Estadística. Tema 5

Convergencia de Variables Aleatorias

Objetivos

1. Motivación estudio secuencias de VAs.
2. Convergencia de valores.
3. Convergencia de distribuciones de probabilidad.

Bibliografía

Dekking et al. (2005). Capítulos 13-14.

Wasserman (2005). Capítulo 5.

Wooldridge (2006). Apéndice B.

Este tema introduce la teoría para grandes muestras o asintótica, y esta basada en el estudio del comportamiento límite de secuencia de VAs. Los dos resultados básicos son la ley de los grandes números y el teorema central del límite.

5.1. Tipos de convergencia

Sea $X_1, X_2, \ldots,$ una secuencia de VAs y sea X otra VA. Sea F_n la CDF de X_n y sea F la CDF de X.

X_n converge a X en **probabilidad**, $X_n \to_p X$, si para cada $\epsilon > 0$,

$$\Pr\left(|X_n - X| > \epsilon\right) \to 0$$

cuando $n \to \infty$.

Si $X_n \to_p X$, también escribimos

$$p \lim_{n\to\infty} X_n = X$$

y también

$$X_n = X + o_p(1)$$

donde $o_p(1)$ indica que la diferencia $X_n - X$ converge a 0 en probabilidad.

X_n converge a X en **distribución**, $X_n \to_d X$, si

$$\lim_{n\to\infty} F_n(t) = F(t)$$

para todo t donde F es continua.

X_n converge a X en **media cuadrática**, $X_n \to_2 X$, si

$$E\left(X_n - X\right)^2 \to 0$$

cuando $n \to \infty$.

Ejemplo: si $X_n \sim N\left(0, 1/n\right)$, entonces $X_n \to_p 0, \to_d 0$.

Relación entre las convergencias:

- $X_n \to_2 X$ implica que $X_n \to_p X$.
- $X_n \to_p X$ implica que $X_n \to_d X$.
- $X_n \to_d X$ y si $\Pr(X = c) = 1$ para algún número real c, entonces $X_n \to_p X$.
- En general ninguna de las implicaciones contrarias es cierta, con esta excepción.
- Tampoco es cierto en general que si $X_n \to_p X$ implica que $EX_n \to EX$.

Propiedades: Sean X_n, X, Y_n, Y RVs y g una función continua.

- $X_n \to_p X$ y $Y_n \to_p Y$, entonces $X_n + Y_n \to_p X + Y$.
- $X_n \to_2 X$ y $Y_n \to_2 Y$, entonces $X_n + Y_n \to_2 X + Y$.
- $X_n \to_d X$ y $Y_n \to_d c$, entonces $X_n + Y_n \to_d X + c$.[Slutsky Th]
- $X_n \to_p X$ y $Y_n \to_p Y$, entonces $X_n \cdot Y_n \to_p X \cdot Y$.
- $X_n \to_d X$ y $Y_n \to_d c$, entonces $X_n \cdot Y_n \to_d cX$. [Slutsky Th]
- $X_n \to_p X$, entonces $g(X_n) \to_p g(X)$.
- $X_n \to_d X$, entonces $g(X_n) \to_d g(X)$. [Continuous Mapping Th]

En general no es verdad que $X_n \to_d X$ y $Y_n \to_d Y$, entonces $X_n + Y_n \to_d X + Y$.

5.2. Ley de los Grandes Números

Sea $X_1, X_2, \ldots, X_n$ una muestra IID con $\mu = EX_1$ y $\sigma^2 = V(X_1)$. La media muestral es

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

con $E\left(\bar{X}_n\right) = \mu$ y $V\left(\bar{X}_n\right) = \sigma^2/n$.

Teorema. **Ley Débil de los Grandes Números**. Si $X_1, X_2, \ldots, X_n$ son una muestra IID, entonces $\bar{X}_n \rightarrow_p \mu$.

Prueba: Desigualdad de Chebysev, asumiendo que $\sigma < \infty$.

Ejemplo. $X_i \sim$ Bernoulli($p = 0{,}5$)

$$\Pr\left(0{,}4 \leq \bar{X}_n \leq 0{,}6\right)?$$

En ese caso decimos que $\bar{X}_n$ es *consistente* para estimar p.

Ejemplo: Convergencia **varianza muestral** es

$$\begin{aligned}
S_n^2 &= \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)^2 \\
&= \left(\frac{n}{n-1}\right)\frac{1}{n}\sum_{i=1}^{n} X_i^2 - \frac{n}{n-1}\bar{X}_n^2 \\
&\rightarrow_p \left(\lim_{n\to\infty}\frac{n}{n-1}\right) p\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n} X_i^2 - \left(\lim_{n\to\infty}\frac{n}{n-1}\right) p\lim_{n\to\infty}\bar{X}_n^2 \\
&= 1\cdot EX^2 - 1\cdot(EX)^2 = EX^2 - \mu^2 = \sigma^2.
\end{aligned}$$

5.3. Teorema Central del Límite

Por la LDGN sabemos que $\bar{X}_n$ está cerca de μ con alta probabilidad según aumenta n, pero eso no nos permite conocer la distribución de probabilidad de $\bar{X}_n$.

