Convergence of random variables (II)

APM 3F004 EP Asymptotic Statistics

S. Gadat

Ecole Polytechnique

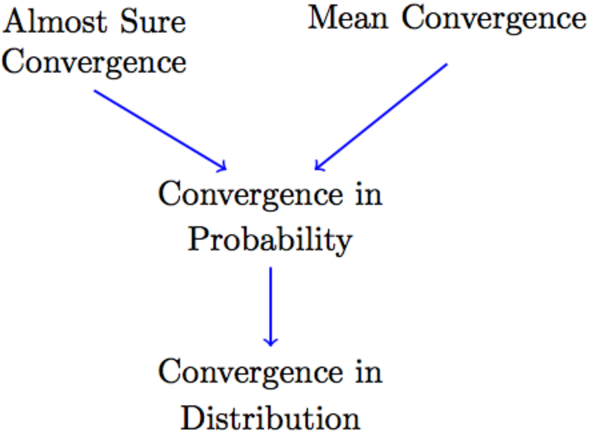
Plan

Assumptions

Let (X_n) be a sequence of real-valued random variables and X be a real-valued random variable, all defined on the same probability space $(\Omega, \mathcal{A}, \mathbb{P})$.
There are mainly four types of convergence involving (X_n) :

- Weak convergence in distribution
- Intermediate convergence in probability
- Convergence in quadratic mean
- Strong convergence almost sure

Link between convergences



Combining convergence statements

Remark

You have to be cautious when combining convergence statements. For example :

- *it is possible to add two convergence in probability,*
- *it is possible to add two convergence in quadratic mean*
- *it is possible to add two almost sure convergence*
- *But it does not necessarily work for convergence in distribution.*

Proposition (Continuous mapping theorem)

h is a continuous function

- *Assume that* $X_n \xrightarrow{d} X$, *then* $h(X_n) \xrightarrow{d} h(X)$.
- *Assume that* $X_n \xrightarrow{\mathbb{P}} X$, *then* $h(X_n) \xrightarrow{\mathbb{P}} h(X)$.
- *Assume that* $X_n \xrightarrow{a.s.} X$, *then* $h(X_n) \xrightarrow{a.s.} h(X)$.

Convergence of moments

- If $X_n \to_d X$ (or $X_n \to_p X$), it is tempting to say that $E(X_n) \to E(X)$; however, this statement is not true in general.
- For example, suppose that $P(X_n = 0) = 1 - n^{-1}$ and $P(X_n = n) = n^{-1}$. Then $X_n \to_p 0$ but $E(X_n) = 1$ for all n (and so converges to 1).
- To ensure convergence of moments, additional conditions are needed; these conditions effectively bound the amount of probability mass in the distribution of X_n concentrated near $\pm\infty$ for large n.

Plan

2 - Slutsky's Lemma

Very important result in statistics.

Lemme (Slutsky's Lemma)

Let $(X_n)_{n\in\mathbb{N}}$ and $(Y_n)_{n\in\mathbb{N}}$ be sequences of random variables defined on a common probability space. Suppose that

$$X_n \xrightarrow{d} X \quad \text{and} \quad Y_n \xrightarrow{\mathbb{P}} c,$$

where X is a random variable and $c \in \mathbb{R}$ is a constant. Then :

1. $X_n + Y_n \xrightarrow{d} X + c$,
2. $X_n Y_n \xrightarrow{d} Xc$,
3. *If* $c \neq 0$, *then* $\frac{X_n}{Y_n} \xrightarrow{d} \frac{X}{c}$.

Remark : Generally, used with CLT and LLN (see later on)

2 - Proof of Slutsky's Lemma

Plan

3 - Delta Method

The Delta method allows to derive some limit theorem for a transformed sequence that already satisfies a limit theorem.

