}) \neq \mathbb{P}(X=i_0)\mathbb{P}(Y=j_0)$, alors $X$ et $Y$ ne sont pas indépendantes.
- S'il existe deux suites de réels positifs $(a_i)_{i\in X(\Omega)}$ et $(b_j)_{j\in Y(\Omega)}$ telles que $\mathbb{P}(\{X=i\}\cap\{Y=j\}) = a_i b_j$, $\sum_{i\in X(\Omega)} a_i = 1$, $\sum_{j\in Y(\Omega)} b_j = 1$, alors $X$ et $Y$ sont indépendantes, et la loi de $X$ est donnée par
  $\mathbb{P}(X=i) = a_i$, $i \in X(\Omega)$ et la loi de $Y$ est donnée par $\mathbb{P}(Y=j) = b_j$, $j \in Y(\Omega)$.
- $X$ et $Y$ indépendantes $\Rightarrow$ $\mathbb{P}(\{X \in \mathcal{D}_1\}\cap\{Y \in \mathcal{D}_2\}) = \mathbb{P}(X\in\mathcal{D}_1)\mathbb{P}(Y\in\mathcal{D}_2)$.
- $X$ et $Y$ indépendantes $\Rightarrow$ $g(X)$ et $h(Y)$ indépendantes.

**"Indépendantes et identiquement distribuées"/*iid* :** On dit que $X$ et $Y$ sont *iid* quand elles sont "indépendantes et identiquement distribuées/suivent la même loi".

**Loi d'une somme de deux *var* indépendantes :** *(produit de convolution)*

$\mathbb{P}(X+Y=k) = \sum_{i\in X(\Omega)} \mathbb{P}(Y=k-i)\mathbb{P}(X=i)$, $k \in (X+Y)(\Omega)$.

**Exemples :**

| $X_i \sim$ | $\mathcal{B}(p)$ | $\mathcal{B}(m_i,p)$ | $\mathcal{P}(\lambda_i)$ | $\mathcal{G}(p)$ |
|---|---|---|---|---|
| $X_1+X_2 \sim$ | $\mathcal{B}(2,p)$ | $\mathcal{B}(m_1+m_2,p)$ | $\mathcal{P}(\lambda_1+\lambda_2)$ | $\mathcal{G}(2,p)$ |

**Formule du transfert :** $\mathbb{E}(g(X,Y)) = \sum\sum_{(i,j)\in(X,Y)(\Omega)} g(i,j)\mathbb{P}(\{X=i\}\cap\{Y=j\})$.

**Linéarité de l'espérance :** $\mathbb{E}(aX+bY) = a\mathbb{E}(X) + b\mathbb{E}(Y)$.

**Inégalité de Cauchy-Schwarz :** $|\mathbb{E}(XY)| \leq \sqrt{\mathbb{E}(X^2)}\sqrt{\mathbb{E}(Y^2)}$.

**Moments et indépendance :** $X$ et $Y$ indépendantes $\Rightarrow$ $\mathbb{E}(g(X)h(Y)) = \mathbb{E}(g(X))\mathbb{E}(h(Y))$. En particulier, $\mathbb{E}(X^kY^m) = \mathbb{E}(X^k)\mathbb{E}(Y^m)$.

**Fonction génératrice :** Si $(X,Y)(\Omega) \subseteq \mathbb{N}^2$, $G(s,t) = \mathbb{E}(s^Xt^Y) = \sum\sum_{(i,j)\in(X,Y)(\Omega)} s^it^j\mathbb{P}(\{X=i\}\cap\{Y=j\})$, $(s,t) \in [-1,1]^2$.

**Propriétés :** $G(s,1) = G_X(s)$, $G(1,t) = G_Y(t)$, $G(s,s) = G_{X+Y}(s)$, $\mathbb{E}(XY) = \dfrac{\partial^2 G}{\partial s \partial t}(1,1)$.

**Fonction génératrice et indépendance :** $X$ et $Y$ indépendantes $\Leftrightarrow$ $G(s,t) = G_X(s)G_Y(t)$ pour tout $(s,t) \in [-1,1]^2$.

**Covariance :** $\mathbb{C}(X,Y) = \mathbb{E}((X - \mathbb{E}(X))(Y - \mathbb{E}(Y))) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$.

**Matrice de covariance :** $\Sigma = \begin{pmatrix} \mathbb{V}(X) & \mathbb{C}(X,Y) \\ \mathbb{C}(Y,X) & \mathbb{V}(Y) \end{pmatrix}$.

**Variance d'une somme de *var* :** $\mathbb{V}(X+Y) = \mathbb{V}(X) + \mathbb{V}(Y) + 2\mathbb{C}(X,Y)$.

**Propriétés :** $\mathbb{C}(X,Y) = \mathbb{C}(Y,X)$, $\mathbb{C}(X,a) = 0$, $\mathbb{C}(aX+b, cY+d) = ac\mathbb{C}(X,Y)$,
$\mathbb{C}(aX+bY, cU+dV) = ac\mathbb{C}(X,U) + ad\mathbb{C}(X,V) + bc\mathbb{C}(Y,U) + bd\mathbb{C}(Y,V)$.

**Paramètres d'un couple de *var* :** Les principaux paramètres d'un couple de *var* (vecteur colonne) $U = (X,Y)^t$ sont son espérance : $\mathbb{E}_2(U) = (\mathbb{E}(X), \mathbb{E}(Y))^t$, et sa matrice de covariance $\Sigma$.

**Covariance, variance et indépendance :** $X$ et $Y$ indépendantes $\Rightarrow$

- $\mathbb{C}(X,Y) = 0$,
- $\mathbb{V}(X+Y) = \mathbb{V}(X) + \mathbb{V}(Y)$.

Si $\mathbb{C}(X,Y) = 0$, on ne peut rien conclure sur l'indépendance des *var* $X$ et $Y$.

**Coefficient de corrélation linéaire :** $\rho(X,Y) = \dfrac{\mathbb{C}(X,Y)}{\sigma(X)\sigma(Y)}$. On a $\rho(X,Y) \in [-1,1]$. Plus $|\rho(X,Y)|$ est proche de 1, plus la dépendance linéaire entre $X$ et $Y$ est forte.

# 7 Vecteurs de *var* discrètes

**Vecteur de *var* :** Soient $X_1, \ldots, X_n$ $n$ *var*. On appelle vecteur de *var* l'application $(X_1, \ldots, X_n) : \Omega \to \mathbb{R}^n$ vérifiant $(X_1, \ldots, X_n)(\omega) = (X_1(\omega), \ldots, X_n(\omega))$, $\omega \in \Omega$.

**Support d'un vecteur de *var* :** Le support d'un vecteur de *var* $(X_1, \ldots, X_n)$ est l'ensemble des valeurs atteintes par $(X_1, \ldots, X_n)$. Il est noté $(X_1, \ldots, X_n)(\Omega)$.

**Loi d'un vecteur de *var* :** La loi d'un vecteur de *var* $(X_1, \ldots, X_n)$ est la probabilité $\mathbb{P}_{(X_1,\ldots,X_n)}$ définie par $\mathbb{P}_{(X_1,\ldots,X_n)}(A) = \mathbb{P}((X_1, \ldots, X_n) \in A) = \mathbb{P}(\{\omega \in \Omega;\ (X_1, \ldots, X_n)(\omega) \in A\})$, $A \subseteq \mathbb{R}^n$.

**Vecteur de *var* discrètes :**

- On appelle vecteur de *var* discrètes tout vecteur de *var* $(X_1, \ldots, X_n)$ prenant un ensemble de valeurs fini, ou infini dénombrable. Dès lors, pour tout $i \in \{1, \ldots, n\}$, $X_i$ est une *var* discrète.
- La loi de $(X_1, \ldots, X_n)$ est donnée par $\mathbb{P}\left(\bigcap_{i=1}^{n} \{X_i = u_i\}\right)$, $(u_1, \ldots, u_n) \in (X_1 \ldots X_n)(\Omega)$.

**Propriétés :** $\displaystyle \sum \ldots \sum_{(u_1,\ldots,u_n) \in (X_1,\ldots,X_n)(\Omega)} \mathbb{P}\left(\bigcap_{i=1}^{n} \{X_i = u_i\}\right) = 1$.

$\displaystyle \mathbb{P}((X_1, \ldots, X_n) \in \mathcal{D}) = \sum \ldots \sum_{(u_1,\ldots,u_n) \in (X_1,\ldots,X_n)(\Omega) \cap \mathcal{D}} \mathbb{P}\left(\bigcap_{i=1}^{n} \{X_i = u_i\}\right)$.

**Caractérisation :** $q_{u_1,\ldots,u_n}$, $(u_1, \ldots, u_n) \in \mathcal{D}$ caractérise la loi d'un vecteur de *var* discrètes $\Leftrightarrow$ $q_{u_1,\ldots,u_n} \geq 0$ et $\displaystyle \sum \ldots \sum_{(u_1,\ldots,u_n) \in \mathcal{D}} q_{u_1,\ldots,u_n} = 1$.

**Loi marginale d'une *var* composante de $(X_1,\ldots,X_n)$ :** La loi de $X_1$ est donnée par

$$\mathbb{P}(X_1=u_1)=\sum_{(u_2,\ldots,u_n)\in(X_2,\ldots,X_n)(\Omega)}\!\!\!\sum\ldots\sum \mathbb{P}\left(\bigcap_{i=1}^{n}\{X_i=u_i\}\right),\ u_1\in X_1(\Omega).$$

On définit de la même manière les lois marginales des autres *var* $X_2,\ldots,X_n$.

**Loi de vecteurs de *var* composantes de $(X_1,\ldots,X_n)$ :** La loi de $(X_1,X_2)$ est donnée par

$$\mathbb{P}(\{X_1=u_1\}\cap\{X_2=u_2\})=\sum_{(u_3,\ldots,u_n)\in(X_3,\ldots,X_n)(\Omega)}\!\!\!\sum\ldots\sum \mathbb{P}\left(\bigcap_{i=1}^{n}\{X_i=u_i\}\right),\ (u_1,u_2)\in(X_1,X_2)(\Omega).$$

On définit de la même manière les lois d'autres vecteurs de *var*.

