L3 Economie

Statistiques inférentielles

Chapitre 1 : Outils de probabilité pour la statistique

Cécile Durot
cecile.durot@gmail.com

Université Paris Nanterre

Définitions

- Une expérience est dite aléatoire si elle peut produire des résultats différents lorsqu'on la renouvelle dans les mêmes conditions.
- Une expérience aléatoire définit un ensemble Ω de résultats possibles appelé univers.
- On appelle tribu sur Ω tout ensemble $\mathcal{A}$ de parties de Ω satisfaisant les trois conditions :
  - $\emptyset \in \mathcal{A}$,
  - Si $A \in \mathcal{A}$, alors $\overline{A} \in \mathcal{A}$ où $\overline{A}$ désigne le complémentaire de A dans Ω,
  - Si $A_n \in \mathcal{A}$ pour tout n, alors $\cup_n A_n \in \mathcal{A}$ (union finie ou dénombrable).
- Tout élément $A \in \mathcal{A}$ est appelé événement. Par exemple, l'ensemble vide $\emptyset$ et l'univers Ω sont toujours des événements.

Remarque : Si $A_n \in \mathcal{A}$ pour tout n, alors on a $\cap_n A_n \in \mathcal{A}$: toute union ou intersection dénombrable d'événements est un événement.

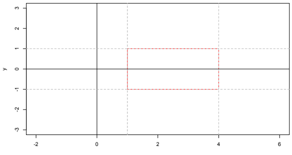
Exemples de tribus

1. Si Ω est fini, alors $\mathcal{A}$ est l'ensemble de toutes les parties de Ω. Par exemple pour un lancer de pièce de monnaie, on a $\Omega = \{p, f\}$ et
$$\mathcal{A} = \{\emptyset, \{p\}, \{f\}, \Omega\}.$$
Pour deux lancers successifs, on a $\Omega = \{pp, pf, ff, fp\}$ et $\mathcal{A}$ comporte $\sum_{k=0}^{4} C_4^k = 16$ éléments :
$$\mathcal{A} = \Big\{\emptyset, \{pp\}, \{pf\}, \{fp\}, \{ff\}, \{pp, pf\}, \{pp, ff\},$$
$$\{pp, fp\}, \{pf, ff\}, \{pf, fp\}, \{ff, fp\}, \{pp, pf, ff\},$$
$$\{pp, pf, fp\}, \{pp, ff, fp\}, \{pf, ff, fp\}, \Omega\Big\}.$$
Ainsi, l'événement "on obtient une fois pile et une fois face" est $\{pf, fp\}$, et l'événement "on obtient pile au premier lancer" est $\{pp, pf\}$.

Exemples de tribus (2)

2. Si Ω est $\mathbb{R}$, $\mathbb{R}^n$, un intervalle de $\mathbb{R}$ ou un pavé de $\mathbb{R}^n$, alors l'ensemble de toutes les parties de Ω possède beaucoup trop d'éléments pour fournir une axiomatique cohérente. Dans le cas $\Omega = \mathbb{R}$, on considère la tribu engendrée par les intervalles (i.e. tous les intervalles sont des événements, de même que toutes les unions ou intersections finites ou dénombrables d'intervalles). Dans le cas $\Omega = \mathbb{R}^n$, on considère la tribu engendrée par les pavés.

Exemple de pavé dans $\mathbb{R}^2$: l'ensemble des (x, y) tels que $x \in [1, 4]$ et $y \in [-1, 1]$ est l'ensemble des points dans le rectangle rouge



Définitions

- Un espace de probablité est un triplet $(\Omega, \mathcal{A}, P)$ où Ω est un univers, $\mathcal{A}$ est une tribu sur Ω et $P : \mathcal{A} \to [0, 1]$ est une fonction satisfaisant les trois conditions :
  - $P(\emptyset) = 0$,
  - $P(\Omega) = 1$,
  - Si $A_n \in \mathcal{A}$ pour tout n, avec $A_i \cap A_j = \emptyset$ pour tous $i \neq j$ (événements deux à deux disjoints), alors pour une union finie ou dénombrable on a
$$P(\cup_n A_n) = \sum_n P(A_n).$$
- La fonction P est appelée mesure de probablité.
- Pour tout événement $A \in \mathcal{A}$, $P(A)$ est appelé probabilité de A.