Théorème

Suppose that (X_n) is a real sequence of random variables such that :

$$a_n(X_n - \theta) \to_d Z$$

where θ is a constant and $\{a_n\}$ is a sequence of constants with $a_n \uparrow \infty$. If $g(x)$ is a function with derivative $g'(\theta)$ at $x = \theta$ then

$$a_n(g(X_n) - g(\theta)) \to_d g'(\theta)Z$$

- a_n is the rate of convergence, and this rate is transported via the Delta method
- The limit distribution of Z becomes the one of $g'(\theta)Z$
- In general : $a_n = \sqrt{n}$ and $Z \sim \mathcal{N}(0, \sigma^2)$.

Example

Exercise 1. Let (Z_n) be a sequence of random variables such that

$$\sqrt{n} Z_n \xrightarrow{d} \mathcal{N}(0, \sigma^2).$$

- Find a limit theorem involving $\cos(Z_n)$.
- Same question with $\exp(Z_n)$.
- Let (X_n) be a sequence of random variables such that

  $$\sqrt{n}(X_n - 1) \xrightarrow{d} \mathcal{N}(0, \sigma^2).$$

  Limit theorem for X_n^{-1} ?

Multivariate Delta Method

Théorème

Suppose that (X_n) *is a sequence of random variables in* $\mathbb{R}^d$ *such that :*

$$a_n(X_n - \theta) \to_d Z$$

where θ *is a constant and* $\{a_n\}$ *is a sequence of constants with* $a_n \uparrow \infty$*. If* $g(x)$ *is a differentiable function with derivative* $Dg(\theta)$ *at* $x = \theta$ *then*

$$a_n(g(X_n) - g(\theta)) \to_d Dg(\theta)Z$$

- Important here to be careful with the derivative of g.
- In general, it reduces to the computation of a covariance matrix after a linear transform of a random centered gaussian random variable

Example

Exercise 2. Let (X_n) be a sequence of random variables such that

$$\sqrt{n}(X_n - \theta) \xrightarrow{d} \mathcal{N}(0, V),$$

with

$$\theta = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \quad \text{and} \quad V = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$$

Find a limit theorem on $\{X_{n,1}\}^2 + \{X_{n,2}\}^3$.

Plan

The weak law

Theorem (The weak law of large numbers)

Let $(X_n)_{n\geq 1}$ *be a sequence of i.i.d. real-valued random variables such that* $\mathbb{E}(|X_1|) < +\infty$. *Then*

$$\overline{X_n} := \frac{1}{n}\sum_{k=1}^{n} X_k \xrightarrow{\mathbb{P}} \mathbb{E}(X_1).$$

- In the case where $\mathbb{V}(X_1) < +\infty$, the proof is straightforward and it relies on the Bienaymé-Tchebychev inequality.
- The finite first moment condition is sufficient but not necessary. One can find *i.i.d.* sequences that satisfy the WLLN and where $\mathbb{E}(|X_1|) = +\infty$.

The strong law

Theorem (The strong law of large numbers)

Let $(X_n)_{n\geq 1}$ be a sequence of i.i.d. real-valued random variables. The two following statements are equivalent :

1. $\mathbb{E}(|X_1|) < +\infty$.
2. $\overline{X_n} := \frac{1}{n}\sum_{k=1}^{n} X_k \xrightarrow{a.s.} \mathbb{E}(X_1)$.

The proof of this result is *difficult !*

Convergence of the sample median in probability

- Suppose that $X_1, \cdots, X_n$ are i.i.d. random variables with a distribution function $F(x)$. Assume that the X_i 's have a unique median $\mu(F(\mu) = 1/2)$; in particular, this implies that for any $\varepsilon > 0, F(\mu + \varepsilon) > 1/2$ and $F(\mu - \varepsilon) < 1/2$.
- Let $X_{(1)}, \cdots, X_{(n)}$ be the order statistics of the X_i 's and define $Z_n = X_{(m_n)}$ where $\{m_n\}$ is a sequence of positive integers with $m_n/n \to 1/2$ as $n \to \infty$. For example, we could take $m_n = n/2$ if n is even and $m_n = (n+1)/2$ if n is odd ; in this case, Z_n is essentially the sample median of the X_i 's.
- Show that $Z_n \to_p \mu$ as $n \to \infty$.