**Fonction de répartition :** $F(x_1,\ldots,x_n)=\mathbb{P}\left(\bigcap_{i=1}^{n}\{X_i\leq x_i\}\right)$, $(x_1,\ldots,x_n)\in\mathbb{R}^n$.

**Fonction génératrice :** Si $(X_1,\ldots,X_n)(\Omega)\subseteq\mathbb{N}^n$, $G(s_1,\ldots,s_n)=\mathbb{E}\left(\prod_{i=1}^{n}s_i^{X_i}\right)$, $(s_1,\ldots,s_n)\in[-1,1]^n$.

**Indépendance de $n$ *var* discrètes :** $X_1,\ldots,X_n$ indépendantes $\Leftrightarrow$ pour tout $(u_1,\ldots,u_n)\in\prod_{i=1}^{n}X_i(\Omega)$ et tout $I\subseteq\{1,\ldots,n\}$, $\mathbb{P}\left(\bigcap_{i\in I}\{X_i=u_i\}\right)=\prod_{i\in I}\mathbb{P}(X_i=u_i)$.

**Sur l'indépendance :**

- Si $(X_1,\ldots,X_n)(\Omega)$ n'est pas un ensemble produit, alors $X_1,\ldots,X_n$ ne sont pas indépendantes.
- $X_1,\ldots,X_n$ indépendantes $\Rightarrow$ $\mathbb{P}\left(\bigcap_{i=1}^{n}\{X_i\in\mathcal{D}_i\}\right)=\prod_{i=1}^{n}\mathbb{P}(X_i\in\mathcal{D}_i)$.
- $X_1,\ldots,X_n$ indépendantes $\Rightarrow$ $g_1(X_1),\ldots,g_n(X_n)$ indépendantes.
- $X_1,\ldots,X_n$ indépendantes $\Rightarrow$ $g(X_1,\ldots,X_q)$ et $h(X_{q+1},\ldots,X_n)$ indépendantes.

**Formule du transfert :** $\mathbb{E}\left(g(X_1,\ldots,X_n)\right)=\sum_{(u_1,\ldots,u_n)\in(X_1,\ldots,X_n)(\Omega)}\!\!\!\sum\ldots\sum g(u_1,\ldots,u_n)\mathbb{P}\left(\bigcap_{i=1}^{n}\{X_i=u_i\}\right)$.

**Espérance et indépendance :** $X_1,\ldots,X_n$ indépendantes $\Rightarrow$ $\mathbb{E}\left(\prod_{i=1}^{n}g_i(X_i)\right)=\prod_{i=1}^{n}\mathbb{E}(g_i(X_i))$.

**Loi d'une somme de $n$ *var* indépendantes :**

| $X_i\sim$ | $\mathcal{B}(p)$ | $\mathcal{B}(m_i,p)$ | $\mathcal{P}(\lambda_i)$ | $\mathcal{G}(p)$ |
|---|---|---|---|---|
| $\sum_{i=1}^{n}X_i\sim$ | $\mathcal{B}(n,p)$ | $\mathcal{B}\left(\sum_{i=1}^{n}m_i,p\right)$ | $\mathcal{P}\left(\sum_{i=1}^{n}\lambda_i\right)$ | $\mathcal{G}(n,p)$ |

**Matrice de covariance :** $\Sigma=\begin{pmatrix}\mathbb{V}(X_1) & \mathbb{C}(X_1,X_2) & \ldots & \ldots & \mathbb{C}(X_1,X_n)\\ \mathbb{C}(X_2,X_1) & \mathbb{V}(X_2) & \ldots & \ldots & \mathbb{C}(X_2,X_n)\\ \vdots & & \ddots & & \vdots\\ \vdots & & & \ddots & \mathbb{C}(X_{n-1},X_n)\\ \mathbb{C}(X_n,X_1) & \ldots & \ldots & \ldots & \mathbb{V}(X_n)\end{pmatrix}$.

**Bilinéarité de la covariance :** $\mathbb{C}\left(\sum_{i=1}^{n}a_iX_i,\sum_{j=1}^{q}b_jY_j\right)=\sum_{i=1}^{n}\sum_{j=1}^{q}a_ib_j\mathbb{C}(X_i,Y_j)$.

**Espérance et variance d'une somme de $n$ *var* :** $\mathbb{E}\left(\sum_{i=1}^{n}X_i\right)=\sum_{i=1}^{n}\mathbb{E}(X_i)$,

$\mathbb{V}\left(\sum_{i=1}^{n}X_i\right)=\sum_{i=1}^{n}\mathbb{V}(X_i)+2\sum_{i=2}^{n}\sum_{j=1}^{i-1}\mathbb{C}(X_i,X_j)$. $X_1,\ldots,X_n$ indépendantes $\Rightarrow$ $\mathbb{V}\left(\sum_{i=1}^{n}X_i\right)=\sum_{i=1}^{n}\mathbb{V}(X_i)$.

**Paramètres d'un vecteur de *var* :** Les principaux paramètres d'un vecteur (colonne) de *var* $X=(X_1,\ldots,X_n)^t$ sont son espérance : $\mathbb{E}_n(X)=(\mathbb{E}(X_1),\ldots,\mathbb{E}(X_n))^t$, et sa matrice de covariance $\Sigma$.

# 8 Convergences de suites de *var* discrètes

**Suite de *var* indépendantes :** $(X_n)_{n\in\mathbb{N}^*}$ est une suite de *var* discrètes indépendantes $\Leftrightarrow$ toute sous-famille finie extraite de $(X_n)_{n\in\mathbb{N}^*}$ est une famille de *var* discrètes indépendantes.

**Convergence en probabilité :** Une suite de *var* discrètes $(X_n)_{n\in\mathbb{N}^*}$ converge en probabilité vers $X$ $\Leftrightarrow$ pour tout $\epsilon > 0$, on a $\lim_{n\to\infty} \mathbb{P}(|X_n - X| \geq \epsilon) = 0$.

**Loi faible des grands nombres :** Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de *var iid* discrètes admettant un moment d'ordre 2. Pour tout $n \in \mathbb{N}^*$, on pose $\overline{X}_n = \dfrac{1}{n}\sum_{i=1}^{n} X_i$. Alors $(\overline{X}_n)_{n\in\mathbb{N}^*}$ converge en probabilité vers $\mathbb{E}(X_1)$.

**Convergence en loi :** Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de *var* discrètes.

$(X_n)_{n\in\mathbb{N}^*}$ converge en loi vers $X$ $\Leftrightarrow$ $\lim\limits_{n\to\infty} \mathbb{P}(X_n = k) = \mathbb{P}(X = k)$ pour tout $k \in X(\Omega)$ $\Leftrightarrow$ pour toute fonction $g : \mathbb{R} \to \mathbb{R}$ continue bornée, on a $\lim\limits_{n\to\infty} \mathbb{E}(g(X_n)) = \mathbb{E}(g(X))$ $\Leftrightarrow$ $\lim\limits_{n\to\infty} G_{X_n}(s) = G_X(s)$, $s \in [-1, 1]$.

**Hiérarchie :**

- Convergence en probabilité vers $X$ $\Rightarrow$ convergence en loi vers $X$.
- Convergence en probabilité vers 0 $\Leftrightarrow$ convergence en loi vers 0.

**Approximations usuelles :**

- Si $p = \dfrac{l}{l+m}$ et $n \leq \dfrac{l+m}{10}$, alors $\mathcal{H}(l, m, n) \approx \mathcal{B}(n, p)$,
- Si $n \geq 31$ et $np \leq 10$, alors $\mathcal{B}(n, p) \approx \mathcal{P}(np)$.

# 9 Variables aléatoires réelles à densité

**Densité :** On appelle densité toute fonction $f : \mathbb{R} \to \mathbb{R}$ vérifiant :

- pour tout $x \in \mathbb{R}$, $f(x) \geq 0$,
- $\int_{-\infty}^{\infty} f(x)dx = 1$.

***Var* à densité :** On dit qu'une *var* $X$ est à densité s'il existe une densité $f$ telle que, pour tous $a \leq b$, on a $\mathbb{P}(a \leq X \leq b) = \int_a^b f(x)dx$. On dit alors que $X$ est de densité $f$ ou $X$ possède la densité $f$ ou $f$ est une densité de $X$. La loi de $X$ est caractérisée par $f$.

**Non-unicité de la densité d'une *var* :** Si $X$ possède la densité $f$, alors elle possède d'autres densités. En effet, si $g$ est est une fonction égale à $f$ sauf en quelques points alors $g$ est aussi une densité de $X$.

**Calculs usuels :**

| $\mathbb{P}(X = a)$ | $\mathbb{P}(X \leq b)$ | $\mathbb{P}(X \geq a)$ | $\mathbb{P}(a \leq X \leq b)$ |
|---|---|---|---|
| 0 | $\int_{-\infty}^b f(x)dx$ | $\int_a^{\infty} f(x)dx$ | $\int_{-\infty}^b f(x)dx - \int_{-\infty}^a f(x)dx$ |

Ainsi, pour tout $a \in \mathbb{R}$, on a $\mathbb{P}(X = a) = 0$ (en cela, le comportement d'une *var* à densité diffère radicalement de celui d'une *var* discrète). On peut donc remplacer les inégalités larges par des inégalités strictes :

$\mathbb{P}(a \leq X \leq b) = \mathbb{P}(a < X \leq b) = \mathbb{P}(a \leq X < b) = \mathbb{P}(a < X < b) \ldots$

**Fonction de répartition :** $F(x) = \mathbb{P}(X \leq x) = \int_{-\infty}^x f(t)dt$, $x \in \mathbb{R}$. La fonction de répartition caractérise la loi de $X$ ; pour tous $a \leq b$, on a $\mathbb{P}(a \leq X \leq b) = F(b) - F(a)$.

**Support :** Soit $F$ la fonction de répartition d'une *var* $X$ de densité $f$. Alors $X(\Omega)$ est l'adhérence de $\{x \in \mathbb{R};\ F(x) \in ]0,1[\}$. Si $\{x \in \mathbb{R};\ f(x) \neq 0\}$ est un intervalle ou $\mathbb{R}$ *(ce qui est souvent le cas)*, alors $X(\Omega)$ est aussi l'adhérence de $\{x \in \mathbb{R};\ f(x) \neq 0\}$.

