APUNTES PROBABILIDAD

Tema 3: Convergencias Estocásticas. Teoremas Límite

Doble Grado en Informática y Estadística (INdat)

Autor

Víctor Elvira Fernández

Universidad de Valladolid

November 6, 2024

Contents

1 Convergencias estocásticas

1.1 Convergencia casi seguro

Definición 1.1. *Sea $\{X_n\}_{n=1}^{\infty}$ una sucesión de v.a. definidas en un espacio probabilístico (Ω, p) y sea X una v.a. sobre el mismo espacio probabilístico. Se dice que $\{X_n\}_{n=1}^{\infty}$* ***converge casi seguro*** *hacia X si*

$$\lim_{n\to\infty} X_n(\omega) = X(\omega) \quad \forall \omega \in A \subseteq \Omega \text{ con } P(A) = 1$$

Y lo representaremos como

$$X_n \xrightarrow{c.s} X$$

Esta definición suaviza el concepto de convergencia funcional que exige la la convergencia de de $X_n(\omega)$ a $X(\omega)$ para todo $\omega \in \Omega$. La convergencia c.s. resulta todavía un concepto demasiado exigente para las situaciones prácticas en las que se necesita, por ello definimos una convergencia menos fuerte.

1.2 Convergencia en probabilidad

Definición 1.2. *Sea $\{X_n\}_{n=1}^{\infty}$ una sucesión de v.a. definidas en un espacio probabilístico (Ω, p) y sea X una v.a. sobre el mismo espacio probabilístico. Se dice que $\{X_n\}_{n=1}^{\infty}$* ***converge en probabilidad*** *hacia X si*

$$\lim_{n\to\infty} P(|X_n - X| > \varepsilon) = 0 \quad \forall \varepsilon > 0$$

Y lo representaremos como

$$X_n \xrightarrow{p} X$$

- **Nota:** La condición de convergencia en probabilidad también se puede escribir como

$$\lim_{n\to\infty} P(|X_n - X| \leq \varepsilon) = 1 \quad \forall \varepsilon > 0$$

Propiedades de la convergencia en probabilidad

1. $X_n \xrightarrow{p} X \iff X_n - X \xrightarrow{p} 0$
2. Si $X_n \xrightarrow{p} X$ y $X_n \xrightarrow{p} Y \Longrightarrow P(X = Y) = 1$, se dice que X es c.s igual a Y
3. Si $X_n \xrightarrow{p} X$ y g es una función continua $g : \mathbb{R} \longrightarrow \mathbb{R}$ entonces $g(X_n) \xrightarrow{p} g(X)$
   - $kX_n \xrightarrow{p} kX \quad \forall k \in \mathbb{R}$
   - $X_n^k \xrightarrow{p} X^k \quad \forall k > 0$
4. Si $X_n \xrightarrow{p} X$ y $Y_n \xrightarrow{p} Y$ y g es una función continua $g : \mathbb{R}^2 \longrightarrow \mathbb{R}$ entonces $g(X_n, Y_n) \xrightarrow{p} g(X, Y)$
   - $X_n + Y_n \xrightarrow{p} X + Y$

- $X_n \times Y_n \xrightarrow{p} X \times Y$
- $\frac{X_n}{Y_n} \xrightarrow{p} \frac{X}{Y}$ siempre que $P(Y_n \neq 0) = 1$ *y* $P(Y \neq 0) = 1$

5. Si $X_n \xrightarrow{c.s} X \Longrightarrow X_n \xrightarrow{p} X$ pero el recíproco no es cierto

1.3 Convergencia en ley

Definición 1.3. *Sea $\{X_n\}_{n=1}^{\infty}$ una sucesión de v.a. y sea X otra v.a. Sean también F_n y F las funciones de distribución de X_n y X respectivamente. Se dice que $\{X_n\}_{n=1}^{\infty}$* ***converge en ley o distribución*** *hacia X si*

$$\lim_{n\to\infty} F_n(x) = F(X) \quad \forall x \in \mathbb{R} \text{ en los que } F \text{ es continua}$$

Y lo representaremos como

$$X_n \xrightarrow{L} X$$

- **Nota:** Observar que las v.a. X_n y X no tienen por que estar definidas en el mismo espacio probabilístico (Ω, p)