Remarque : Il découle de la loi des grands nombres (voir plus loin dans ce cours) que $P(A)$ correspond à la fréquence d'observation de l'évenement A lorsque l'on renouvelle l'expérience un grand nombre de fois.

Quelques propriétés

Ci-dessous, A, B, $A_1, A_2, \ldots$ désignent des événements. On a :

- $P(\overline{A}) = 1 - P(A)$,
- Si $A_1 \cup \cdots \cup A_n = \Omega$ et $A_i \cap A_j = \emptyset$ pour tous $i \neq j$ (i.e. $A_1, \ldots, A_n$ forment une partition de Ω) alors $P(A_1) + \cdots + P(A_n) = 1$,
- Si $A \subset B$, alors $P(A) \leqslant P(B)$,
- $P(A \cup B) \leqslant P(A) + P(B)$.

Définition

Deux événements A et B sont dits indépendants entre eux si $P(A \cap B) = P(A)P(B)$.

Définition

Soit $(\Omega, \mathcal{A}, P)$ un espace de probabilité. Soit $\mathcal{B}$ la tribu engendrée par les intervalles dans $\mathbb{R}$. Une variable aléatoire réelle (en abrégé, v.a.r.) est une fonction

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

satisfaisant la propriété suivante :

$$\text{pour tout } B \in \mathcal{B}, \text{ on a } X^{-1}(B) \in \mathcal{A},$$

c'est-à-dire que $X^{-1}(B)$ est un événement.

Notations

Ici, $X^{-1}(B)$ est l'image réciproque de B par X, c'est-à-dire que

$$X^{-1}(B) = \{\omega \in \Omega \text{ tels que } X(\omega) \in B\}.$$

Pour alléger les notations, on note également $\{X \in B\}$ cet événement. Puisque $\{X \in B\} \in \mathcal{A}$, la probabilité $P(\{X \in B\})$ est bien définie ; on la note simplement $P(X \in B)$.

Remarque

De cette définition abstraite, on retiendra surtout qu'une v.a.r. est une grandeur dépendant du résultat d'une expérience aléatoire, c'est-à-dire non prévisible à l'avance avec certitude. On peut définir plusieurs v.a.r. attachées à une même expérience aléatoire.

Définition

Pour une expérience réalisée, une v.a.r. X prend une valeur donnée $X(\omega)$ appelée réalisation de X. On note traditionnellement x une réalisation de X.

Remarque

Si l'on renouvelle l'expérience aléatoire, alors on obtient différentes réalisations de la même v.a.r. X.

Exemples

- Considérant une population $\mathcal{P}$ (i.e. un ensemble fini d'individus), une expérience aléatoire consiste à choisir un individu de $\mathcal{P}$ au hasard (i.e. de sorte que chaque individu de $\mathcal{P}$ ait les mêmes chances que les autres d'être choisi). L'univers est $\Omega = \mathcal{P}$ et on peut considérer par exemple

$$\begin{array}{rcl} X : \Omega & \to & \mathbb{R} \\ \omega & \to & \text{revenu net mensuel de l'individu } \omega \end{array}$$

- Une autre expérience consiste à choisir n individus de $\mathcal{P}$ au hasard, sans remise. Dans ce cas, Ω est l'ensemble des parties de $\mathcal{P}$ à n éléments et on peut définir par exemple $X(\omega)$ comme le salaire net mensuel moyen des n individus dans ω.
- Une autre encore consiste à choisir n individus de $\mathcal{P}$ au hasard, avec remise. Dans ce cas, $\Omega = \mathcal{P}^n$ et on peut définir X comme ci-dessus.

Définitions et notations

Soit X une v.a.r. et $\mathcal{B}$ la tribu engendrée par les intervalles dans $\mathbb{R}$.

1. La loi de probablité, en abrégé "loi", de X est la donnée de $P(X \in B)$ pour tout $B \in \mathcal{B}$. On note $P_X(B) = P(X \in B)$.
2. La fonction de répartition de X est la fonction notée F_X définie par

$$\begin{array}{rcl} F_X : \mathbb{R} & \to & [0,1] \\ t & \to & P(X \in]-\infty, t]), \end{array}$$

ce qu'on note aussi $F_X(t) = P(X \leqslant t)$ pour tout $t \in \mathbb{R}$.

