Tarea 3: Corrección de errores páginas 46 a 60 de la ficha de Convergencia

Inferencia Estadística

Lin Chiu Chen Yang

November 26, 2022

Universidad Nacional de Ingeniería

Facultad de Ingeniería

Económica, Estadística y Ciencias Sociales

FIEECS

que $\dfrac{\mathbf{n}}{\mathbf{n}+1} \to 1$ cuando $\mathbf{n} \to \infty$. Ahora,

$$\frac{1}{\mathbf{n}}\sum_{i=1}^{\mathbf{n}}(X_i - \overline{X})^2 = \frac{1}{\mathbf{n}}\sum_{i=1}^{\mathbf{n}} X_i^2 - \overline{X}^2$$

(una identidad algebraica). Como F tiene una varianza finita, también posee un segundo momento finito, es decir,$\mathbf{E}(X^2) = \sigma^2+\mu^2 < \infty$. Aplicando la ley fuerte de los grandes números a la sucesión $X_1^2, X_2^2, \ldots$ se obtiene $\frac{1}{\mathbf{n}}\sum_{i=1}^{\mathbf{n}} X_i^2 \xrightarrow{c.s.} \mathbf{E}(X^2) = \sigma^2 + \mu^2$. Aplicando el LDGN a la secuencia $X_1, X_2, \ldots$, se obtiene $\overline{X}_{\mathbf{n}} \xrightarrow{c.s.} \mu$, y por lo tanto por el teorema de aplicación continuo, $(\overline{X})^2 \xrightarrow{c.s.} \mu^2$. Ahora, por el teorema sobre la preservación de la convergencia, se obtiene que $\frac{1}{\mathbf{n}}\sum_{i=1}^{\mathbf{n}} X_i^2 - \overline{X}^2 \xrightarrow{c.s.} \sigma^2 + \mu^2 - \mu^2 = \sigma^2$, que termina la prueba.

Ejemplo 1.30. (**Convergencia de la correlación muestral**). Suponga $F(x, y)$ is a de distribución conjunta en $\mathbb{R}^2$, y suponga que $\mathbf{E}(X^2)$, $\mathbf{E}(Y^2)$ son ambos finitos. Sea

$$\rho = \frac{\mathbf{Cov}(X,Y)}{\sqrt{\mathbf{V}ar(X)\mathbf{V}ar(Y)}} = \frac{\mathbf{E}(XY) - \mathbf{E}(X)\mathbf{E}(Y)}{\sqrt{\mathbf{V}ar(X)\mathbf{V}ar(Y)}}$$

la correlación entre X y Y. Suponga $(X_i, Y_i), 1 \leq i \leq \mathbf{n}$ son $\mathbf{n}$ observaciones independientes de la del función de distribución conjunta F. El coeficiente de **correlación muestral** se define como

$$r = \frac{\displaystyle\sum_{i=1}^{\mathbf{n}}(X_i - \overline{X})(Y_i - \overline{Y})}{\sqrt{\displaystyle\sum_{i=1}^{\mathbf{n}}(X_i - \overline{X})^2 \sum_{i=1}^{\mathbf{n}}(Y_i - \overline{Y})^2}}$$

El propósito de este ejemplo es mostrar que r converge casi con seguridad a ρ.

Es conveniente reescribir r en la forma equivalente

$$r = \frac{\displaystyle\frac{1}{\mathbf{n}}\sum_{i=1}^{\mathbf{n}}(X_i - \overline{X})(Y_i - \overline{Y})}{\sqrt{\displaystyle\frac{1}{\mathbf{n}}\sum_{i=1}^{\mathbf{n}}(X_i - \overline{X})^2}\sqrt{\displaystyle\frac{1}{\mathbf{n}}\sum_{i=1}^{\mathbf{n}}(Y_i - \overline{Y})^2}}.$$

