Statistiques inférentielles

-

CY-Tech ING2-GI

-

Année universitaire 2021-2022

-

A. BOURHATTAS

Chapitre 0

Rappels sur les probabilités et les Convergences de v.a.

0.1 Espaces probabilisés

0.1.1 Cas général

En probabilités, on travaille dans un espace probabilisé $(\Omega, \mathcal{A}, P)$. Pour celà, on commence par :

1. définir l'ensemble de tous les évènements élémentaires, appelé *espace fondamental* ou *univers*. Il est généralement noté Ω.
2. définir, parmi les parties de Ω, l'ensemble $\mathcal{A}$ des évènements dont on pourra calculer la probabilité. Il s'agit d'une tribu ou σ-algèbre. Autrement dit elle contient $\emptyset, \Omega$, est stable par complémentaire et par union dénombrable.
3. définir une mesure de probabilités, càd une application
   $P : \mathcal{A} \longrightarrow [0,1]$ avec $P(\Omega) = 1$ et $P(\cup_{i=1}^{\infty} A_i) = \sum\limits_{i=1}^{\infty} P(A_i)$ pour toute famille $(A_i)_{i \in \mathbb{N}}$ d'évènements disjoints 2 à 2.

0.1.2 Cas discret

Lorsque l'espace fondamental Ω est discret, càd fini ou dénombrable, on prend :

1. $\mathcal{A} = \mathcal{P}(\Omega)$, toutes les parties de Ω sont des évènements.
2. La mesure de probabilité est définie par la donnée des $P(\omega), \omega \in \Omega$ avec $\sum\limits_{\omega \in \Omega} P(\omega) = 1$.

0.1.3 Cas continu

Lorsque l'espace fondamental Ω est continu, càd $\Omega = \mathbb{R}$ ou un intervalle de $\mathbb{R}$, muni de la mesure de Lebesgue (μ), on prend :

1. $\mathcal{A} = \mathcal{B}_{\mathbb{R}}$, ensemble des boréliens = plus petite tribu engendrée par les intervalles.
2. La probabilité est définie par la donnée d'une fonction de densité $f : \Omega \longrightarrow \mathbb{R}^+$ positive, μ-mesurable et qui vérifie $\int_\Omega f(x) d\mu = 1$. Ainsi la probabilité d'un évènement $A \in \mathcal{A}$ est obtenue par : $P(A) = \int_A f d\mu$.

0.2 Variable aléatoire

0.2.1 Définition

Etant donné un espace probabilisé $(\Omega, \mathcal{A}, P)$, on appelle variable aléatoire, toute application : $X : \Omega \longrightarrow \mathbb{R}$ qui vérifie : $\forall B \in \mathcal{B}_{\mathbb{R}}, X^{-1}(B) \in \mathcal{A}$.

Une variable aléatoire permet de "transporter" la structure probabiliste définie sur Ω vers $\mathbb{R}$. On obtient ainsi une mesure de probabilité P_X sur $\mathbb{R}$ définie par : $\forall B \in \mathcal{B}_{\mathbb{R}},\ P_X(B) = P(X^{-1}(B))$.
Une v.a. est discrète ou continue selon que $X(\Omega)$ est discret ou continu.

0.2.2 Loi de probabilités, fonction de répartition

1. Lorsque la v.a. est discrète, sa loi de probabilité est définie à l'aide la fonction masse $p_X(x_k) = P(X = x_k) = P(\{\omega \in \Omega, X(\omega) = x_k\}$ pour les valeurs x_k du support de X.
2. Dans le cas continu, cette loi est définie grâce à une densité de probabilité $f_X : \mathbb{R} \longrightarrow \mathbb{R}$, fonction à valeurs positives, continue presque partout et vérifiant : $\displaystyle\int_{-\infty}^{+\infty} f(x)dx = 1$.
3. La fonction de répartition est par définition : $F_X : \mathbb{R} \longrightarrow [0,1]$ avec :
   $\forall x \in \mathbb{R}, F_X(x) = P(X \leq x) = P_X(]-\infty, x])$.
   Dans le cas discret : $F_X(x) = \displaystyle\sum_{k,\ x_k \leq x} P(X = x_k) = \sum_{k,\ x_k \leq x} p_X(x_k)$.
   Dans le cas continu : $F_X(x) = \displaystyle\int_{-\infty}^{x} f_X(t)dt$.
   La fonction de répartition est toujours *croissante*, partout *continue à droite* avec une limite égale à 0 en $-\infty$ et égale à 1 en $+\infty$.