Propriétés

- La loi de X est entièrement caractérisée par F_X,
- F_X est une fonction croissante (au sens large),
- $\lim_{t \to -\infty} F_X(t) = 0$ et $\lim_{t \to +\infty} F_X(t) = 1$,
- Pour tous réels $a < b$ on a $P(X \in]a, b]) = F_X(b) - F_X(a)$,
- P_X est une mesure de probablité.

V.a.r. discrète et v.a.r. continue

Soit X une v.a.r. et $\Omega_X = \{X(\omega),\ \omega \in \Omega\}$ l'image de Ω par X, i.e. l'ensemble des valeurs possibles pour X.

1. X est dite discrète si Ω_X est un ensemble fini ou dénombrable. Par exemple :
    - Face obtenue lors du lancer d'un dé, $\Omega_X = \{1, 2, 3, 4, 5, 6\}$.
    - Nombre de bugs dans un programme informatique, $\Omega_X = \mathbb{N}$.
2. Dans certains cas, Ω_X n'est pas dénombrable. Par exemple :
    - Appel de la fonction random d'une calculatrice, $\Omega_X = [0, 1]$.
    - Durée de bon fonctionnement d'un système, $\Omega_X = [0, +\infty)$.

   Si Ω_X n'est pas dénombrable, nous ne considèrerons dans ce cours que le cas de v.a.r. dont la fonction de répartition est continue. Nous dirons alors que X est continue.

Propriétés d'une v.a.r. discrète

Soit X une v.a.r. discrète et $\Omega_X = \{X(\omega),\ \omega \in \Omega\}$.

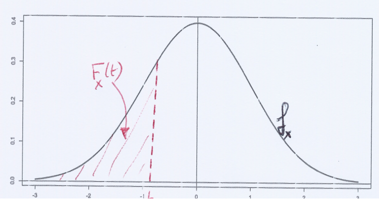
1. La loi de X est entièrement déterminée par les probabilités
$$P(X = x),\ x \in \Omega_X.$$
2. $\displaystyle\sum_{x\in\Omega_X} P(X = x) = 1.$
3. Pour tout $t \in \mathbb{R}$, on a $F_X(t) = \displaystyle\sum_{x\in\Omega_X,\ x\leqslant t} P(X = x)$. La fonction F_X est donc une fonction en escalier.

Densité d'une v.a.r. continue

Définition : Soit X une v.a.r. Si il existe une fonction $f_X : \mathbb{R} \to [0, +\infty[$ continue sauf éventuellement en un nombre fini de points et satisfaisant

$$F_X(t) = \int_{-\infty}^{t} f_X(x)dx \text{ pour tout } t \in \mathbb{R},$$

alors f_X est appelée densité de X.



Densité d'une v.a.r. continue

Propriétés : Soit X une v.a.r. possédant une densité f_X.

1. La densité ne dépend que de la loi de probablité.
2. $f_X(x) = F'_X(x)$ en tout point de continuité de f_X.
3. $P(X \in]a, b]) = \displaystyle\int_a^b f_X(x)dx$ pour tous $a < b$
4. $\int_{-\infty}^{+\infty} f_X(x)dx = 1$.
5. $P(X = x) = 0$ pour tout $x \in \mathbb{R}$.
6. F_X est continue.
7. Pour tout $t \in \mathbb{R}$, $F_X(t)$ est l'aire sous la courbe de f_X à gauche de t.
8. Si f_X est une fonction paire, alors $F_X(-t) = 1 - F_X(t)$ pour tout $t \in \mathbb{R}$.

Espérance

Définition : Sous réserve d'être bien définie, l'espérance de X est

- $E(X) = \sum_{x \in \Omega_X} x P(X = x)$ si X est discrète,
- $E(X) = \int_{-\infty}^{+\infty} x f_X(x) dx$ si X possède une densité f_X.

Remarque : Il découle de la loi des grands nombres (voir plus loin dans ce cours) que $E(X)$ correspond à la valeur moyenne observée pour X lorsque l'on renouvelle l'expérience un grand nombre de fois.

Formule de transfert : Soit $g : \Omega_X \to \mathbb{R}$ continue. Alors

- $E(g(X)) = \sum_{x \in \Omega_X} g(x) P(X = x)$ si X est discrète,
- $E(g(X)) = \int_{-\infty}^{+\infty} g(x) f_X(x) dx$ si X possède une densité f_X,

sous réserve que cette quantité soit bien définie.