Por la LDGN, $\frac{1}{\mathbf{n}}\sum_{i=1}^{\mathbf{n}} X_iY_i$ converge casi con seguridad a $\mathbf{E}(XY)$ y $\overline{X}, \overline{Y}$ convergen casi con seguridad a $\mathbf{E}(X)$, $\mathbf{E}(Y)$. Por el ejemplo 1.29, $\frac{1}{\mathbf{n}}\sum_{i=1}^{\mathbf{n}}(X_i - \overline{X})^2$ converge casi con seguridad a $\mathbf{V}ar(X)$, y $\frac{1}{\mathbf{n}}\sum_{i=1}^{\mathbf{n}}(Y_i - \overline{Y})^2$ converge casi con seguridad a $\mathbf{V}ar(Y)$. Ahora considere la función $f(s,t,u,v,w) = \frac{s - tu}{\sqrt{v}\sqrt{w}}$, $-\infty < s,t,u < \infty$, $v,w > 0$. Esta función es continua en el conjunto $S = \{(s,t,u,v,w)\colon -\infty < s,t,u < \infty, v,w > 0, (s-tu)^2 \leq vw\}$. La distribución conjunta de $\frac{1}{\mathbf{n}}\sum_{i=1}^{\mathbf{n}} X_iY_i$, $\overline{X}$, $\overline{Y}$, $\frac{1}{\mathbf{n}}\sum_{i=1}^{\mathbf{n}}(X_i - \overline{X})^2$,$\frac{1}{\mathbf{n}}\sum_{i=1}^{\mathbf{n}}(Y_i - \overline{Y})^2$ asigna probabilidad uno al conjunto S. Por el teorema de función continua, se deduce que $r \xrightarrow{c.s} \rho$.

1.10 Convergencia en Distribución

En lo que sigue, $\{X_{\mathbf{n}}\}_{\mathbf{n}\in\mathbb{N}}$ es una sucesión de variables aleatorias independientes e idénticamente distribuidas que pueden ser discretas o continuas. En las aplicaciones, los elementos de la sucesión representan **n** observaciones independientes en una variable aleatoria, asociado con un fenómeno subyacente que es importante para para el investigador. En un entorno probabilístico/estadístico, el interés radica en conocer la distribución de la variable aleatoria X, ya sea que esté representada por las probabilidades $P(X \in B), B \subseteq \mathbb{R}$, por la función de distribución F de las variables aleatorias X_n o por su función de densidad o función de probabilidad f. En la práctica, esta distribución es desconocida por el investigador. Entonces sería deseable aproximar, en algún sentido, la distribución desconocida, por una distribución conocida. El tipo de concepto de convergencia que justifica este tipo de aproximación se llama *convergencia en la distribución* o *convergencia en la ley.* De todos los conceptos de convergencia que estamos discutiendo, la convergencia en la distribución es uno de los más útiles para responder preguntas prácticas. Por ejemplo, los estadísticos suelen estar mucho más interesados en construir intervalos de confianza y no solo en obtener estimadores puntuales, y es necesario un teorema central de límite de algún tipo para producir un intervalo de confianza. En esta sección, se dan las bases para tal aproximación.

Definición 1.13. Sean $X, X_1, X_2, \ldots$ variables aleatorias con funciones de distribución $F, F_1, F_2, \ldots$ respectivamente. Las variables aleatorias pueden ser discretas o continuas y no necesariamente sean independientes ni estar distribuidas en forma idéntica. $\{X_{\mathbf{n}}\}_{\mathbf{n}\geq 1}$ *converge en distribución o en ley* para la variable aleatoria X, si $F_{X_{\mathbf{n}}}(t) \to F_X(t)$ en todo punto t en la cual F_X es continua. (Recuerde que F_X es continua en t si y solo si $P(X = t) = 0$). Este es el modo de convergencia necesario para la aproximación de una distribución por otra.

Notaciones: $X_{\mathbf{n}} \xrightarrow{D} X$ o $X_{\mathbf{n}} \xrightarrow{D} F$ o $X_{\mathbf{n}} \xrightarrow{\mathscr{L}} X$

(Véase también la Figura 1.2)

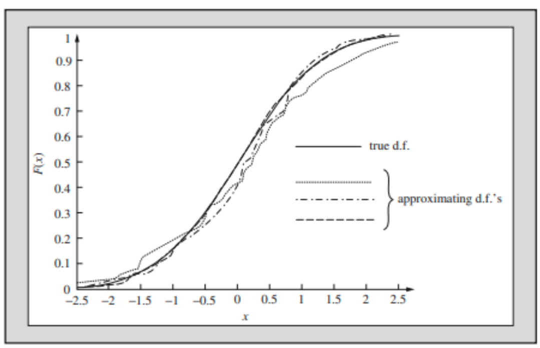

Figura 1.2: la función de distribución representada por la curva continua se aproxima mediante los gl representados por las curvas ...., .-. .-. .-. .-. y ——-.

El siguiente ejemplo ilustra la definición.