**Exemples :**

| Loi | uniforme | exponentielle | gamma | normale |
|---|---|---|---|---|
| paramètres | $(a,b) \in \mathbb{R}^2,\ \ a < b$ | $\lambda > 0$ | $m > 0,\ \lambda > 0$ | $\mu \in \mathbb{R},\ \ \sigma > 0$ |
| $X \sim$ | $\mathcal{U}([a,b])$ | $\mathcal{E}(\lambda)$ | $\Gamma(m,\lambda)$ | $\mathcal{N}(\mu,\sigma^2)$ |
| $X(\Omega)$ | $[a,b]$ | $[0,\infty[$ | $[0,\infty[$ | $\mathbb{R}$ |
| $f(x)$ | $\dfrac{1}{b-a}$ | $\lambda e^{-\lambda x}$ | $\dfrac{\lambda^m}{\Gamma(m)} x^{m-1} e^{-\lambda x}$ | $\dfrac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ |

**Propriétés :** $F(x) \in [0,1]$, $F$ est continue sur $\mathbb{R}$, $F$ est croissante, $\lim_{x\to-\infty} F(x) = 0$ et $\lim_{x\to\infty} F(x) = 1$.

**Densité et fonction de répartition :** Soit $X$ une *var* à densité de fonction de répartition $F$. Alors une densité $f$ de $X$ est donnée par $f(x) = F'(x)$ pour tout $x \in X(\Omega)$ sauf les points où $F$ n'est pas dérivable, et par ce qu'on veut ailleurs (du moment que $f$ reste une densité).

**Égalité en loi :** $X$ et $Y$ suivent la même loi $\Leftrightarrow$ $F_X(x) = F_Y(x)$ pour tout $x \in \mathbb{R}$ $\Leftrightarrow$ $f_X(x) = f_Y(x)$ pour tout $x \in \mathbb{R}$ sauf, éventuellement, en un nombre fini de points.

***Var* symétrique :** $X$ est symétrique $\Leftrightarrow$ $X$ et $-X$ suivent la même loi.

$X$ possède une densité $f$ paire $\Rightarrow$ $X$ est symétrique.

*Exemple :* $X \sim \mathcal{N}(0,1)$, *i.e. de densité :* $f(x) = \dfrac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$, $x \in \mathbb{R}$, *est symétrique car* $f$ *est paire.*

**Propriétés :** $X$ symétrique $\Rightarrow$ $\mathbb{P}(X \leq -x) = 1 - \mathbb{P}(X \leq x)$, $x \in \mathbb{R}$, $\mathbb{P}(|X| \leq x) = 2\mathbb{P}(X \leq x) - 1$, $x \geq 0$.

**Espérance :** $\mathbb{E}(X) = \int_{-\infty}^{\infty} x f(x) dx$.

**Formule du transfert :** $\mathbb{E}(g(X)) = \int_{-\infty}^{\infty} g(x) f(x) dx$.

**Propriétés :** $\mathbb{E}(a) = a$, $X(\Omega) \subseteq [0,\infty[$ $\Rightarrow$ $\mathbb{E}(X) \geq 0$, $\mathbb{E}(aX+b) = a\mathbb{E}(X) + b$, $\mathbb{E}(aX^2 + bX + c) = a\mathbb{E}(X^2) + b\mathbb{E}(X) + c$.

**Moment d'ordre $p$ :** $m_p = \mathbb{E}(X^p) = \int_{-\infty}^{\infty} x^p f(x) dx$.

**Méthode de la fonction muette :** Si, pour toute fonction continue bornée $g : \mathbb{R} \to \mathbb{R}$, on a $\mathbb{E}(g(X)) = \mathbb{E}(g(Y))$, alors $X$ et $Y$ suivent la même loi.

**Espérance et *var* symétrique :** $X$ est symétrique $\Rightarrow$ pour toute fonction $g : \mathbb{R} \to \mathbb{R}$ impaire (si existence) $\mathbb{E}(g(X)) = 0$.

**Variance :** $\mathbb{V}(X) = \mathbb{E}((X - \mathbb{E}(X))^2)$.

**Propriétés :** $\mathbb{V}(X) \geq 0$, $\mathbb{V}(X) = 0$ $\Leftrightarrow$ $X$ est une *var* constante, $\mathbb{V}(aX+b) = a^2\mathbb{V}(X)$.

**Formule de König-Huyghens :** $\mathbb{V}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$.

**Écart-type :** $\sigma(X) = \sqrt{\mathbb{V}(X)}$.

**Exemples :**

| $X \sim$ | $\mathcal{U}([a,b])$ | $\mathcal{E}(\lambda)$ | $\Gamma(m,\lambda)$ | $\mathcal{N}(\mu,\sigma^2)$ |
|---|---|---|---|---|
| $\mathbb{E}(X)$ | $\dfrac{a+b}{2}$ | $\dfrac{1}{\lambda}$ | $\dfrac{m}{\lambda}$ | $\mu$ |
| $\mathbb{V}(X)$ | $\dfrac{(b-a)^2}{12}$ | $\dfrac{1}{\lambda^2}$ | $\dfrac{m}{\lambda^2}$ | $\sigma^2$ |
| $\sigma(X)$ | $\dfrac{b-a}{2\sqrt{3}}$ | $\dfrac{1}{\lambda}$ | $\dfrac{\sqrt{m}}{\lambda}$ | $\sigma$ |

**Inégalité de Markov :** Soit $h : \mathbb{R} \to ]0, \infty[$ une fonction croissante. Pour tout $\epsilon > 0$, $\mathbb{P}(|X| \geq \epsilon) \leq \dfrac{\mathbb{E}(h(|X|))}{h(\epsilon)}$.

**Inégalité de Bienaymé-Tchebychev :** Pour tout $\epsilon > 0$, $\mathbb{P}(|X - \mathbb{E}(X)| \geq \epsilon) \leq \dfrac{\mathbb{V}(X)}{\epsilon^2}$.

**Transformée de Laplace :** $L(t) = \mathbb{E}(e^{tX}) = \int_{-\infty}^{\infty} e^{tx} f(x) dx$, $t \in \mathbb{R}$.

**Transformée de Laplace et ses dérivées :** $L'(t) = \mathbb{E}(Xe^{tX})$, $L''(t) = \mathbb{E}(X^2 e^{tX})$, $L^{(k)}(t) = \mathbb{E}(X^k e^{tX})$.

**Propriétés :** $\mathbb{E}(X) = L'(0)$, $\mathbb{V}(X) = L''(0) - (L'(0))^2$.

**Transformée de Laplace et égalité en loi :** $X$ et $Y$ suivent la même loi $\Leftrightarrow$ $L_X(t) = L_Y(t)$ pour tout $t \in \mathbb{R}$.

**Fonction caractéristique :** $\varphi(t) = \mathbb{E}(e^{itX}) = \int_{-\infty}^{\infty} e^{itx} f(x) dx$, $t \in \mathbb{R}$.

**Exemples :**

| $X \sim$ | $\mathcal{U}([a,b])$ | $\mathcal{E}(\lambda)$ | $\Gamma(m,\lambda)$ | $\mathcal{N}(\mu,\sigma^2)$ |
|---|---|---|---|---|
| $\varphi(t)$ | $\dfrac{e^{itb} - e^{ita}}{it(b-a)}$ | $\dfrac{\lambda}{\lambda - it}$ | $\dfrac{\lambda^m}{(\lambda - it)^m}$ | $e^{it\mu - \sigma^2 \frac{t^2}{2}}$ |

**Fonction caractéristique et densité :** $f(x) = \dfrac{1}{2\pi} \int_{-\infty}^{\infty} \varphi(t) e^{-itx} dt$, $x \in \mathbb{R}$.

**Fonction caractéristique et ses dérivées :** $\varphi'(t) = i\mathbb{E}(Xe^{itX})$, $\varphi''(t) = -\mathbb{E}(X^2 e^{itX})$, $\varphi^{(k)}(t) = i^k \mathbb{E}(X^k e^{itX})$.

**Fonction caractéristique et égalité en loi :** $X$ et $Y$ suivent la même loi $\Leftrightarrow$ $\varphi_X(t) = \varphi_Y(t)$ pour tout $t \in \mathbb{R}$.

## 10 Retour sur la loi normale

**Loi normale :** *Rappel :* La *var* $X$ suit la loi normale $\mathcal{N}(\mu, \sigma^2)$, $\mu \in \mathbb{R}$, $\sigma > 0$, si elle possède la densité :

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \qquad x \in \mathbb{R}.$$

La loi normale est particulièrement intéressante car

- elle modélise correctement de nombreuses mesures physiques pouvant prendre une infinité de valeurs (avec les décimales) : poids, tailles, durées, vitesses...
- elle est une loi limite universelle. Cela découle du théorème central limite présenté plus loin.

**Graphe :** Le graphe de $f$ est symétrique par rapport à $\mu$. Celui-ci est en forme de "cloche" plus ou moins arrondie selon les valeurs de $\sigma$.

**Espérance, variance et écart-type :** $X \sim \mathcal{N}(\mu, \sigma^2) \Rightarrow \mathbb{E}(X) = \mu$, $\mathbb{V}(X) = \sigma^2$, $\sigma(X) = \sigma$. $\mu$ est la valeur moyenne de $X$ et $\sigma$ mesure la dispersion des valeurs de $X$ autour de $\mu$.

**Loi normale centrée réduite :** La *var* $X$ suit la loi normale centrée réduite $\mathcal{N}(0,1)$ si elle possède la densité :

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \qquad x \in \mathbb{R}.$$

On a $\mathbb{E}(X) = 0$, d'où le "centrée", et $\mathbb{V}(X) = 1$, d'où le "réduite". De plus, $X$ est symétrique. Cela implique :

- pour tout $x \in \mathbb{R}$, on a $\mathbb{P}(-X \leq x) = \mathbb{P}(X \leq x)$,
- pour tout $x \in \mathbb{R}$, on a $\mathbb{P}(X \leq -x) = 1 - \mathbb{P}(X \leq x)$,
- pour tout $x \geq 0$, on a $\mathbb{P}(|X| \leq x) = 2\mathbb{P}(X \leq x) - 1$,
- pour toute fonction impaire $g$ (si existence) on a $\mathbb{E}(g(X)) = 0$.