Variance

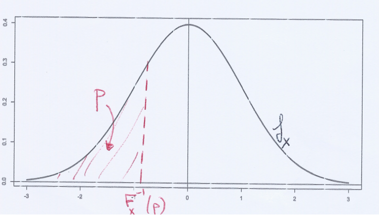
Définition : Sous réserve d'être bien définie, la variance de X est $var(X) = E((X - E(X))^2)$.

Remarque : La variance est un indicateur de la dispersion de X : plus la variance de X est petite, plus les réalisations de X sont concentrées autour de l'espérance $E(X)$. Plutôt que la variance, on considère souvent l'écart-type $\sqrt{var(X)}$ comme indicateur de dispersion.

Quantile

Définition : Soit F_X continue et strictement croissante.

1. La fonction réciproque F_X^{-1} est appelée fonction quantile.
2. Pour tout $p \in]0,1[$, $F_X^{-1}(p)$ est appelé le quantile d'ordre p de X. Il s'agit de la solution q de l'équation $F_X(q) = p$. Si X possède une densité f_X, il s'agit donc du point q tel que l'aire sous la courbe de f_X à gauche de q soit égale à p.



Remarque : Il existe une définition plus générale couvrant entre autres le cas discret.

Vecteurs aléatoires

Définition : Soient $X_1, \ldots, X_n$ des de v.a.r. définies sur un même univers Ω. Alors la fonction $(X_1, \ldots, X_n) : \Omega \to \mathbb{R}^n$ est appelée vecteur aléatoire. La réalisation d'un tel vecteur prend la forme

$$(X_1(\omega), \ldots, X_n(\omega)).$$

Exemple : On choisit successivement avec remise n individus dans une population $\mathcal{P}$. Alors $\Omega = \mathcal{P}^n$ et pour tout $\omega \in \Omega$, on définit $X_i(\omega)$ le revenu mensuel net du i-ème individu dans ω.

Suites : La notion de vecteur aléatoire peut être étendue à la notion de suite aléatoire $(X_i)_{i \in \mathbb{N}}$.

Indépendance

Définition : Dans un vecteur aléatoire $(X_1, \dots, X_n)$, les variables aléatoires $X_1, \dots, X_n$ sont dite (mutuellement) indépendantes si pour tous éléments $B_1, \dots, B_n$ de la tribu engendrée par les intervalles dans $\mathbb{R}$,

$$P((X_1, \dots, X_n) \in B_1 \times \cdots \times B_n) = P(X_1 \in B_1) \times \cdots \times P(X_n \in B_n).$$

Remarque : Concrètement, l'indépendance signifie que la valeur prise par certaines variables n'a aucune influence sur la valeur prise par d'autres. C'est le cas lors de tirages au hasard successifs avec remise si pour tout i, la v.a.r. X_i ne porte que sur le résultat du i-ème tirage. Ce n'est pas le cas pour un tirage avec remise.

Calculs d'espérance et de variance

Propriétés : Considérons des nombres réels $a_0, \dots, a_n$ et un vecteur aléatoire $(X_1, \dots, X_n)$.

1. $a_0 + a_1 X_1 + \dots + a_n X_n$ est une v.a.r.
2. Si $X_1, \dots, X_n$ possèdent une espérance, alors (linéarité)
$$E(a_0 + a_1 X_1 + \dots + a_n X_n) = a_0 + a_1 E(X_1) + \dots + a_n E(X_n).$$
3. Si $X_1, \dots, X_n$ possèdent une variance et sont indépendantes,
$$var(a_0 + a_1 X_1 + \dots + a_n X_n) = a_1^2 var(X_1) + \dots + a_n^2 var(X_n).$$

Exemples (cf exercices 9 et 10) : • Soit $X_1, \dots, X_n$ des v.a.r. indépendantes possédant une espérance m et une variance σ^2. Soit

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$$

Alors, $E(\bar{X}_n) = m$ et $var(\bar{X}_n) = \sigma^2 / n$.

• Si X possède une variance, alors $var(X) = E(X^2) - (E(X))^2$.