Ejemplo 1.31. Para $\mathbf{n} \geq 1$ sean las funciones de distribuciones $F_{\mathbf{n}}(x)$ y la función de distribución F dados por:

$$F_{\mathbf{n}}(x) = \begin{cases} 0 & \text{si} \quad x < 1 - \frac{1}{\mathbf{n}} \\ \frac{1}{2} & \text{si} \quad 1 - \frac{1}{\mathbf{n}} \leq x < 1 + \frac{1}{\mathbf{n}} \\ 1 & \text{si} \quad x \geq 1 + \frac{1}{\mathbf{n}} \end{cases}, \qquad F(x) = \begin{cases} 0 & \text{si} \quad x < 1 \\ 1 & \text{si} \quad x \geq 1 \end{cases}$$

Analicemos si $F_{\mathbf{n}}(x)$ converge o no a $F(x)$ cuando $\mathbf{n} \to \infty$. (Véase también la Figura 12.2.)

Análisis:

La función de distribución F es continua en todas partes excepto en el punto $x = 1$. Por ejemplo, para $x < 1$, sea $\mathbf{n}_0 > \frac{1}{1-x}$. Entonces

$$1 - x > \frac{1}{\mathbf{n}_0} \Rightarrow x < 1 - \frac{1}{n_0}$$

y por tanto $x < 1 - \frac{1}{\mathbf{n}} \forall \mathbf{n} \geq \mathbf{n}_0$. De esta manera, $F_{\mathbf{n}}(x) = 0$ para $\mathbf{n} \geq \mathbf{n}_0$.

Ahora, para x > 1, sea $\mathbf{n}_0 \geq \frac{1}{x-1} \Rightarrow x - 1 \geq \frac{1}{\mathbf{n}_0} \Rightarrow x \geq 1 + \frac{1}{\mathbf{n}_0}$ y $x \geq 1 + \frac{1}{\mathbf{n}}$ para todo $\mathbf{n} \geq \mathbf{n}_0$, de modo que $F_{\mathbf{n}}(x) = 1 \forall \mathbf{n} \geq \mathbf{n}_0$. De esta manera, para $x \neq 1$, $F_{\mathbf{n}} \to F$, Así que si $X_{\mathbf{n}}$ y X son variables aleatorias tales que $X_{\mathbf{n}} \sim F_{\mathbf{n}}$ y $X \sim F$, y $X_{\mathbf{n}} \xrightarrow[\mathbf{n}\to\infty]{D} X$.

Observación 7. El ejemplo también ilustra el punto de que si x es un punto de discontinuidad de F, entonces $F_{\mathbf{n}}(x)$ no necesita converger a $F(x)$. En el Ejemplo 1.31, $F_{\mathbf{n}}(1) = \frac{1}{2}$ para todo $\mathbf{n}$, y $F(1) = 1$.

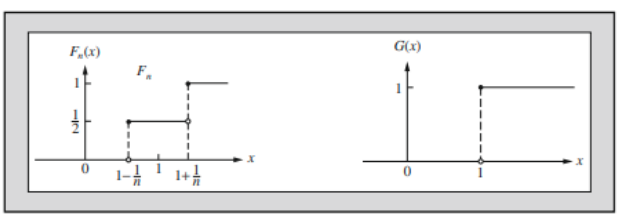

Figura 1.3: la fd F se aproxima por la funciones de distribuciones de $F_{\mathbf{n}}$ en todos los puntos $x \neq 1$.

La idea, detrás de la Definición 1.13 es la aproximación de la probabilidad (presumiblemente desconocida) $P(X \leq x) = F(x)$ por las probabilidades (presumiblemente conocidas) $P(X_{\mathbf{n}} \leq x) = F_{\mathbf{n}}(x)$, para $\mathbf{n}$ suficientemente grande. La convergencia en distribución también permite la aproximación de probabilidades de la forma $P(x < X \leq y)$ por las probabilidades $P(x < X_{\mathbf{n}} \leq y)$, para puntos de continuidad x e y de F con $x < y$. Esto es así debido aque:

$$\begin{aligned} P(x < X_{\mathbf{n}} \leq y) &= P(X_{\mathbf{n}} \leq y) - P(X_{\mathbf{n}} \leq x) \\ &= F_{\mathbf{n}}(y) - F_{\mathbf{n}}(x) \xrightarrow[\mathbf{n}\to\infty]{} F(y) - F(x) = P(x < X \leq y). \end{aligned}$$

Mientras que la convergencia en distribución permite la comparación de ciertas probabilidades, calculadas en términos de $X_{\mathbf{n}}$ y e X de las variables aleatorias individuales, no proporciona una evaluación de las probabilidades calculadas sobre el comportamiento conjunto de $X_{\mathbf{n}}$ y X.