**Propriétés :** $X \sim \mathcal{N}(\mu, \sigma^2) \;\Rightarrow$

- $X + b \sim \mathcal{N}(\mu + b, \sigma^2)$, $aX \sim \mathcal{N}(a\mu, a^2\sigma^2)$ et $aX + b \sim \mathcal{N}(a\mu + b, a^2\sigma^2)$,
- $Z = \dfrac{X - \mu}{\sigma} \sim \mathcal{N}(0, 1)$. Ainsi, la transformation opérée par $Z$ a centrée et réduite la *var* $X$. Cette transformation est souvent utilisée en pratique.

**Règle des $4\sigma$ :** $X \sim \mathcal{N}(\mu, \sigma^2) \;\Rightarrow\; \mathbb{P}(\mu - 4\sigma \leq X \leq \mu + 4\sigma) \simeq 1$. Ainsi, la plupart des réalisations d'une *var* $X \sim \mathcal{N}(\mu, \sigma^2)$ sont contenues dans l'intervalle $[\mu - 4\sigma, \mu + 4\sigma]$.

**Moments :** $X \sim \mathcal{N}(0, 1) \;\Rightarrow\; \mathbb{E}(X^{2k}) = \prod_{i=1}^{k}(2i - 1)$, $\mathbb{E}(X^{2k+1}) = 0$.

**Fonction de répartition :** $X \sim \mathcal{N}(0, 1) \;\Rightarrow\; F(x) = \mathbb{P}(X \leq x) = \int_{-\infty}^{x} \dfrac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$, $x \in \mathbb{R}$.

On a $F(x) = F(y) \;\Leftrightarrow\; x = y$, et $f(x) = F'(x)$, $x \in \mathbb{R}$.

La fonction $F$ ne peut pas s'écrire sous une forme analytique, sans intégrale. C'est pourquoi on utilise une table de valeurs associée donnant $F(x)$ avec $x \in [0; 3, 99]$ et $x = x_1 + x_2$.

**Approximation : loi binomiale et loi normale :** Si $n \geq 31$, $np \geq 5$ et $n(1 - p) \geq 5$, alors

$\mathcal{B}(n, p) \approx \mathcal{N}(np, np(1 - p))$. Avec une correction de continuité, $X \sim \mathcal{B}(n, p)$ et $Y \sim \mathcal{N}(np, np(1 - p)) \;\Rightarrow$

- pour tout $k \in \mathbb{N}$, on a $\mathbb{P}(X \leq k) \simeq \mathbb{P}(Y \leq k + 0, 5)$,
  $\mathbb{P}(X \geq k) \simeq \mathbb{P}(Y \geq k - 0, 5)$ et $\mathbb{P}(X = k) \simeq \mathbb{P}(k - 0, 5 \leq Y \leq k + 0, 5)$,
- pour tout $(k_1, k_2) \in \mathbb{N}^2$ tel que $k_1 < k_2$, on a $\mathbb{P}(k_1 \leq X \leq k_2) \simeq \mathbb{P}(k_1 - 0, 5 \leq Y \leq k_2 + 0, 5)$.

**Approximation : loi de Poisson et la loi normale :** Si $\lambda \geq 15$, alors $\mathcal{P}(\lambda) \approx \mathcal{N}(\lambda, \lambda)$.

## 11 Couples de *var* à densité

**Densité (bidimensionnelle) :** On appelle densité (bidimensionnelle) toute fonction $f : \mathbb{R}^2 \to \mathbb{R}$ vérifiant :

- pour tout $(x, y) \in \mathbb{R}^2$, $f(x, y) \geq 0$,
- $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy = 1$.

**Couple de *var* à densité :** On dit qu'un couple de *var* $(X, Y)$ est à densité s'il existe une densité $f$ telle que, pour tout $\mathcal{D} \subseteq \mathbb{R}^2$, on a $\mathbb{P}((X, Y) \in \mathcal{D}) = \int\int_{\mathcal{D}} f(x, y) dx dy$, ($x$ est associé à $X$ et $y$ est associé à $Y$).

On dit alors que $(X, Y)$ est de densité $f$ ou $(X, Y)$ possède la densité $f$ ou $f$ est une densité de $(X, Y)$. La loi de $(X, Y)$ est caractérisée par $f$.

**Propriétés :** Pour tous $a \leq b$, $c \leq d$ et toutes fonctions

$g : \mathbb{R} \to \mathbb{R}$ et $h : \mathbb{R} \to \mathbb{R}$ telles que $g(x) \leq h(x)$, $x \in \mathbb{R}$, on a

- $\mathbb{P}(\{a \leq X \leq b\} \cap \{c \leq Y \leq d\}) = \int_c^d \left( \int_a^b f(x, y) dx \right) dy$,
- $\mathbb{P}(\{a \leq X \leq b\} \cap \{g(X) \leq Y \leq h(X)\}) = \int_a^b \left( \int_{g(x)}^{h(x)} f(x, y) dy \right) dx$,
- $\mathbb{P}(g(X) \leq Y \leq h(X)) = \int_{-\infty}^{\infty} \left( \int_{g(x)}^{h(x)} f(x, y) dy \right) dx$.

**Fonction de répartition :** $F(x, y) = \mathbb{P}(\{X \leq x\} \cap \{Y \leq y\})$, $(x, y) \in \mathbb{R}^2$. On a $F_X(x) = \lim_{y \to \infty} F(x, y)$, $x \in \mathbb{R}$, et $F_Y(y) = \lim_{x \to \infty} F(x, y)$, $y \in \mathbb{R}$.

**Support :** Soit $F$ la fonction de répartition d'un couple de *var* $(X, Y)$ de densité $f$. Alors $(X, Y)(\Omega)$ est l'adhérence de $\{(x, y) \in \mathbb{R}^2;\ F(x, y) \in ]0, 1[\}$.

Si $\{(x,y)\in\mathbb{R}^2;\ f(x,y)\neq 0\}$ est un ensemble d'aire non nulle "sans points ou courbes isolés" *(ce qui est souvent le cas)*, alors $(X,Y)(\Omega)$ est aussi l'adhérence de $\{(x,y)\in\mathbb{R}^2;\ f(x,y)\neq 0\}$. De plus, $X(\Omega)$ est l'adhérence de l'ensemble des $x$ vérifiant $f(x,y)\neq 0$ pour tout $y\in\mathbb{R}$, et $Y(\Omega)$ est l'adhérence de l'ensemble des $y$ vérifiant $f(x,y)\neq 0$ pour tout $x\in\mathbb{R}$.

**Densité et fonction de répartition :** Soit $(X,Y)$ un couple de *var* à densité de fonction de répartition $F$. Alors une densité $f$ de $(X,Y)$ est donnée par $f(x,y)=\dfrac{\partial^2 F}{\partial x\partial y}(x,y)$ pour tout $(x,y)\in(X,Y)(\Omega)$ sauf les points of $F$ n'est pas différentiable, et par ce qu'on veut ailleurs (du moment que $f$ reste une densité).

**Densités marginales :**

- Une densité de $X$ est $f_X(x)=\int_{-\infty}^{\infty} f(x,y)dy$, $x\in\mathbb{R}$.
- Une densité de $Y$ est $f_Y(y)=\int_{-\infty}^{\infty} f(x,y)dx$, $y\in\mathbb{R}$.
- Si $f(x,y)=f(y,x)$ pour tout $(x,y)\in\mathbb{R}^2$ sauf, éventuellement, sur un ensemble d'aire nulle, alors $X$ et $Y$ suivent la même loi ; une densité de $Y$ est $f_Y(y)=f_X(y)$, $y\in\mathbb{R}$.

**Indépendance de 2 *var* à densité :** $X$ et $Y$ indépendantes $\Leftrightarrow$ $(X,Y)$ possède la densité produit : $f_X(x)f_Y(y)$, *i.e.* $f(x,y)=f_X(x)f_Y(y)$ pour tout $(x,y)\in\mathbb{R}^2$ sauf, éventuellement, sur un ensemble d'aire nulle (impliquant $(X,Y)(\Omega)=X(\Omega)\times Y(\Omega)$) $\Leftrightarrow$ $F(x,y)=F_X(x)F_Y(y)$ pour tout $(x,y)\in\mathbb{R}^2$ $\Leftrightarrow$ pour tous $a\leq b$, $c\leq d$, $\mathbb{P}(\{a\leq X\leq b\}\cap\{c\leq Y\leq d\})=\mathbb{P}(a\leq X\leq b)\mathbb{P}(c\leq Y\leq d)$.

**Sur l'indépendance :**

- Si $(X,Y)(\Omega)$ n'est pas un ensemble produit, *i.e.* on ne peut pas écrire $(X,Y)(\Omega)=I\times J$ avec $(I,J)\subseteq\mathbb{R}^2$, alors $X$ et $Y$ ne sont pas indépendantes.
- S'il existe deux fonctions $g:\mathbb{R}\to\mathbb{R}$ et $h:\mathbb{R}\to\mathbb{R}$ telles que $f(x,y)=g(x)h(y)$, $\int_{-\infty}^{\infty} g(x)dx=1$, $\int_{-\infty}^{\infty} h(y)dy=1$, alors $X$ et $Y$ sont indépendantes, une densité de $X$ est $g$ et une densité de $Y$ est $h$.
- $X$ et $Y$ indépendantes $\Rightarrow$ $\mathbb{P}(\{X\in\mathcal{D}_1\}\cap\{Y\in\mathcal{D}_2\})=\mathbb{P}(X\in\mathcal{D}_1)\mathbb{P}(Y\in\mathcal{D}_2)$.
- $X$ et $Y$ indépendantes $\Rightarrow$ $g(X)$ et $h(Y)$ indépendantes.

**"Indépendantes et identiquement distribuées"/*iid* :** On dit que $X$ et $Y$ sont *iid* quand elles sont "indépendantes et identiquement distribuées/suivent la même loi".

**Loi d'un couple de *var* $(\phi(X,Y),\psi(X,Y))$ :** *(théorème du changement de variable)* Soient $(X,Y)$ un couple de *var* de densité $f_{(X,Y)}$, et $\phi:(X,Y)(\Omega)\to\mathbb{R}$ et $\psi:(X,Y)(\Omega)\to\mathbb{R}$ deux fonctions.