Propiedades de la convergencia en Ley

1. Sea $\{X_n\}_{n=1}^{\infty}$ una sucesión de v.a. que toma valores enteros, entonces:

$$X_n \xrightarrow{L} X \iff P(X_n = k) \xrightarrow[n\to\infty]{} P(X = k) \;\; \forall k \in \mathbb{N}$$

2. Sea $\{X_n\}_{n=1}^{\infty}$ una sucesión de v.a. y sea X una v.a., todas continuas, y sean f_n y f las funciones de densidad de X_n y X respectivamente

$$f_n(x) \xrightarrow[n\to\infty]{} f(x) \;\; \forall x \Longrightarrow X_n \xrightarrow{L} X$$

3. Sea $\{X_n\}_{n=1}^{\infty}$ una sucesión de v.a. tal que $X_n \xrightarrow{L} X$ y sea $g : \mathbb{R} \longrightarrow \mathbb{R}$ función continua, entonces

$$g(X_n) \xrightarrow{L} g(X)$$

4. **Teorema de Polya:** Si $\lim_{n\to\infty} F_n(x) = F(X)$ para todo $x \in \mathbb{R}$ y F es continua entonces la convergencia es uniforme; es decir

$$\lim_{n\to\infty} \{\sup_{x\in\mathbb{R}} |F_n(x) - F(x)|\} = 0$$

5. Si $X_n \xrightarrow{p} X \Longrightarrow X_n \xrightarrow{L} X$ pero el recíproco no es cierto

6. Sea $k \in \mathbb{R}$, $X_n \xrightarrow{p} k \iff X_n \xrightarrow{L} k$ (converge a una variable degenerada)

Para las siguientes propiedades sea $\{(X_n, Y_n)\}_{n=1}^{\infty}$ una sucesión de vectores aleatorios definidos sobre el mismo (Ω, p) y sea $c \in \mathbb{R}$ una constante

1. Si $|Y_n - X_n| \xrightarrow{p} 0$ y $Y_n \xrightarrow{L} Y$ entonces $X_n \xrightarrow{L} Y$
2. Si $X_n \xrightarrow{L} X$ y $Y_n \xrightarrow{p} c$ entonces
   - $X_n \pm Y_n \xrightarrow{L} X \pm c$
   - $X_n \times Y_n \xrightarrow{L} cX$
   - $\frac{X_n}{Y_n} \xrightarrow{L} \frac{X}{c}$ si $c \neq 0$

2 Teoremas Límite

2.1 Leyes de los grandes números

El objetivo de las leyes de los grandes números es estudiar el comportamiento probabilístico de las medias muestrales de v.a. cuando el número de variables promediadas tiende hacia infinito, es decir, sea $\{X_n\}_{n=1}^{\infty}$ una sucesión de v.a. se trata de estudiar el comportamiento de $\{\overline{X}_n\}_{n=1}^{\infty}$ donde

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

Si el comportamiento asintótico se refleja mediante una convergencia en probabilidad se hablará de ley débil y si es mediante una convergencia casi seguro se hablará de ley fuerte

Ley fuerte de los grandes números de Kolmogorov

Definición 1.4 *Sea $\{X_n\}_{n=1}^{\infty}$ una sucesión de v.a.i.i.d.*

$$\frac{S_n}{n} \xrightarrow{c.s} \mu = EX_i \iff \textit{Existe la esperanza de } X_i$$

Si hay convergencia hacia una constante, esa es su esperanza.

Ley débil de los grandes números (L.g.n.)

Definición 1.5 *Sea $\{X_n\}_{n=1}^{\infty}$ una sucesión de v.a. incorreladas para las que existe el momento de segundo orden y tienen media y varianza común; es decir, podemos escribir $EX_i = \mu$ y $Var(X_i) = \sigma^2$, entonces*

$$\frac{S_n}{n} \xrightarrow{p} \mu$$

- **Nota:** Para la débil no es necesario que las variables sean independientes ni estén igualmente distribuidas, pero tiene que existir el momento de orden dos, para la fuerte, es necesario que sean v.a.i.i.d. pero basta con que exista el momento de primer orden

Prueba de la L.g.n.