Ejemplo 1.32. Sea $\mathbf{\Omega} = \{1, 2, 3, 4\}$, y sobre los subconjuntos de $\mathbf{\Omega}$, sea P la función de probabilidad uniforme discreta. Defina las siguientes variables aleatorias:

$$X_{\mathbf{n}}(1) = X_{\mathbf{n}}(2) = X_{\mathbf{n}}(3) = 1, \quad X_{\mathbf{n}}(3) = X_{\mathbf{n}}(4) = 0, \quad \mathbf{n} = 1, 2, \ldots$$

y

$$X(1) = X(2) = 0, \quad X_(3) = X_(4) = 1.$$

Análisis: Entonces: $|X_{\mathbf{n}}(\omega) - X(\omega)| = 1$ para todo $\omega \in \mathbf{\Omega}$.

Por lo tanto, $X_{\mathbf{n}}$ no converge en probabilidad a X, cuando $\mathbf{n} \to \infty$. Ahora,

$$F_{\mathbf{n}}(x) = \begin{cases} 0 & \text{si} \quad x < 0 \\ \frac{1}{2} & \text{si} \quad 0 \leq x < 1 \\ 1 & \text{si} \quad x \geq 1 \end{cases}, \qquad F(x) = \begin{cases} 0 & \text{si} \quad x < 0 \\ \frac{1}{2} & \text{si} \quad 0 \leq x < 1 \\ 1 & \text{si} \quad x \geq 1 \end{cases}$$

de modo que $F_{\mathbf{n}}(x) = F(x)$ para todo $x \in \mathbb{R}$. Así, trivialmente, $F_{\mathbf{n}}(x) \xrightarrow[\mathbf{n}\to\infty]{} F(x)$ para todos los puntos de continuidad de F; es decir, $X_{\mathbf{n}} \xrightarrow[\mathbf{n}\to\infty]{D} X$, pero $X_{\mathbf{n}}$ no converge en probabilidad a X.

Ejemplo 1.33. Sea $X_{\mathbf{n}} \sim \mathscr{N}\left(0, \frac{1}{\mathbf{n}}\right)$. Entonces la función de distribución de $X_{\mathbf{n}}$ es

$$\begin{aligned} F_{X_{\mathbf{n}}}(x) = P(X_{\mathbf{n}} \leq x) &= \int_{-\infty}^{x} \frac{e^{-\frac{1}{2}\mathbf{n}t^2}}{\sqrt{\frac{2\pi}{\mathbf{n}}}} dt \\ &= \int_{-\infty}^{\sqrt{\mathbf{n}}x} \frac{e^{-\frac{1}{2}s^2}}{\sqrt{2\pi}} ds \\ &= \Phi(\sqrt{\mathbf{n}}x) \end{aligned}$$

Luego,

$$F_{X_n}(x) \to \begin{cases} \Phi(\infty) = 1 & \text{si} \quad x > 0 \\ \Phi(0) = \frac{1}{2} & \text{si} \quad x = 0 \\ \Phi(-\infty) = 0 & \text{si} \quad x < 0 \end{cases}$$

Si

$$F_X(x) = \begin{cases} 1 & \text{si} \quad x \geq 0 \\ 0 & \text{si} \quad x < 0 \end{cases}$$

el único punto de discontinuidad está en $x = 0$. En todos lados, $\Phi(\sqrt{\mathbf{n}}x) = F_{X_\mathbf{n}}(x) \to F_X(x) = \Phi(z)$, donde $\Phi(z) = P(Z \leq z)$ con $Z \sim \mathcal{N}(0,1)$.

Por lo tanto $X_\mathbf{n} \xrightarrow{d} X$, donde $P(X = 0) = 1$, o $X_\mathbf{n} \xrightarrow{d} 0$ ya que la variable aleatoria aquí es degenerado.