On considère le couple de *var* : $(U,V)=(\phi(X,Y),\psi(X,Y))$. On suppose que $g(x,y)=(\phi(x,y),\psi(x,y))$ est injective et qu'il existe deux fonctions $h:\mathbb{R}^2\to X(\Omega)$ et $k:\mathbb{R}^2\to Y(\Omega)$ différentiables telles que $u=\phi(x,y)$ et $v=\psi(x,y)$ $\Leftrightarrow$ $x=h(u,v)$ et $y=k(u,v)$. Soit $J(u,v)$ le jacobien associé à $(x,y)=(h(u,v),k(u,v))$ défini par le déterminant : $J(u,v)=\begin{vmatrix}\dfrac{\partial h}{\partial u}(u,v) & \dfrac{\partial h}{\partial v}(u,v)\\[2ex] \dfrac{\partial k}{\partial u}(u,v) & \dfrac{\partial k}{\partial v}(u,v)\end{vmatrix}=\dfrac{\partial h}{\partial u}(u,v)\dfrac{\partial k}{\partial v}(u,v)-\dfrac{\partial k}{\partial u}(u,v)\dfrac{\partial h}{\partial v}(u,v).$

Alors une densité de $(U,V)$ est donnée par $f_{(U,V)}(u,v)=f_{(X,Y)}(h(u,v),k(u,v))|J(u,v)|$, $(u,v)\in\mathbb{R}^2$.

**Densité d'une somme de deux *var* $X$ et $Y$ indépendantes :** *(produit de convolution)*

$$f_{X+Y}(x)=(f_X*f_Y)(x)=\int_{-\infty}^{\infty} f_X(u)f_Y(x-u)du,\qquad x\in\mathbb{R}.$$

**Exemples :**

| $X_i \sim$ | $\mathcal{E}(\lambda)$ | $\Gamma(m_i, \lambda)$ | $\mathcal{N}(\mu_i, \sigma_i^2)$ | $\chi^2(\nu_i)$ |
|---|---|---|---|---|
| $X_1 + X_2 \sim$ | $\Gamma(2, \lambda)$ | $\Gamma(m_1 + m_2, \lambda)$ | $\mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$ | $\chi^2(\nu_1 + \nu_2)$ |

**Caractérisations de quelques lois connues :**

- Loi du Chi-deux : $X$ et $Y$ *iid* avec $X \sim \mathcal{N}(0,1) \ \Rightarrow \ X^2 + Y^2 \sim \chi^2(2)$.
- Loi de Student : $X$ et $Y$ indépendantes avec $X \sim \mathcal{N}(0,1)$ et $Y \sim \chi^2(\nu) \ \Rightarrow \ \dfrac{X}{\sqrt{\dfrac{Y}{\nu}}} = \sqrt{\nu}\dfrac{X}{\sqrt{Y}} \sim \mathcal{T}(\nu)$.
- Loi de Fisher : $X$ et $Y$ indépendantes avec $X \sim \chi^2(\nu_1)$ et $Y \sim \chi^2(\nu_2) \ \Rightarrow \ \dfrac{\dfrac{X}{\nu_1}}{\dfrac{Y}{\nu_2}} = \dfrac{\nu_2 X}{\nu_1 Y} \sim \mathcal{F}(\nu_1, \nu_2)$.

**Formule du transfert :** $\mathbb{E}(g(X,Y)) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)f(x,y)dxdy$.

**Méthode de la fonction muette :** Si, pour toute fonction continue bornée $g : \mathbb{R}^2 \to \mathbb{R}$, on a

$\mathbb{E}(g(X,Y)) = \mathbb{E}(g(U,V))$, alors $(X,Y)$ et $(U,V)$ suivent la même loi.

**Linéarité de l'espérance :** $\mathbb{E}(aX + bY) = a\mathbb{E}(X) + b\mathbb{E}(Y)$.

**Inégalité de Cauchy-Schwarz :** $|\mathbb{E}(XY)| \leq \sqrt{\mathbb{E}(X^2)}\sqrt{\mathbb{E}(Y^2)}$.

**Moments et indépendance :** $X$ et $Y$ indépendantes $\ \Rightarrow \ \mathbb{E}(g(X)h(Y)) = \mathbb{E}(g(X))\mathbb{E}(h(Y))$. En particulier, on a $\mathbb{E}(X^k Y^m) = \mathbb{E}(X^k)\mathbb{E}(Y^m)$.

**Covariance :** $\mathbb{C}(X,Y) = \mathbb{E}((X - \mathbb{E}(X))(Y - \mathbb{E}(Y))) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$.

**Matrice de covariance :** $\Sigma = \begin{pmatrix} \mathbb{V}(X) & \mathbb{C}(X,Y) \\ \mathbb{C}(Y,X) & \mathbb{V}(Y) \end{pmatrix}$.

**Variance d'une somme de *var* :** $\mathbb{V}(X+Y) = \mathbb{V}(X) + \mathbb{V}(Y) + 2\mathbb{C}(X,Y)$.

**Propriétés :** $\mathbb{C}(X,Y) = \mathbb{C}(Y,X)$, $\mathbb{C}(X,a) = 0$, $\mathbb{C}(aX+b, cY+d) = ac\mathbb{C}(X,Y)$,

$\mathbb{C}(aX + bY, cU + dV) = ac\mathbb{C}(X,U) + ad\mathbb{C}(X,V) + bc\mathbb{C}(Y,U) + bd\mathbb{C}(Y,V)$,

**Covariance, variance et indépendance :** $X$ et $Y$ indépendantes $\Rightarrow$

- $\mathbb{C}(X,Y) = 0$,
- $\mathbb{V}(X+Y) = \mathbb{V}(X) + \mathbb{V}(Y)$.

$\mathbb{C}(X,Y) = 0 \ \not\Rightarrow$ on ne peut rien dire sur l'indépendance de $X$ et $Y$.

**Coefficient de corrélation linéaire :** $\rho(X,Y) = \dfrac{\mathbb{C}(X,Y)}{\sigma(X)\sigma(Y)}$. On a $\rho(X,Y) \in ]-1,1[$. Plus $|\rho(X,Y)|$ est proche de 1, plus la liaison linéaire entre $X$ et $Y$ est forte.

**Paramètres d'un couple de *var* :** Les principaux paramètres d'un couple de *var* (vecteur colonne) $Z = (X,Y)^t$ sont son espérance : $\mathbb{E}_2(Z) = (\mathbb{E}(X), \mathbb{E}(Y))^t$, et sa matrice de covariance $\Sigma$.

**Propriété :** Soient $Z$ un couple de *var* (vecteur colonne) d'espérance $\mu$ et de matrice de covariance $\Sigma$, $A$ une matrice à 2 lignes et 2 colonne, $c$ un vecteur colonne à 2 composantes et $W = AZ + c$. Alors

- l'espérance de $W$ est $\mu_* = A\mu + c$,
- la matrice de covariance de $W$ est $\Sigma_* = A\Sigma A^t$.

**Couple gaussien :** Soient $X$ et $Y$ deux *var*. Alors $(X,Y)^t$ est un couple gaussien $\Leftrightarrow$

pour tout $(a,b) \in \mathbb{R}^2 - \{(0,0)\}$, la *var* $aX + BY$ suit une loi normale.

En particulier, si $(X,Y)^t$ est gaussien, alors $X$ et $Y$ suivent des lois normales.

**Caractérisation :** Un couple gaussien $(X,Y)^t$ est caractérisé par son espérance $\mu$ et sa matrice de covariance $\Sigma$. La loi de $(X,Y)^t$ est notée $\mathcal{N}_2(\mu,\Sigma)$.

**Propriété d'indépendance :** Si $(X,Y)^t$ est un couple gaussien, alors $\mathbb{C}(X,Y)=0 \;\Rightarrow\; X$ et $Y$ indépendantes.

**Densité et fonction caractéristique d'un couple gaussien :** Soit $(X,Y)^t \sim \mathcal{N}_2(\mu,\Sigma)$ avec $\mu=(\mu_1,\mu_2)^t$.

- Si $\det(\Sigma)>0$, alors $(X,Y)^t$ possède la densité : $f(x,y)=\dfrac{1}{2\pi\sqrt{\det(\Sigma)}}e^{-\frac{1}{2}(x-\mu_1,y-\mu_2)\Sigma^{-1}(x-\mu_1,y-\mu_2)^t}$, $(x,y)\in\mathbb{R}^2$.
- La fonction caractéristique de $(X,Y)^t$ est $\varphi(s,t)=e^{i(\mu_1,\mu_2)(s,t)^t-\frac{1}{2}(s,t)\Sigma(s,t)^t}$, $(s,t)\in\mathbb{R}^2$.

**Cas particulier :** Soit $(X,Y)^t\sim\mathcal{N}_2(\mu,\Sigma)$ avec $\mu=(0,0)^t=0_2$ et $\Sigma=\begin{pmatrix}1 & \rho\\ \rho & 1\end{pmatrix}$, $\rho\in]-1,1[$. Alors

- $(X,Y)^t$ possède la densité : $f(x,y)=\dfrac{1}{2\pi\sqrt{1-\rho^2}}e^{-\frac{1}{2(1-\rho^2)}(x^2+y^2-2\rho xy)}$, $(x,y)\in\mathbb{R}^2$,
- la fonction caractéristique de $(X,Y)^t$ est $\varphi(s,t)=e^{-\frac{1}{2}(s^2+t^2+2\rho st)}$, $(s,t)\in\mathbb{R}^2$.

**Propriété :** Soient $Z\sim\mathcal{N}_2(\mu,\Sigma)$, $A$ une matrice à 2 lignes et 2 colonnes, $c$ un vecteur colonne à 2 composantes et $W=AZ+c$. Alors on a $W\sim\mathcal{N}_2(A\mu+c,A\Sigma A^t)$.

## 12 Vecteurs de *var* à densité

**Densité ($n$-dimensionnelle) :** On appelle densité ($n$-dimensionnelle) toute fonction $f:\mathbb{R}^n\to\mathbb{R}$ vérifiant :

- pour tout $(x_1,\ldots,x_n)\in\mathbb{R}^n$, $f(x_1,\ldots,x_n)\geq 0$,
- $\int_{-\infty}^{\infty}\ldots\int_{-\infty}^{\infty} f(x_1,\ldots,x_n)dx_1\ldots dx_n=1$.