Sea $\varepsilon > 0$, utilizando la desigualdad de Chebyshev tenemos

$$0 \leq P(|\frac{S_n}{n} - \mu| > \varepsilon) \leq \frac{Var(\frac{S_n}{n})}{\varepsilon^2} = \frac{Var(S_n)}{\varepsilon^2 n^2} = \frac{n\sigma^2}{\varepsilon^2 n^2} = \frac{\sigma^2}{\varepsilon^2 n} \xrightarrow[x\longrightarrow\infty]{} 0$$

2.2 Teorema Central del Límite (T.C.L)

Definición 1.6 *Sea* $\{X_n\}_{n=1}^{\infty}$ *una sucesión de v.a.i.i.d para las que existe el momento de segundo orden, con* $EX_i = \mu$ *y* $Var(X_i) = \sigma^2 \neq 0$ *se verifica que*

$$\frac{S_n - n\mu}{\sigma\sqrt{n}} \xrightarrow{L} Z \quad con \quad Z \sim N(0,1)$$

- **Nota:** Puesto que la función de distribución de la ley Normal es continua se tiene la convergencia uniforme de las funciones de distribución de $\frac{S_n - n\mu}{\sigma\sqrt{n}}$ hacia la de la normal.

Este resultado nos asegura que para **n suficientemente grande** podemos aproximar la distribución de S_n por la $N(n\mu, n\sigma^2)$.
La rapidez de convergencia del T.C.L. depende de la distribución exacta de las variables X_i se parezca o no a la normal. Sobre ello se pueden dar las siguientes referencias:

1. Si las X_i son Normales, el resultado es cierto sea cual sea el valor de N puesto que

$$S_n = \sum_{i=1}^{n} X_i \sim N(n\mu, n\sigma^2)$$

2. Si las variables tienen una distribución acampanada y más o menos simétrica, valores de n de 5 o 6 bastan para que la función de distribución de $\frac{S_n - n\mu}{\sigma\sqrt{n}}$ sea prácticamente indistinguible de la $N(0,1)$

3. Si la distribución es poco simétrica, pero sigue siendo razonablemente acampanada y unimodal, pueden conseguirse buenas aproximaciones para valores de n de entre 20 y 30

4. Si la distribución es de las que se denominan en U, pueden ser necesarios valores de n mayores de 50 o incluso de 100 para conseguir una aproximación razonable de las funciones de distribución

2.3 Método Delta

Definición 1.7 *Sea* $\{X_n\}_{n=1}^{\infty}$ *una sucesión de v.a.i.i.d para las que existe el momento de segundo orden, con* $EX_i = \mu$ *y* $Var(X_i) = \sigma^2 \neq 0$ *y sea g una función real de variable real, continua con derivada en* μ *no nula* $g'(\mu) \neq 0$*, entonces*

$$\frac{g(\frac{S_n}{n}) - g(\mu)}{\sigma|g'(\mu)|}\sqrt{n} \xrightarrow{L} N(0,1)$$

Este resultado nos garantiza que para **n suficientemente grande** se puede usar la siguiente distribución aproximada de $g(\overline{X})$

$$g\left(\frac{S_n}{n}\right) \underset{Aprox}{\sim} N\Big(g(\mu), \frac{\sigma^2}{n}(g'(\mu))^2\Big)$$

2.4 Momentos Muestrales

2.4.1 Definiciones importantes

Definiciones 1.8., 1.9., 1.10. y 1.11. *Sean $X_1, X_2, ...X_n$ v.a.i.i.d. con distribución F y sea $k = 1, 2, ...$ definimos*

- Momento **muestral** de orden k respecto del origen

$$a_{nk} = \frac{\sum_{i=1}^{n} X_i^k}{n}$$

- Momento **muestral** de orden k respecto de la media

$$b_{nk} = \frac{\sum_{i=1}^{n} (X_i - \overline{X})^k}{n}$$

- Momento **teórico** de orden k respecto del origen para la distribución F

$$\alpha_k = EX^k \ \ \textit{donde} \ \ X \sim F$$

- Momento **teórico** de orden k respecto de la media para la distribución F

$$\beta_k = E(X - EX)^k \ \ \textit{donde} \ \ X \sim F$$

2.4.2 Propiedades de los momentos muestrales

1. Si existe el momento de orden k para F

$$a_{nk} \xrightarrow{p} \alpha_k \ \ y \ \ b_{nk} \xrightarrow{p} \beta_k$$

2. Si existe el momento de orden 2k para F, para n suficientemente grande podemos aproximar la distribución a_{nk} por una normal de media α_k y varianza $\frac{\alpha_{2k} - (\alpha_k)^2}{n}$