Décomposition biais/variance

Soit T_n une variable aléatoire possédant une variance. On a alors pour tout $a \in \mathbb{R}$:

$$E((T_n - a)^2) = var(T_n) + (E(T_n) - a)^2.$$

En effet, on sait que pour toute variable X possédant une variance, on a $var(X) = E(X^2) - (E(X))^2$. En particulier, avec $X = T_n - a$ on obtient

$$\begin{aligned} var(T_n) &= var(T_n - a) \\ &= E((T_n - a)^2) - (E(T_n) - a)^2. \end{aligned}$$

Lois discrètes

Bernoulli : X suit une loi de Bernoulli si X ne peut prendre que deux valeurs, codées par 0 (échec) et 1 (succès). Cette loi est entièrement déterminée par $p = P(X = 1)$, on la note $\mathcal{B}(p)$.

Binomiale : La loi $\mathcal{B}(n, p)$, avec $n \in \mathbb{N}^\star$ et $p \in [0, 1]$, est la loi de $X_1 + \cdots + X_n$ où $X_1, \ldots, X_n$ sont indépendantes et de même loi $\mathcal{B}(p)$.

Poisson : Pour $\lambda > 0$, X suit une loi de Poisson $\mathcal{P}(\lambda)$ si $\Omega_X = \mathbb{N}$ et

$$P(X = x) = \frac{\lambda^x}{x!} \exp(-\lambda) \text{ pour tout } x \in \mathbb{N}.$$

Cette loi fournit une approximation de $\mathcal{B}(n, p)$ lorsque $\lambda = np$ et n est grand ($n \geqslant 50$ et $p \leqslant 0,1$). Elle modélise le nombre d'occurences d'événements rares.

Lois continues

Exponentielle : Pour $\lambda > 0$, X suit une loi de Exponentielle $\mathcal{E}(\lambda)$ si X possède la densité définie par

$$f_X(x) = \begin{cases} \lambda \exp(-\lambda x) & \text{si } x \geqslant 0 \\ 0 & \text{si } x < 0. \end{cases}$$

Cette loi est utilisée pour modéliser la durée de bon fonctionnement d'un système sans usure.

Gaussienne (aussi appelée Normale) : Une v.a.r. X suit une loi $\mathcal{N}(m, \sigma^2)$, avec $m \in \mathbb{R}$ et $\sigma > 0$ si X possède la densité définie par

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(x-m)^2\right).$$

La loi $\mathcal{N}(0, 1)$ est appelée Gaussienne standard. Cette loi est utilisée pour modéliser des sommes ou moyennes de quantités indépendantes.

Chi-deux : La loi $\chi^2(n)$, avec $n \in \mathbb{N}^\star$, est la loi de $X_1^2 + \cdots + X_n^2$ où $X_1, \ldots, X_n$ sont indépendantes et de même loi Gaussienne standard.

Propriétés de la loi Gaussienne :

1. Si $a_0, \ldots, a_n$ sont des réels et $X_1, \ldots, X_n$ sont des v.a.r. Gaussiennes indépendantes, alors $a_0 + a_1 X_1 + \cdots + a_n X_n$ suit une loi Gaussienne.
2. En particulier, si $X \sim \mathcal{N}(m, \sigma^2)$, alors $\dfrac{X - m}{\sigma} \sim \mathcal{N}(0, 1)$.
3. La fonction de répartition de la loi $\mathcal{N}(m, \sigma^2)$ ne se calcule pas analytiquement. On a recours à un logiciel ou à une table de valeurs numériques pour déterminer le valeur de cette fonction en un point donné. C'est aussi le cas de la loi de Chi-deux.

Espérance et variance des lois usuelles

Loi	Espérance	Variance
$\mathcal{B}(p)$	p	$p(1-p)$
$\mathcal{B}(n,p)$	np	$np(1-p)$
$\mathcal{P}(\lambda)$	λ	λ
$\mathcal{E}(\lambda)$	$1/\lambda$	$1/\lambda^2$
$\mathcal{N}(m,\sigma^2)$	m	σ^2
$\chi^2(n)$	n	$2n$

Définitions : Soit $(T_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires.