**Vecteur de *var* à densité :** On dit qu'un vecteur de *var* $(X_1,\ldots,X_n)$ est à densité s'il existe une densité $f$, telle que, pour tout $\mathcal{D}\subseteq\mathbb{R}^n$, on a $\mathbb{P}((X_1,\ldots,X_n)\in\mathcal{D})=\int\ldots\int_{\mathcal{D}} f(x_1,\ldots,x_n)dx_1\ldots dx_n$. On dit alors que $(X_1,\ldots,X_n)$ est de densité $f$ ou $(X_1,\ldots,X_n)$ possède la densité $f$ ou $f$ est une densité de $(X_1,\ldots,X_n)$. La loi de $(X_1,\ldots,X_n)$ est caractérisée par $f$.

**Fonction de répartition :** $F(x_1,\ldots,x_n)=\mathbb{P}\left(\bigcap_{i=1}^{n}\{X_i\leq x_i\}\right)$, $(x_1,\ldots,x_n)\in\mathbb{R}^n$.

**Retour sur le support :** Soit $F$ la fonction de répartition d'un vecteur de *var* $(X_1,\ldots,X_n)$ à densité. On définit $(X_1,\ldots,X_n)(\Omega)$ par l'adhérence de $\{(x_1,\ldots,x_n)\in\mathbb{R}^n;\ F(x_1,\ldots,x_n)\in]0,1[\}$.

**Densités marginales :** Une densité de $X_1$ est donnée par $f_{X_1}(x_1)=\int_{-\infty}^{\infty}\ldots\int_{-\infty}^{\infty} f(x_1,\ldots,x_n)dx_2\ldots dx_n$, $x_1\in\mathbb{R}$. On définit de la même manière des densités pour d'autres *var* $X_2,\ldots,X_n$.

**Densité d'un vecteur de *var* composantes de $(X_1,\ldots,X_n)$ :** Une densité de $(X_1,X_2)$ est donnée par $f_{(X_1,X_2)}(x_1,x_2)=\int_{-\infty}^{\infty}\ldots\int_{-\infty}^{\infty} f(x_1,\ldots,x_n)dx_3\ldots dx_n$, $(x_1,x_2)\in\mathbb{R}^2$. On définit de la même manière des densités d'autres vecteurs de *var*.

**Indépendance de $n$ *var* à densité :** $X_1,\ldots,X_n$ indépendantes $\Leftrightarrow$ $(X_1,\ldots,X_n)$ possède la densité produit : $\prod_{i=1}^{n} f_{X_i}(x_i)$, *i.e.* $f(x_1,\ldots,x_n)=\prod_{i=1}^{n} f_{X_i}(x_i)$ pour tout $(x_1,\ldots,x_n)\in\mathbb{R}^n$ sauf, éventuellement, sur un ensemble de volume nul (impliquant $(X_1,\ldots,X_n)(\Omega)=X_1(\Omega)\times\ldots\times X_n(\Omega)$) $\Leftrightarrow$ $F(x_1,\ldots,x_n)=\prod_{i=1}^{n} F_{X_i}(x_i)$ pour tout $(x_1,\ldots,x_n)\in\mathbb{R}^n \Leftrightarrow$ pour tous $a_i\leq b_i$, $i\in\{1,\ldots,n\}$, $\mathbb{P}\left(\bigcap_{i=1}^{n}\{a_i\leq X_i\leq b_i\}\right)=\prod_{i=1}^{n}\mathbb{P}(a_i\leq X_i\leq b_i)$.

**Sur l'indépendance :**

- Si $(X_1, \ldots, X_n)(\Omega)$ n'est pas un ensemble produit, alors $X_1, \ldots, X_n$ ne sont pas indépendantes.
- $X_1, \ldots, X_n$ indépendantes $\Rightarrow$ $\mathbb{P}\left(\bigcap_{i=1}^{n} \{X_i \in \mathcal{D}_i\}\right) = \prod_{i=1}^{n} \mathbb{P}(X_i \in \mathcal{D}_i)$.
- $X_1, \ldots, X_n$ indépendantes $\Rightarrow$ $g_1(X_1), \ldots, g_n(X_n)$ indépendantes.
- $X_1, \ldots, X_n$ indépendantes $\Rightarrow$ $g(X_1, \ldots, X_q)$ et $h(X_{q+1}, \ldots, X_n)$ indépendantes.

**Formule du transfert :** $\mathbb{E}\left(g(X_1, \ldots, X_n)\right) = \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} g(x_1, \ldots, x_n) f(x_1, \ldots, x_n) dx_1 \ldots dx_n$.

**Espérance et indépendance :** $X_1, \ldots, X_n$ indépendantes $\Rightarrow$ $\mathbb{E}\left(\prod_{i=1}^{n} g_i(X_i)\right) = \prod_{i=1}^{n} \mathbb{E}(g_i(X_i))$.

**Quelques lois usuelles d'une somme de $n$ *var* indépendantes :**

| $X_i \sim$ | $\mathcal{E}(\lambda)$ | $\Gamma(m_i, \lambda)$ | $\mathcal{N}(\mu_i, \sigma_i^2)$ | $\chi^2(\nu_i)$ |
|---|---|---|---|---|
| $\sum_{i=1}^{n} X_i \sim$ | $\Gamma(n, \lambda)$ | $\Gamma\left(\sum_{i=1}^{n} m_i, \lambda\right)$ | $\mathcal{N}\left(\sum_{i=1}^{n} \mu_i, \sum_{i=1}^{n} \sigma_i^2\right)$ | $\chi^2\left(\sum_{i=1}^{n} \nu_i\right)$ |

Ainsi, toute *var* définie comme une combinaison linéaire de $n$ *var* indépendantes suivant chacune une loi normale suit une loi normale.

**Résultat clé sur la loi du Chi-deux :** $X_1, \ldots, X_n$ *iid*, $X_i \sim \mathcal{N}(0,1)$ $\Rightarrow$ $\sum_{i=1}^{n} X_i^2 \sim \chi^2(n)$.

**Matrice de covariance :** $\Sigma = \begin{pmatrix} \mathbb{V}(X_1) & \mathbb{C}(X_1, X_2) & \ldots & \ldots & \mathbb{C}(X_1, X_n) \\ \mathbb{C}(X_2, X_1) & \mathbb{V}(X_2) & \ldots & \ldots & \mathbb{C}(X_2, X_n) \\ \vdots & & \ddots & & \vdots \\ \vdots & & & \ddots & \mathbb{C}(X_{n-1}, X_n) \\ \mathbb{C}(X_n, X_1) & \ldots & & \ldots\ \ldots & \mathbb{V}(X_n) \end{pmatrix}$.

**Espérance et variance d'une somme de $n$ *var* :** $\mathbb{E}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathbb{E}(X_i)$,

$\mathbb{V}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathbb{V}(X_i) + 2\sum_{i=2}^{n}\sum_{j=1}^{i-1} \mathbb{C}(X_i, X_j)$. $X_1, \ldots, X_n$ indépendantes $\Rightarrow \mathbb{V}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathbb{V}(X_i)$.

**Bilinéarité de la covariance :** $\mathbb{C}\left(\sum_{i=1}^{n} a_i X_i, \sum_{j=1}^{q} b_j Y_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{q} a_i b_j \mathbb{C}(X_i, Y_j)$.

**Paramètres d'un vecteur de *var* :** Les principaux paramètres d'un vecteur (colonne) de *var* $X = (X_1, \ldots, X_n)^t$ sont son espérance : $\mathbb{E}_n(X) = (\mathbb{E}(X_1), \ldots, \mathbb{E}(X_n))^t$, et sa matrice de covariance $\Sigma$.

**Propriété :** Soient $Z$ un vecteur de $n$ *var* d'espérance $\mu$ et de matrice de covariance $\Sigma$, $A$ une matrice à $n$ colonnes et $q$ lignes, $c$ un vecteur colonne à $q$ composantes et $W = AZ + c$. Alors

- l'espérance de $W$ est $\mu_* = A\mu + c$,
- la matrice de covariance de $W$ est $\Sigma_* = A\Sigma A^t$.

**Fonction caractéristique d'un vecteur de *var* à densité :** $\varphi(t_1, \ldots, t_n) = \mathbb{E}(e^{it_1X_1 + \ldots + it_nX_n})$, $(t_1, \ldots, t_n) \in \mathbb{R}^n$.

**Vecteur gaussien :** Soient $X_1, \ldots, X_n$ $n$ *var*. Alors $(X_1, \ldots, X_n)^t$ est un vecteur gaussien $\Leftrightarrow$ toute les combinaisons linéaires de $X_1, \ldots, X_n$ suivent une loi normale : pour tout $(a_1, \ldots, a_n) \in \mathbb{R}^n - \{(0, \ldots, 0)\}$, $\sum_{i=1}^{n} a_i X_i$ suit une loi normale.

En particulier, si $(X_1, \ldots, X_n)^t$ est un vecteur gaussien, alors, pour tout $i \in \{1, \ldots, n\}$, $X_i$ suit une loi normale.

**Caractérisation :** Un vecteur gaussien $X = (X_1, \ldots, X_n)^t$ est caractérisé par son espérance $\mu$ et sa matrice de covariance $\Sigma$. La loi de $X$ est notée $\mathcal{N}_n(\mu, \Sigma)$.

**Propriété d'indépendance :** Si $(X_1, \ldots, X_n)^t$ est un vecteur gaussien alors $X_i$ et $X_j$ indépendantes $\Leftrightarrow$ $\mathbb{C}(X_i, X_j) = 0$.

**Densité et fonction caractéristique d'un vecteur gaussien :** Soit $X = (X_1, \ldots, X_n)^t \sim \mathcal{N}_n(\mu, \Sigma)$ avec $\mu = (\mu_1, \ldots, \mu_n)^t$.

- Si $\det(\Sigma) > 0$, alors $X$ possède la densité :
  $f(x_1, \ldots, x_n) = \frac{1}{(2\pi)^{\frac{n}{2}} \sqrt{\det(\Sigma)}} e^{-\frac{1}{2}(x_1-\mu_1,\ldots,x_n-\mu_n)\Sigma^{-1}(x_1-\mu_1,\ldots,x_n-\mu_n)^t}$, $(x_1, \ldots, x_n) \in \mathbb{R}^n$.
- La fonction caractéristique de $X$ est $\varphi(t_1, \ldots, t_n) = e^{i(\mu_1,\ldots,\mu_d)(t_1,\ldots,t_n)^t - \frac{1}{2}(t_1,\ldots,t_n)\Sigma(t_1,\ldots,t_n)^t}$, $(t_1, \ldots, t_n) \in \mathbb{R}^n$.

