Asymptotic Statistics : Random vectors

M2RI - Toulouse University

Clément Lalanne

Basic definitions

Definition: Let $X = (X_1, \dots, X_p)$ be a random vector in $\mathbb{R}^p$, its cumulative distribution function is defined as
$F_X : \mathbb{R}^p \to [0,1], \qquad (x_1, \dots, x_p) \mapsto \mathbb{P}(X_1 \le x_1, \dots, X_p \le x_p)\,.$

Definition: Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of random vectors in $\mathbb{R}^p$, and let X be a random vector in $\mathbb{R}^p$. We say that (X_n) converges to X

- **In distribution**: if $F_{X_n}(x) \xrightarrow[n\to+\infty]{} F_X(x)$ for every continuity point x of F_X (noted $X_n \xrightarrow[n\to+\infty]{\mathscr{L},d} X$)
- **In probability**: if for all $\varepsilon > 0$, $\mathbb{P}\Big(\|X_n - X\| \ge \varepsilon\Big) \xrightarrow[n\to+\infty]{} 0$ (noted $X_n \xrightarrow[n\to+\infty]{\mathbb{P}} X$)
- **Almost surely**: if $\mathbb{P}\Big(\|X_n - X\| \to 0\Big) = 1$ (noted $X_n \xrightarrow[n\to+\infty]{a.s.} X$).

Remark: Convergence in law does not require that all random variables be defined on the same probability space.

Convergence in distribution

Theorem (Portmanteau): The following statements are equivalent

- $X_n \xrightarrow{d} X$
- For every bounded continuous function f, $\mathbb{E}[f(X_n)] \to \mathbb{E}[f(X)]$
- For every bounded Lipschitz function f, $\mathbb{E}[f(X_n)] \to \mathbb{E}[f(X)]$
- For every nonnegative continuous function f, $\liminf_{n\to\infty} \mathbb{E}[f(X_n)] \geq \mathbb{E}[f(X)]$
- For every open set $O \subset \mathbb{R}^p$, $\liminf_{n\to\infty} \mathbb{P}(X_n \in O) \geq \mathbb{P}(X \in O)$
- For every closed set $F \subset \mathbb{R}^p$, $\limsup_{n\to\infty} \mathbb{P}(X_n \in F) \leq \mathbb{P}(X \in F)$
- For every Borel set B such that $\mathbb{P}(X \in \partial B) = 0$, $\mathbb{P}(X_n \in B) \to \mathbb{P}(X \in B)$.

Exercice: Show that $X_n \stackrel{d}{\to} X$ does not always imply $X_n - X \stackrel{d}{\to} 0$.

Exercice: Show that $\delta_{x_n} \stackrel{d}{\to} \delta_x$ i.f.f. $x_n \to x$.

Fourier transform

Definition: Let X be a random vector in $\mathbb{R}^p$. The characteristic function of X, denoted by φ_X, is the Fourier transform of its distribution and is defined by $\forall t \in \mathbb{R}^p, \quad \varphi_X(t) = \mathbb{E}\left[e^{i\langle t,X\rangle}\right] = \int_{\mathbb{R}^p} e^{i\langle t,x\rangle}\, dP_X(x)$

Proposition: Let X, Y be random vectors in $\mathbb{R}^p$ with characteristic functions φ_X, φ_Y. We have:

- $\varphi_X = \varphi_Y \iff P_X = P_Y \iff X \overset{d}{=} Y$
- For all t, $|\varphi_X(t)| \leq 1$ and $\varphi_X(0) = 1$
- φ_X is uniformly continuous
- For any $A \in \mathbb{R}^{p\times p}$ and $b \in \mathbb{R}^p$, $\varphi_{AX+b}(t) = \varphi_X(A^T t)\, e^{i\langle b,t\rangle}, \quad \forall t$
- $\varphi_{-X}(t) = \varphi_X(-t) = \overline{\varphi_X(t)}, \quad \forall t$
- If $\mathbb{E}(\|X\|^p) < +\infty$ then $\varphi_X \in C^p$. Moreover, for any multi-index $m = (m_1, \ldots, m_p)$ with $|m| \leq p$, $\varphi_X^{(m)}(t) = i^{|m|}\, \mathbb{E}\big(X^m e^{i\langle t,X\rangle}\big)$
- The converse of the previous point is false, but still: If φ_X admits a derivative of even order p at 0, then X has finite absolute moments up to order \$p\$
- If X admits a density with respect to Lebesgue measure, then $\varphi_X(t) \to 0 \quad$ as $\|t\| \to \infty$
- If $\varphi_X \in L^1$, then X admits a continuous bounded density f and $f(x) = \dfrac{1}{(2\pi)^p} \displaystyle\int_{\mathbb{R}^p} e^{-i\langle t,x\rangle}\, \varphi_X(t)\, dt$
- X and Y are independent if and only if $\varphi_{(X,Y)}(s,t) = \varphi_X(s)\,\varphi_Y(t), \quad \forall (s,t)$
- If X and Y are independent, then $P_{X+Y} = P_X * P_Y, \quad$ and $\quad \varphi_{X+Y} = \varphi_X\,\varphi_Y$.

Exercice: Compute the Fourier transform of the standard normal distribution on $\mathbb{R}$ with density

$$x \mapsto \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right),$$

and deduce the Fourier transform of a general Gaussian distribution $\mathcal{N}(\mu, \Sigma)$ on $\mathbb{R}^d$ with density

$$x \mapsto \frac{1}{(2\pi)^{d/2}\sqrt{\det \Sigma}} \exp\left(-\tfrac{1}{2}\|x-\mu\|^2_{\Sigma^{-1}}\right)$$

Hint: You can try to find an ordinary differential equation satisfied by the characteristic function.

EQ: CV in distribution - CV Fourier transform

Definition: A family of random vectors $(X_\alpha)_{\alpha \in A}$ is said to be uniformly tight if for every $\varepsilon > 0$ there exists $M > 0$ such that $\sup_{\alpha \in A} \mathbb{P}\Big(\|X_\alpha\| > M \Big) \leq \varepsilon$.

Prohorov's Theorem:

- If $X_n \xrightarrow{d} X$, then $(X_n)_{n \geq 1}$ is uniformly tight.

- If $(X_n)_{n \geq 1}$ is uniformly tight then there exists a random vector X and a subsequence (X_{n_k}) such that $X_{n_k} \xrightarrow{d} X$.

Levy's Theorem:

- **If** $X_n \xrightarrow{d} X_n$, then $\varphi_{X_n}(t) \to \varphi_X(t), \quad \forall t$
- If there exists φ continuous at 0 such that $\varphi_{X_n}(t) \to \varphi(t), \quad \forall t$, then there exists a random vector X such that $\varphi = \varphi_X$ and $X_n \xrightarrow{d} X_n$.

Example: Weak law of large numbers

Let $X_1, \ldots, X_n \overset{\text{i.i.d.}}{\sim} X$. Then $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{d} \mu$ if the characteristic function φ_X is differentiable at 0 and $\varphi_X'(0) = i\mu$.

Remark: if $X \in L^1$, the **strong law of large numbers** guarantees that the convergence happens a.s. (but it is harder to prove).

Example: Central limit theorem

Let $X_1, \ldots, X_n \overset{\text{i.i.d.}}{\sim} X$ with $\mathbb{E}(X) = 0, \mathbb{E}(X^2) = 1$.. Then $\sqrt{n}\,\bar{X}_n \xrightarrow{d} \mathcal{N}(0,1)$.