Sea $X_1, X_2, \ldots, X_n$ una muestra IID con $\mu = EX_1$ y $\sigma^2 = V(X_1)$. Entonces, para n grande, $\bar{X}_n$ tiene una distribución aproximadamente Normal con media $E\left(\bar{X}_n\right) = \mu$ y varfianza $V\left(\bar{X}_n\right) = \sigma^2/n$.

Teorema. **Teorema Central del Límite.** Si $X_1, X_2, \ldots, X_n$ son una muestra IID con $\mu = EX_1$ y $\sigma^2 = V(X_1)$, entonces

$$Z_n = \frac{\bar{X}_n - \mu}{\sqrt{V\left(\bar{X}_n\right)}} = \frac{\sqrt{n}\left(\bar{X}_n - \mu\right)}{\sigma} \to_d Z,$$

donde $Z \sim N(0,1)$.

Notación. Alternativas:

$$\begin{aligned}
Z_n &\to {}_d N(0,1) \\
Z_n &\approx N(0,1) \\
\bar{X}_n &\approx N\left(\mu, \frac{\sigma^2}{n}\right) \\
\bar{X}_n - \mu &\approx N\left(0, \frac{\sigma^2}{n}\right) \\
\sqrt{n}\left(\bar{X}_n - \mu\right) &\approx N\left(0, \frac{\sigma^2}{n}\right) \\
\frac{\sqrt{n}\left(\bar{X}_n - \mu\right)}{\sigma} &\approx N(0,1).
\end{aligned}$$

En general no sabemos cuanto vale σ, pero aplicando Slutsky, tenemos que:

Teorema. Bajo las mismas condiciones del TCL, como $S_n \to_p \sigma$ porque $g(x) = x^2$ es una función continua,

$$\frac{\sqrt{n}\left(\bar{X}_n - \mu\right)}{S_n} \to_d N(0,1).$$

Teorema Central del Límite Multivariante. Sean $X_1, X_2, \ldots, X_n$ vectores aleatorios IID con

$$X_i = \begin{pmatrix} X_{1i} \\ X_{2i} \\ \vdots \\ X_{ki} \end{pmatrix}$$

con media

$$\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_k \end{pmatrix} = \begin{pmatrix} EX_{1i} \\ EX_{2i} \\ \vdots \\ EX_{ki} \end{pmatrix}$$

y matriz de varianzas-covarianzas Σ,

$$\Sigma = \begin{pmatrix} V\left(X_{1i}\right) & \cdots & C\left(X_{1i}, X_{ki}\right) \\ \vdots & \ddots & \vdots \\ C\left(X_{ki}, X_{1i}\right) & \cdots & V\left(X_{ki}\right) \end{pmatrix}.$$

Entonces

$$\begin{aligned} \sqrt{n}\left(\bar{X}_n - \mu\right) &\to_d N\left(0, \Sigma\right), \\ \sqrt{n}\Sigma^{1/2}\left(\bar{X}_n - \mu\right) &\to_d N\left(0, I_k\right) \end{aligned}$$

donde I_k es la matriz identidad de dimensión k.

5.4. Método delta

Teorema. (**Método Delta**) Si

$$\frac{\sqrt{n}\left(\bar{Y}_n-\mu\right)}{\sigma}\rightarrow_d N\left(0,1\right)$$

y g es diferenciable, tal que $g'\left(\mu\right)\neq 0$, entonces

$$\frac{\sqrt{n}\left(g\left(\bar{Y}_n\right)-g\left(\mu\right)\right)}{\left|g'\left(\mu\right)\right|\sigma}\rightarrow_d N\left(0,1\right),$$

es decir si

$$\bar{Y}_n\approx N\left(\mu,\frac{\sigma^2}{n}\right)$$

entonces

$$g\left(\bar{Y}_n\right)\approx N\left(g\left(\mu\right),\left(g'\left(\mu\right)\right)^2\frac{\sigma^2}{n}\right).$$

Teorema. **(Método Delta Multivariante)**. Si $Y_n=\left(Y_{n1},\ldots Y_{nk}\right)$ es una secuencia de vectores aleatorios tal que

$$\sqrt{n}\left(\bar{Y}_n-\mu\right)\rightarrow_d N\left(0,\Sigma\right)$$

y $g:\mathbb{R}^k\rightarrow\mathbb{R}$ es diferenciable, tal que $\nabla_\mu=\nabla g\left(\mu\right)$, entonces

$$\sqrt{n}\left(g\left(\bar{Y}_n\right)-g\left(\mu\right)\right)\rightarrow_d N\left(0,\nabla_\mu^T\Sigma\nabla_\mu\right).$$

Ejemplo. Distribución asintótica de $\bar{X}_n^2$. Sea $g\left(x\right)=x^2$, entonces $g'\left(x\right)=2x$. Por tanto, si

$$\bar{X}_n\approx N\left(\mu,\frac{\sigma^2}{n}\right)$$

entonces

$$\begin{aligned} g\left(\bar{X}_n\right) &\approx N\left(g\left(\mu\right),\left(g'\left(\mu\right)\right)^2\frac{\sigma^2}{n}\right) \\ &= N\left(\mu^2,\left(2\mu\right)^2\frac{\sigma^2}{n}\right) \end{aligned}$$

es decir

$$\sqrt{n}\frac{\bar{X}_n^2-\mu^2}{2\mu\sigma}\approx N\left(0,1\right).$$

¿Qué ocurre si $\mu=0$?