**Propriété :** Soient $Z \sim \mathcal{N}_n(\mu, \Sigma)$, $A$ une matrice à $n$ colonnes et $q$ lignes, $c$ un vecteur colonne à $q$ composantes et $W$ tel que $W = AZ + c$. Alors on a $W \sim \mathcal{N}_q(A\mu + c, A\Sigma A^t)$.

**Lois usuelles associées à la loi normale :** Soient $X_1, \ldots, X_n$ $n$ *var iid* suivant chacune la loi normale $\mathcal{N}(\mu, \sigma^2)$. On pose $\overline{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$, $S_n^2 = \frac{1}{n-1} \sum_{i=1}^{n} \left(X_i - \overline{X}_n\right)^2$. Alors $\mathbb{E}(\overline{X}_n) = \mu$, $\mathbb{E}(S_n^2) = \sigma^2$ et

| *var* | $\sqrt{n}\left(\frac{\overline{X}_n - \mu}{\sigma}\right)$ | $\frac{n-1}{\sigma^2} S_n^2$ | $\sqrt{n}\left(\frac{\overline{X}_n - \mu}{S_n}\right)$ | $\sum_{i=1}^{n} \left(\frac{X_i - \mu}{\sigma}\right)^2$ |
|---|---|---|---|---|
| Loi | $\mathcal{N}(0,1)$ | $\chi^2(n-1)$ | $\mathcal{T}(n-1)$ | $\chi^2(n)$ |

## 13 Convergences de suites de *var*; généralité

**Suite de *var* indépendantes :** $(X_n)_{n \in \mathbb{N}^*}$ est une suite de *var* indépendantes $\Leftrightarrow$ toute sous-famille finie extraite de $(X_n)_{n \in \mathbb{N}^*}$ est une famille de *var* indépendantes.

**Convergence en probabilité :** $(X_n)_{n \in \mathbb{N}^*}$ converge en probabilité vers $X$ $\Leftrightarrow$ pour tout $\epsilon > 0$, on a $\lim_{n \to \infty} \mathbb{P}(|X_n - X| \geq \epsilon) = 0$.

**Loi faible des grands nombres :** Soit $(X_n)_{n \in \mathbb{N}^*}$ une suite de *var iid* admettant un moment d'ordre 2. Pour tout $n \in \mathbb{N}^*$, on pose $\overline{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$. Alors $(\overline{X}_n)_{n \in \mathbb{N}^*}$ converge en probabilité vers $\mathbb{E}(X_1)$.

**Convergence en loi :** Une suite de *var* $(X_n)_{n \in \mathbb{N}^*}$ converge en loi vers $X$ $\Leftrightarrow$ $\lim_{n \to \infty} F_{X_n}(x) = F_X(x)$ en tout point de continuité $x$ de $F_X$ $\Leftrightarrow$ pour toute fonction $g : \mathbb{R} \to \mathbb{R}$ continue bornée, on a $\lim_{n \to \infty} \mathbb{E}(g(X_n)) = \mathbb{E}(g(X))$ $\Leftrightarrow$ $\lim_{n \to \infty} \varphi_{X_n}(t) = \varphi_X(t)$, $t \in \mathbb{R}$.

**Théorème central limite :** Soit $(X_n)_{n \in \mathbb{N}^*}$ une suite de *var iid* admettant un moment d'ordre 2. Pour tout $n \in \mathbb{N}^*$, on pose $\overline{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$, et $Z_n = \frac{\overline{X}_n - \mathbb{E}(\overline{X}_n)}{\sigma(\overline{X}_n)} = \sqrt{n}\left(\frac{\overline{X}_n - \mathbb{E}(X_1)}{\sigma(X_1)}\right)$. Alors la suite de *var* $(Z_n)_{n \in \mathbb{N}^*}$ converge en loi vers une *var* $Z$ suivant la loi normale $\mathcal{N}(0,1)$.

**Théorème de continuité (continuous mapping theorem) :**

- $(X_n)_{n \in \mathbb{N}^*}$ converge en probabilité vers $X$ $\Rightarrow$ pour toute fonction $g : X(\Omega) \to \mathbb{R}$ continue, sauf, éventuellement, en un nombre fini de points, $(g(X_n))_{n \in \mathbb{N}^*}$ converge en probabilité vers $g(X)$.
- $(X_n)_{n \in \mathbb{N}^*}$ converge en loi vers $X$ $\Rightarrow$ pour toute fonction $g : X(\Omega) \to \mathbb{R}$ continue, sauf, éventuellement, en un nombre fini de points, $(g(X_n))_{n \in \mathbb{N}^*}$ converge en loi vers $g(X)$.

**Convergence dans $\mathbb{L}^p$ :** $(X_n)_{n \in \mathbb{N}^*}$ converge dans $\mathbb{L}^p$ vers $X$ $\Leftrightarrow$ $\lim_{n \to \infty} \mathbb{E}\left(|X_n - X|^p\right) = 0$.

**Hiérarchie :**

- Convergence dans $\mathbb{L}^p$ vers $X$ $\Rightarrow$ convergence en probabilité vers $X$ $\Rightarrow$ convergence en loi vers $X$.
- Convergence en probabilité vers 0 $\Leftrightarrow$ convergence en loi vers 0.

**Convergence presque sûre :** $(X_n)_{n\in\mathbb{N}^*}$ converge presque sûrement ($ps$) vers $X$ $\Leftrightarrow$

$\mathbb{P}(\lim_{n\to\infty} X_n = X) = \mathbb{P}(\{\omega \in \Omega;\ \lim_{n\to\infty} X_n(\omega) = X(\omega)\}) = 1.$

**Critère de convergence** ***ps*** **:** Pour tout $\epsilon > 0$, $\sum_{n=1}^{\infty} \mathbb{P}(|X_n - X| \geq \epsilon) < \infty$ $\Rightarrow$ $(X_n)_{n\in\mathbb{N}^*}$ converge $ps$ vers $X$.

**Loi forte des grands nombres :** Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de *var iid* admettant une espérance. Pour tout $n \in \mathbb{N}^*$, on pose $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$. Alors $(\overline{X}_n)_{n\in\mathbb{N}^*}$ converge $ps$ vers $\mathbb{E}(X_1)$.

**Hiérarchie :** Convergence $ps$ $\Rightarrow$ convergence en probabilité $\Rightarrow$ convergence en loi.

Il n'y a pas d'implication entre la convergence dans $\mathbb{L}^p$ et la convergence $ps$.

# 14 Introduction à l'estimation paramétrique

**$n$-échantillon :** On appelle $n$-échantillon d'une *var* $X$ $n$ *var* $X_1, \ldots, X_n$ *iid* suivant la même loi que celle de $X$.

**Estimation paramétrique :** Soit $X$ une *var* dont la loi dépend d'au moins un paramètre $\theta$ inconnu. L'objectif est d'évaluer ce paramètre à l'aide d'un $n$-échantillon $X_1, \ldots, X_n$ de $X$.

**Estimateur :** On appelle estimateur $\widehat{\theta}_n$ de $\theta$ toute fonction de $X_1, \ldots, X_n$.

**Biais :** $B(\widehat{\theta}_n, \theta) = \mathbb{E}(\widehat{\theta}_n) - \theta$. $\widehat{\theta}_n$ est sans biais (de $\theta$) $\Leftrightarrow$ $B(\widehat{\theta}_n, \theta) = 0$ $\Leftrightarrow$ $\mathbb{E}(\widehat{\theta}_n) = \theta$.

**Risque quadratique :** $R(\widehat{\theta}_n, \theta) = \mathbb{E}((\widehat{\theta}_n - \theta)^2)$. Plus $R(\widehat{\theta}_n, \theta)$ est petit, plus l'estimateur $\widehat{\theta}_n$ estime bien $\theta$.

Si $\widehat{\theta}_n$ est sans biais, alors on a $R(\widehat{\theta}_n, \theta) = \mathbb{V}(\widehat{\theta}_n)$.

**Estimateur consistant :** $\widehat{\theta}_n$ est consistant $\Leftrightarrow$ $(\widehat{\theta}_n)_{n\in\mathbb{N}^*}$ converge en probabilité vers $\theta$.

**Méthode des moments :** La méthode des moments permet de construire un estimateur $\widehat{\theta}_n$ de $\theta$ :

1. On détermine une fonction $g : \mathbb{R} \to \mathbb{R}$ telle que $\mathbb{E}(X_1) = g(\theta)$.
2. On isole $\theta$ en fonction de $\mathbb{E}(X_1)$ : $\theta = g^{-1}(\mathbb{E}(X_1))$.
3. On estime $\mathbb{E}(X_1)$ par $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$.
4. L'estimateur $\widehat{\theta}_n$ de $\theta$ par la méthode des moments est : $\widehat{\theta}_n = g^{-1}(\overline{X}_n)$.

**Intervalle de confiance :** On appelle intervalle de confiance pour $\theta$ au niveau $100(1-\alpha)\%$, $\alpha \in ]0,1[$, un intervalle de la forme $I_\theta = [a(X_1, \ldots, X_n), b(X_1, \ldots, X_n)]$, où $a(X_1, \ldots, X_n)$ et $b(X_1, \ldots, X_n)$ désignent deux fonctions dépendantes de $X_1, \ldots, X_n$, tel que $\mathbb{P}(\theta \in I_\theta) = 1 - \alpha$. Pour que $I_\theta$ contienne avec une forte probabilité $\theta$, on prend $\alpha$ petit : $\alpha = 0,05$, $\alpha = 0,01$ ...

En pratique, on dispose de données $x_1, \ldots, x_n$, lesquelles sont des réalisations de $X_1, \ldots, X_n$. Une réalisation de $I_\theta$ est donc : $i_\theta = [a(x_1, \ldots, x_n), b(x_1, \ldots, x_n)]$, donnant ainsi un intervalle de valeurs dans lequel $\theta$ a $100(1-\alpha)\%$ de chances d'appartenir.

Les intervalles de confiance associés à une *var* $X$ suivant une loi normale sont donnés page 23.