Not so easy !

The central limit theorem

Theorem (CLT)

Let $(X_n)_{n\geq 1}$ be a sequence of i.i.d. real-valued random variables such that $\mathbb{V}(X_1) < +\infty$. We set $\mu = \mathbb{E}(X_1)$ and $\sigma^2 = \mathbb{V}(X_1)$. The central limit theorem states that

$$\sqrt{n}\frac{(\overline{X_n} - \mu)}{\sigma} \xrightarrow{d} \mathcal{N}(0,1).$$

<u>Informal Proof :</u>

CLT and approximations

Remark

One may use the CLT to find approximate distributions involving the normal distribution. For example :

- *For* $np(1-p) \geq 20$ *and* p *close to* $\frac{1}{2}$*, we have*

$$\mathcal{B}(n,p) \approx \mathcal{N}(np, np(1-p)).$$

- *For* $n > 30$*, we have*

$$\mathcal{P}(n) \approx \mathcal{N}(n,n).$$

- *For* $n > 30$*, we have*

$$\chi^2(n) \approx \mathcal{N}(n,2n).$$

Approximation of the binomial distribution

- We want to evaluate $P[a \le X \le b]$ for some integers a and b.
- A naive application of the CLT gives

$$
\begin{aligned}
& P[a \le X \le b] \\
& = P\left[\frac{a - n\theta}{\sqrt{n\theta(1-\theta)}} \le \frac{X - n\theta}{\sqrt{n\theta(1-\theta)}} \le \frac{b - n\theta}{\sqrt{n\theta(1-\theta)}}\right] \\
& \approx \Phi\left(\frac{b - n\theta}{\sqrt{n\theta(1-\theta)}}\right) - \Phi\left(\frac{a - n\theta}{\sqrt{n\theta(1-\theta)}}\right)
\end{aligned}
$$

Normal approximation of the binomial distribution

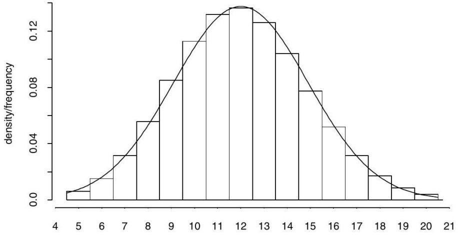


Figure – Binomial distribution ($n = 40, \theta = 0.3$) and approximating Normal density

The multivariate CLT

Theorem (The multivariate CLT)

Let $(X_n)_{n\geq 1}$ be a sequence of $i.i.d.$ d-dimensional random vectors with a finite covariance matrix. We set $\mu = \mathbb{E}(X_1)$ and Σ the covariance matrix. The multivariate central limit theorem states that

$$\sqrt{n}\left(\overline{X_n} - \mu\right) \xrightarrow{d} \mathcal{N}_d(0, \Sigma).$$

Plan

5 - Applications

Exercise 3. Let $(X_n)_{n\geq 1}$ be a sequence of $i.i.d.$ real-valued random variables such that $\mathbb{V}(X_1) < +\infty$. We set $\mu = \mathbb{E}(X_1)$ and $\sigma^2 = \mathbb{V}(X_1)$. Find the limit in distribution of the sequence

$$\sqrt{n}\left(\overline{X_n}^2 - \mu^2\right).$$

Exercise 4. Let (X_n) be a sequence of random variables $\mathcal{P}(\theta)$. Find a limit theorem on $\sqrt{\bar{X}_n}$.

Variance Stabilizing transform for Bernoulli random variables

Exercise 5.

- Suppose that $X_1, \cdots, X_n$ are i.i.d. Bernoulli random variables with parameter θ. Then

$$\sqrt{n}\left(\bar{X}_n - \theta\right) \to_d Z \sim N(0, \theta(1-\theta))$$

- Find g such that

$$\sqrt{n}\left(g\left(\bar{X}_n\right) - g(\theta)\right) \to_d N(0,1)$$