2.4.3 Distribución asintótica de la varianza muestral

Proposición 1.12. *Sea $\{X_n\}_{n=1}^{\infty}$ una sucesión de v.a.i.i.d para las que existe el momento de cuarto orden y $\beta_4 > \sigma_4$, denotando $EX_i = \mu$, $Var(X_i) = \sigma^2 \neq 0$ y $\beta_4 = E((X_i - EX_i)^4)$, entonces*

$$\sqrt{n}(S_n^2 - \sigma^2) \xrightarrow{L} N(0, \beta_4 - \sigma_4)$$

Por lo que, para n suficientemente grande, se puede utilizar la siguiente distribución aproximada de la varianza muestral

$$S_n^2 = \frac{\sum_{i=1}^{n} (X_i - \overline{X})^2}{n} \underset{Aprox.}{\sim} N\Big(\sigma^2, \frac{\beta_4 - \sigma_4}{n}\Big)$$

2.5 Teorema Fundamental de la Estadística. Función de distribución muestral

Definición 1.13. *Sean $X_1, X_2, ..., X_n$ v.a. igualmente distribuidas con distribución F, definidas en un mismo espacio probabilístico (Ω, p), se define la **función de distribución muestral o empírica** como la función $F_n : \mathbb{R} \longrightarrow [0,1]$ que a cada $x \in \mathbb{R}$ le asigna el valor*

$$F_n(x) = \frac{\textit{número de } X_i \leq x}{n}$$

Podemos escribir $F_n(x)$ en función de las siguientes variables de Bernouilli

$$F_n(x) = \frac{\sum_{i=1}^{n} Y_i}{n} \quad \textit{donde } Y_i = \begin{cases} 1 & si \quad X_i \leq x \\ 0 & si \quad X_i > x \end{cases}$$

O bien como una función escalonada como

$$F_n(x) = \begin{cases} 0 & si & x < X_{(1)} \\ \frac{1}{n} & si & X_{(1)} \leq x < X_{(2)} \\ & & ... \\ \frac{k}{n} & si & X_{(k)} \leq x < X_{(k+1)} \\ & & ... \\ \frac{n-1}{n} & si & X_{(n-1)} \leq x < X_{(n)} \\ 1 & si & x \geq X_{(n)} \end{cases}$$

Se observa que:

1. Fijo ω, esta función es la función de distribución teórica de una v.a. uniforme discreta que toma valores en $X_1(\omega), X_2(\omega), ..., X_n(\omega)$

2. Fijo $x \in \mathbb{R}$, $F_n(x)$ es una v.a. definida sobre (Ω, p) en función de la muestra $X_1, X_2, ..., X_n$

   Si suponemos la independencia entre las X_i podemos además observar que:

3. Fijo $x \in \mathbb{R}$, $F_n(x) \xrightarrow{p} F(x)$. La prueba es la aplicación de la L.g.n. aplicada a las variables $Y_1, Y_2, ..., Y_n$ definidas arriba

4. Fijo $x \in \mathbb{R}$, para n suficientemente grande la distribución aproximada de $F_n(x)$ (aplicando el T.C.L. a las variables $Y_1, Y_2, ..., Y_n$) es

$$N\left(F(x), \frac{F(x)(1 - F(x))}{n}\right)$$

Además, 3. no es solo una convergencia puntual para cada $x \in \mathbb{R}$, si no que la convergencia es uniforme en x; es decir,

$$D_n = \underset{x \in \mathbb{R}}{Sup} |F_n(x) - F(x)| \xrightarrow{p} 0$$

Cuando el resultado anterior se enuncia con convergencia casi seguro y uniforme en x se le conoce como **Teorema de Glivenko Cantelli o Fundamental de la estadística.**

References

[1] Pilar Rodríguez de Tío,
Vectores aleatorios y distribuciones multivariantes.
Universidad de Valladolid, 2023.

[2] Pilar Rodríguez de Tío,
Distribuciones asociadas a variables independientes igualmente distribuidas.
Universidad de Valladolid, 2023.

[3] Pilar Rodríguez de Tío,
Convergencias estocásticas. Teoremas límite.
Universidad de Valladolid, 2023.