**Hypothèses statistiques :** On appelle hypothèse statistique une hypothèse inhérente au contexte de l'étude.

À la base d'un test statistique, il y a deux hypothèses complémentaires, notées $H_0$ et $H_1$,

- ◦ l'hypothèse $H_0$ formule ce que l'on souhaite rejeter/réfuter,
- ◦ l'hypothèse $H_1$ formule ce que l'on souhaite montrer.

Par exemple, si on veut montrer l'hypothèse "le produit est non conforme", $H_0$ et $H_1$ s'opposent sous la forme :

$$H_0 : \text{"le produit est conforme"} \qquad \text{contre} \qquad H_1 : \text{"le produit est non conforme".}$$

**Test statistique :** On appelle test statistique une démarche/procédure qui vise à apporter une réponse à la question : est-ce que les données nous permettent de rejeter $H_0$/accepter $H_1$ avec un faible risque de se tromper ?

**p-valeur :** On appelle p-valeur le plus petit réel $\alpha \in ]0,1[$ calculé à partir des données tel que l'on puisse se permettre de rejeter $H_0$ au risque $100\alpha\%$.

**Degré de significativité :** La p-valeur nous donne un degré de significativité du rejet de $H_0$.

Le rejet de $H_0$ sera :

- ◦ significatif si p-valeur $\in ]0.01,\ 0.05]$, symbolisé par $*$,
- ◦ très significatif si p-valeur $\in ]0.001,\ 0.01]$, symbolisé par $**$,
- ◦ hautement significatif si p-valeur $< 0.001$, symbolisé par $***$.

Il y a non rejet de $H_0$ si p-valeur $> 0.05$.

S'il y a non-rejet de $H_0$, sauf convention, on ne peut rien conclure du tout (avec le risque considéré). En revanche, peut-être qu'un risque de départ plus élevé ou la disposition de plus de données peuvent conduire à un rejet de $H_0$.

Les tests statistiques associés à une *var* $X$ suivant une loi normale sont présentés page 24.

# 15 Intervalles de confiance

Lois : $Z \sim \mathcal{N}(0,1)$, $T \sim \mathcal{T}(\nu)$, $K \sim \chi^2(\nu)$, $\nu = n-1$. Niveau : $100(1-\alpha)\%$, $\alpha \in ]0,1[$.

| $X \sim \mathcal{N}(\mu, \sigma^2)$ | valeurs | $i_\mu$ |
|---|---|---|
| $\sigma$ connu<br>Z-IntConf | $\mathbb{P}(\|Z\| \geq z_\alpha) = \alpha$ | $\left[\overline{x} - z_\alpha \frac{\sigma}{\sqrt{n}},\ \overline{x} + z_\alpha \frac{\sigma}{\sqrt{n}}\right]$ |
| $\sigma$ inconnu<br>T-IntConf | $\mathbb{P}(\|T\| \geq t_\alpha(\nu)) = \alpha$ | $\left[\overline{x} - t_\alpha(\nu) \frac{s}{\sqrt{n}},\ \overline{x} + t_\alpha(\nu) \frac{s}{\sqrt{n}}\right]$ |
| $X \sim \mathcal{N}(\mu, \sigma^2)$ | valeurs | $i_{\sigma^2}$ |
| Chi-Square-IntConf | $\mathbb{P}(K \geq a_\alpha(\nu)) = 1 - \frac{\alpha}{2}$<br>$\mathbb{P}(K \geq b_\alpha(\nu)) = \frac{\alpha}{2}$ | $\left[\frac{n-1}{b_\alpha(\nu)} s^2,\ \frac{n-1}{a_\alpha(\nu)} s^2\right]$ |
| $n \geq 31$<br>(parfois utilisé) | $\mathbb{P}(\|Z\| \geq z_\alpha) = \alpha$ | $\left[\frac{2(n-1)s^2}{\left(\sqrt{2n-3} + z_\alpha\right)^2},\ \frac{2(n-1)s^2}{\left(\sqrt{2n-3} - z_\alpha\right)^2}\right]$ |
| $X \sim \mathcal{B}(p)$ | valeurs | $i_p$ |
| $n \geq 31$<br>$np_1 \geq 5$, $n(1-p_2) \geq 5$<br>Prop-Z-IntConf | $\mathbb{P}(\|Z\| \geq z_\alpha) = \alpha$ | $\left[\overline{x} - z_\alpha \frac{\sqrt{\overline{x}(1-\overline{x})}}{\sqrt{n-1}},\ \overline{x} + z_\alpha \frac{\sqrt{\overline{x}(1-\overline{x})}}{\sqrt{n-1}}\right] = [p_1, p_2]$ |
| $X$ de loi quelconque | valeurs | $i_{\mathbb{E}(X)}$ |
| $n \geq 1000$<br>Z-IntConf-Lim | $\mathbb{P}(\|Z\| \geq z_\alpha) = \alpha$ | $\left[\overline{x} - z_\alpha \frac{s}{\sqrt{n}},\ \overline{x} + z_\alpha \frac{s}{\sqrt{n}}\right]$ |

# 16 Tests de conformité

Lois : $Z \sim \mathcal{N}(0,1)$, $T \sim \mathcal{T}(\nu)$, $K \sim \chi^2(\nu)$, $\nu = n-1$. Rejet de $H_0$ au risque $100\alpha\%$ $\Leftrightarrow$ p-valeur $\leq \alpha$.

| $X \sim \mathcal{N}(\mu, \sigma^2)$ | $H_0$ | $H_1$ | Stat. test obs. | p-valeurs |
|---|---|---|---|---|
| $\sigma$ connu :<br>Z-Test | $\mu = \mu_0$<br>$\mu \leq \mu_0$<br>$\mu \geq \mu_0$ | $\mu \neq \mu_0$<br>$\mu > \mu_0$<br>$\mu < \mu_0$ | $z_{obs} = \sqrt{n}\left(\frac{\overline{x}-\mu_0}{\sigma}\right)$ | $\mathbb{P}(\lvert Z\rvert \geq \lvert z_{obs}\rvert)$<br>$\mathbb{P}(Z \geq z_{obs})$ $\left(= \frac{1}{2}\mathbb{P}(\lvert Z\rvert \geq z_{obs}) \text{ si } z_{obs} > 0\right)$<br>$\mathbb{P}(Z \leq z_{obs})$ $\left(= \frac{1}{2}\mathbb{P}(\lvert Z\rvert \geq -z_{obs}) \text{ si } z_{obs} < 0\right)$ |
| $\sigma$ inconnu :<br>T-Test | $\mu = \mu_0$<br>$\mu \leq \mu_0$<br>$\mu \geq \mu_0$ | $\mu \neq \mu_0$<br>$\mu > \mu_0$<br>$\mu < \mu_0$ | $t_{obs} = \sqrt{n}\left(\frac{\overline{x}-\mu_0}{s}\right)$ | $\mathbb{P}(\lvert T\rvert \geq \lvert t_{obs}\rvert)$<br>$\mathbb{P}(T \geq t_{obs})$ $\left(= \frac{1}{2}\mathbb{P}(\lvert T\rvert \geq t_{obs}) \text{ si } t_{obs} > 0\right)$<br>$\mathbb{P}(T \leq t_{obs})$ $\left(= \frac{1}{2}\mathbb{P}(\lvert T\rvert \geq -t_{obs}) \text{ si } t_{obs} < 0\right)$ |
| 1-Chi2-Test | $\sigma^2 = \sigma_0^2$<br>$\sigma^2 \leq \sigma_0^2$<br>$\sigma^2 \geq \sigma_0^2$ | $\sigma^2 \neq \sigma_0^2$<br>$\sigma^2 > \sigma_0^2$<br>$\sigma^2 < \sigma_0^2$ | $\chi^2_{obs} = \frac{n-1}{\sigma_0^2}s^2$ | $2\min\left(\mathbb{P}(K \geq \chi^2_{obs}),\ \mathbb{P}(K \leq \chi^2_{obs})\right)$<br>$\mathbb{P}(K \geq \chi^2_{obs})$<br>$\mathbb{P}(K \leq \chi^2_{obs})$ |
| $n \geq 31$<br>(parfois utilisé) | $\sigma^2 = \sigma_0^2$<br>$\sigma^2 \leq \sigma_0^2$<br>$\sigma^2 \geq \sigma_0^2$ | $\sigma^2 \neq \sigma_0^2$<br>$\sigma^2 > \sigma_0^2$<br>$\sigma^2 < \sigma_0^2$ | $z_{obs} = \sqrt{\frac{2(n-1)}{\sigma_0^2}s^2} - \sqrt{2n-3}$ | $\mathbb{P}(\lvert Z\rvert \geq \lvert z_{obs}\rvert)$<br>$\mathbb{P}(Z \geq z_{obs})$ $\left(= \frac{1}{2}\mathbb{P}(\lvert Z\rvert \geq z_{obs}) \text{ si } z_{obs} > 0\right)$<br>$\mathbb{P}(Z \leq z_{obs})$ $\left(= \frac{1}{2}\mathbb{P}(\lvert Z\rvert \geq -z_{obs}) \text{ si } z_{obs} < 0\right)$ |
| $X \sim \mathcal{B}(p)$ | $H_0$ | $H_1$ | stat. test obs. et *var* | p-valeurs |
| $n \geq 31$,<br>$np_0 \geq 5$,<br>$n(1-p_0) \geq 5$ :<br>1-Prop-Z-Test | $p = p_0$<br>$p \leq p_0$<br>$p \geq p_0$ | $p \neq p_0$<br>$p > p_0$<br>$p < p_0$ | $z_{obs} = \sqrt{n}\left(\frac{\overline{x}-p_0}{\sqrt{p_0(1-p_0)}}\right)$ | $\mathbb{P}(\lvert Z\rvert \geq \lvert z_{obs}\rvert)$<br>$\mathbb{P}(Z \geq z_{obs})$ $\left(= \frac{1}{2}\mathbb{P}(\lvert Z\rvert \geq z_{obs}) \text{ si } z_{obs} > 0\right)$<br>$\mathbb{P}(Z \leq z_{obs})$ $\left(= \frac{1}{2}\mathbb{P}(\lvert Z\rvert \geq -z_{obs}) \text{ si } z_{obs} < 0\right)$ |

# Index