0.2.3 Espérance, variance

1. Pour X v.a. discrète, l'espérance ou moyenne de X est :
   $E(X) = \mu = \sum_k x_k P(X = x_k) = \sum_k x_k p_X(x_k)$.
   La variance (carré de l'écart-type) est donnée par :
   $V(X) = E((X - E(X))^2) = E((X - \mu)^2)$, et donc :
   $V(X) = \sigma^2 = \sum_k (x_k - \mu)^2 P(X = x_k) = \left[\sum_k x_k^2 P(X = x_k)\right] - \mu^2$.
2. Pour X v.a. discrète, l'espérance ou moyenne de X est :
   $E(X) = \mu = \displaystyle\int_{-\infty}^{+\infty} xf(x)dx$. La variance est :
   $V(X) = \sigma^2 = \displaystyle\int_{-\infty}^{+\infty} (x - \mu)^2 f(x)dx = \int_{-\infty}^{+\infty} x^2 f(x)dx - \mu^2$.
3. L'espérance est linéaire, c à d : $E(\lambda X + \mu Y) = \lambda E(X) + \mu E(Y)$.
4. Quant varoiance, on a : $V(\lambda X) = \lambda^2 V(X)$,
   et si X et Y sont indépendantes, $V(X + Y) = V(X) + V(Y)$.

0.3 Les principales lois de probabilité usuelles

0.3.1 Lois discrètes usuelles

Nom	Notation	Loi de probabilité	Espérance	Variance
Uniforme		Valeurs prises $x_k, 1 \leq k \leq n$ $P(X = x_k) = \frac{1}{n}$	E(X)= $\frac{\sum\limits_{k=1}^{n} x_k}{n}$	$V(X) = \frac{\sum\limits_{k=1}^{n} x_k^2}{n} - E(X)^2$
Bernouilli	$\mathcal{B}(p)$	$X = 1$ (succès) ou $X = 0$ $P(X = 1) = p$ $P(X = 0) = 1 - p = q$	$E(X) = p$	$V(X) = p(1-p) = pq$
Binomiale	$\mathcal{B}(n,p)$	pour $0 \leq k \leq n$ $P(X = k) = C_n^k p^k q^{n-k}$	$E(X) = np$	$V(X) = np(1-p) = npq$
Poisson	$\mathcal{P}(\lambda)$	pour tout $k \in \mathbb{N}$ $P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}$	$E(X) = \lambda$	$V(X) = \lambda$

0.3.2 Loi continues usuelles

Nom	Notation	Loi de probabilité (densité)	Espérance	Variance
Uniforme	$\mathcal{U}\,[a,b]$	$f_X(x) = \frac{1}{b-a}$ sur $[a,b]$ $f_X(x) = 0$ ailleurs	$E(X) = \frac{a+b}{2}$	$V(X) = \frac{(b-a)^2}{12}$
Exponentielle	$\mathcal{E}xp(\theta)$	$f_X(x) = \theta e^{-\theta x}$ sur $[0, +\infty]$ $f_X(x) = 0$ ailleurs	$E(X) = \frac{1}{\theta}$	$V(X) = \frac{1}{\theta^2}$
Normale Centrée Réduite	$\mathcal{N}(0,1)$	$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$	$E(X) = 0$	$V(X) = 1$
Normale	$\mathcal{N}(\mu, \sigma^2)$	$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$	$E(X) = \mu$	$V(X) = \sigma^2$
Khi deux	$\chi^2(n)$ n dégrés de liberté	$\chi^2(n) = \sum\limits_{i=1}^{n} X_i^2$ $X_i \sim \mathcal{N}(0,1)$	$E(X) = n$	$V(X) = 2n$
Student	$T(n)$	$T(n) = \frac{U}{\sqrt{\frac{V}{n}}}$ $U \sim \mathcal{N}(0,1)$ $V \sim \chi^2(n)$	$E(T_n) = 0$ si $n > 0$	$V(T_n) = \frac{n}{n-2}$ si $n > 2$

0.4 Convergence de suites de v.a.

0.4.1 Inégalités importantes

1. **Inégalité de Markov** Si X est une v.a. qui vérifie : $E(|X|) < +\infty$.

$$\forall a > 0 \quad , \quad P\left(|X| \geq a\right) \leq \frac{E(|X|)}{a}$$

2. **Inégalité de Bienaymé - Tchebychev** Si X est une v.a. qui vérifie : $E(X^2) < +\infty$.

$$\forall a > 0 \quad , \quad P\left(|X - E(X)| \geq a\right) \leq \frac{V(X)}{a^2}$$

0.4.2 Types de convergence

Soit $(X_n)_{n \geq 0}$ une suite de v.a. et X une autre v.a. donnée.