1. Soit $a \in \mathbb{R}$. On dit que T_n converge en probabilité vers a, ce que l'on note
$$T_n \xrightarrow[n\to\infty]{P} a,$$
si
$$\forall \epsilon > 0, \ \lim_{n\to\infty} P(|T_n - a| > \epsilon) = 0,$$
ou de façon équivalente, si
$$\forall \epsilon > 0, \ \lim_{n\to\infty} P(T_n \in [a-\epsilon, a+\epsilon]) = 1.$$

2. Soit L une loi de probablité possédant une densité et T une v.a.r. de loi L. On dit que T_n converge en loi vers L, ce que l'on note
$$T_n \xrightarrow[n\to\infty]{\mathcal{L}} L,$$
si
$$\forall t \in \mathbb{R}, \ \lim_{n\to\infty} P(T_n \leqslant t) = P(T \leqslant t).$$

Propriétés :

1. Si $T_n \xrightarrow[n\to\infty]{\mathcal{L}} L$, alors on a aussi
$$\lim_{n\to\infty} P(T_n \in I) = P(T \in I)$$
pour tout intervalle $I \subset \mathbb{R}$.
2. Pour une loi limite L continue, on peut montrer (théorème de Dini) que la convergence des fonctions de répartition est uniforme :
$$\lim_{n\to\infty} P(T_n \leqslant t_n) - P(T \leqslant t_n) = 0$$
pour tout $t_n \in \mathbb{R}$ dépendant éventuellement de n.
3. Dans ce cours, nous ne considèrerons que des situations de convergence en loi pour lesquelles la loi limite L est soit une loi Gaussienne soit une loi de Chi-deux.
4. Il existe une définition plus générale couvrant, entre autres, le cas d'une loi limite discrète.

Loi des grands nombres et Théorème de la limite centrale

Soit $X_1, X_2, \ldots$ une suite de v.a.r. indépendantes et de même loi. Pour tout $n \in \mathbb{N}$, on note

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i.$$

Théorème (LGN) : Si $X_1, X_2, \ldots$ admettent une espérance m, alors

$$\bar{X}_n \xrightarrow[n\to\infty]{P} m.$$

Théorème (TLC) : Si $X_1, X_2, \ldots$ admettent une espérance m et une variance σ^2, alors

$$\sqrt{n}\frac{\bar{X}_n - m}{\sigma} \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}(0,1).$$

Remarques

1. Dans le TLC, le terme de gauche et la loi limite sont tous deux centrés (espérance $=0$) et réduits (variance$=1$).
2. En particulier, si $T_n \sim \mathcal{B}(n,p)$ pour tout $n \in \mathbb{N}$, où $p \in]0,1[$, alors (voir exercice 13)
$$\frac{T_n - np}{\sqrt{np(1-p)}} \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}(0,1).$$
On utilise cette approximation lorsque $n \geqslant 30$, $np \geqslant 15$ et $np(1-p) > 5$.
3. Si $X_1, X_2, \ldots$ sont indépendantes et de même loi gaussienne $\mathcal{N}(m,\sigma^2)$ alors pour tout n, on a la loi exacte (voir exercice 14)
$$\sqrt{n}\frac{\bar{X}_n - m}{\sigma} \sim \mathcal{N}(0,1).$$

Théorème

Théorème : Soit (T_n) une suite de variables aléatoires et a un réel. Si $\lim_{n\to\infty} E((T_n - a)^2) = 0$ alors

$$T_n \xrightarrow[n\to\infty]{P} a.$$

En particulier, si $\lim_{n\to\infty} E(T_n) = a$ et $\lim_{n\to\infty} var(T_n) = 0$, il découle de la décomposition biais/variance que T_n converge en probabilité vers a quand $n \to \infty$.

Théorème de Slutsky

Soit $T_1, T_2, \dots$ une suite de v.a.r.

1. Soit $a \in \mathbb{R}$ et g une fonction continue sur Ω_{T_n} pour tout n. On a l'implication suivante :
$$T_n \xrightarrow[n\to\infty]{P} a \quad \Rightarrow \quad g(T_n) \xrightarrow[n\to\infty]{P} g(a).$$
2. Soit $a \in \mathbb{R}$ et $a_1, a_2, \dots$ une suite réelle convergeant vers 1. On a l'implication suivante :
$$T_n \xrightarrow[n\to\infty]{P} a \quad \Rightarrow \quad a_n T_n \xrightarrow[n\to\infty]{P} a.$$
3. Soit $S_1, S_2, \dots$ une suite de v.a.r. convergeant en probabilité vers $\sigma > 0$. On a l'implication suivante :
$$\frac{T_n}{\sigma} \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}(0,1) \quad \Rightarrow \quad \frac{T_n}{S_n} \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}(0,1).$$