1. **Convergence presque sûre**
   On dit que $(X_n)_{n\geq 0}$ converge presque sûrement vers X, si et ssi :
   $$P\left(X_n \to X\right) = 1 \Longleftrightarrow P\left(\left\{\omega \in \Omega, \lim_{n\to\infty} X_n(\omega) = X\right\}\right) = 1$$
   On note $\quad X_n \xrightarrow{p.s.} X$
2. **Convergence en moyenne quadratique**
   On dit que $(X_n)_{n\geq 0}$ converge en moyenne quadratique vers X, si et ssi :
   $$\lim_{n\to\infty} E\left((X_n - X)^2\right) = 0$$
   On note $\quad X_n \xrightarrow{m.q.} X$
3. **Convergence en probabilité**
   On dit que $(X_n)_{n\geq 0}$ converge en probabilité vers X, si et ssi :
   $$\forall\varepsilon > 0, \quad \lim_{n\to\infty} P\left(|X_n - X| \geq \varepsilon\right) = 0 \quad \text{ou bien} \quad \forall\varepsilon > 0, \quad \lim_{n\to\infty} P\left(|X_n - X| < \varepsilon\right) = 1$$
   On note $\quad X_n \xrightarrow{P} X$
4. **Convergence en loi**
   Si on note F_n et F les fonctions de répartition de X_n et de X, on dit que $(X_n)_{n\geq 0}$ converge en loi vers X, si et ssi :
   $$\lim_{n\to\infty} F_n(x) = F(x) \quad \text{partout où F est continue}$$
   On note $\quad X_n \xrightarrow{\mathcal{L}} X$

0.4.3 Relations entre les types de convergence

Entre ces différents types de convergence, on a les relations d'implication suivantes :

$$\begin{tikzcd}
\text{Cvgce en m.q.} \arrow[dr] & & \\
& \text{Cvgce en proba} \arrow[r] & \text{Cvgce en loi} \\
\text{Cvgce p.s.} \arrow[ur] & &
\end{tikzcd}$$

0.5 Théorèmes limites

0.5.1 Loi des grands nombres

1. **Loi faible des grands nombres**
   Si $(X_n)_{n\geq 0}$ est une suite de v.a. non corrélées (qui ne suivent pas forcément la même loi) de même variance finie σ^2, et de moyenne commune μ, alors
   la moyenne empirique : $\overline{X_n} = \frac{1}{n}\sum\limits_{i=1}^{n} X_i$ converge en probabilité vers μ. Autrement dit :
   $$\forall\varepsilon > 0, \quad \lim_{n\to\infty} P\left(|X_n - \mu| \geq \varepsilon\right) = 0$$
2. **Loi forte des grands nombres**
   Si $(X_n)_{n\geq 0}$ est une suite de v.a.i.i.d (indépendantes et qui suivent la même loi), intégrables $(E(|X_i|) < \infty)$, de moyenne commune μ, alors
   la moyenne empirique : $\overline{X_n} = \frac{1}{n}\sum\limits_{i=1}^{n} X_i$ converge presque sûrement vers μ. Autrement dit :
   $$P\left(X_n \to \mu\right) = 1 \Longleftrightarrow P\left(\left\{\omega \in \Omega, \lim_{n\to\infty} X_n(\omega) = \mu\right\}\right) = 1$$

0.5.2 Théorème de la limite centrale T.C.L.

Soit $(X_n)_{n\geq 0}$ une suite de v.a.i.i.d de variance commune finie σ^2 et de moyenne commune μ, alors la suite de v.a. $\dfrac{\overline{X_n}-\mu}{\sigma}\sqrt{n} = \dfrac{\sum\limits_{i=1}^{n} X_i - n\mu}{\sigma\sqrt{n}}$ converge en loi vers $\mathcal{N}(0,1)$.

En pratique, due $n > 30$, on consid que $\dfrac{\overline{X_n}-\mu}{\sigma}\sqrt{n} = \dfrac{\sum\limits_{i=1}^{n} X_i - n\mu}{\sigma\sqrt{n}}$ suit la loi normale centrt rite.