Ejemplo 1.34. Suponga que,

$$X_\mathbf{n} \sim U\left[\frac{1}{2} - \frac{1}{\mathbf{n}+1}; \frac{1}{2} + \frac{1}{\mathbf{n}+1}\right], \mathbf{n} \geq 1.$$

Debido a que el intervalo $\left[\frac{1}{2} - \frac{1}{\mathbf{n}+1}, \frac{1}{2} + \frac{1}{\mathbf{n}+1}\right]$ se reduce al único punto de $\frac{1}{2}$, intuitivamente sentimos que la distribución de $X_\mathbf{n}$ se acerca a una distribución concentrada en $\frac{1}{2}$, es decir, una *distribución de un punto*. La función de distribución F de la distribución concentrada en $\frac{1}{2}$ es igual a la función $F(x) = 0$ para $x < 0$, y $F(x) = 1$ para $x \geq \frac{1}{2}$. Ahora consideremos la función de distribución de $X_\mathbf{n}$; denotemos por $F_\mathbf{n}(x)$. Sea $x < \frac{1}{2}$ fijo. Entonces, cuando $\mathbf{n} \to \infty$, $F_\mathbf{n}(x) = 0$, y por lo tanto $\lim_{\mathbf{n}\to\infty} F_\mathbf{n}(x) = 0$. A continuación, fije $x > \frac{1}{2}$. Entonces, cuando $\mathbf{n} \to \infty$ $F_\mathbf{n}(x) = 1$, y por tanto $\lim_{\mathbf{n}\to\infty} F_\mathbf{n}(x) = 1$.
En consecuencia
Como resultado
Por lo tanto, si $x < \frac{1}{2}$, o si $x > \frac{1}{2}$, el $\lim_{\mathbf{n}\to\infty} F_\mathbf{n}(x) = F(x)$. Si x es exactamente igual a $\frac{1}{2}$, entonces $F_\mathbf{n}(x) = \frac{1}{2}$. Pero, $F\left(\frac{1}{2}\right) = 1$. Entonces $x = \frac{1}{2}$ es un punto problemático, y el único punto problemático, en que $F_\mathbf{n}\left(\frac{1}{2}\right) \nrightarrow F\left(\frac{1}{2}\right)$. Curiosamente, $x = \frac{1}{2}$ también es exactamente el único punto en el que F no es continua. Sin embargo, no queremos que este punto problemático arruine nuestra sensación intuitiva de que $X_\mathbf{n}$ se acerca a la distribución de un punto concentrada en $\frac{1}{2}$. Es decir, no tenemos en cuenta ningún punto en el que la función de distribución límite no sea continua.
continúa

La mayor parte de la probabilidad puede verse como variaciones, extensiones y generalizaciones de dos resultados básicos, el **teorema central del límite** y la **ley de los grandes números**. Ambos teoremas tratan del comportamiento límite de sucesión de variables aleatorias. Las nociones de límite involucradas son el tema de esta sección. Todos los límites en la sección son cuando $\mathbf{n} \to \infty$.

La relación precisa entre convergencia en distribución y convergencia en probabilidad se establece en el siguiente teorema.

Proposición 1.4. Si $X_{\mathbf{n}} \xrightarrow[\mathbf{n}\to\infty]{P} X \Rightarrow X_{\mathbf{n}} \xrightarrow[\mathbf{n}\to\infty]{D} X$. El recíproco también es verdadero si X es una variable aleatoria *degenerada*; esto es, $P(X = c) = 1$ para alguna constante c.

Prueba. Sea $x \in \mathbb{R}$ un punto de continuidad de F y sea $\boldsymbol{\epsilon} > 0$ dado. Entonces se tiene el evento

$$\begin{aligned}
[X \leq x - \boldsymbol{\epsilon}] &= \Big\{[X_{\mathbf{n}} \leq x] \cap [X \leq x - \boldsymbol{\epsilon}]\Big\} \cup \Big\{[X_{\mathbf{n}} > x] \cap [X \leq x - \boldsymbol{\epsilon}]\Big\} \\
&\subseteq [X_{\mathbf{n}} \leq x] \cup \Big\{[X_{\mathbf{n}} > x] \cap [X \leq x - \boldsymbol{\epsilon}]\Big\} \\
&\subseteq [X_{\mathbf{n}} \leq x] \cup [|X_{\mathbf{n}} - X| \geq \boldsymbol{\epsilon}]
\end{aligned}$$

puesto que,

$$\begin{aligned}
[X_{\mathbf{n}} > x] \cap [X \leq x - \boldsymbol{\epsilon}] &= [X_{\mathbf{n}} > x] \cap [-X \geq -x + \boldsymbol{\epsilon}] \\
&\subseteq [X_{\mathbf{n}} - X \geq \boldsymbol{\epsilon}] \subseteq [|X_{\mathbf{n}} - X| \geq \boldsymbol{\epsilon}].
\end{aligned}$$

Por lo tanto

$$[X \leq x - \boldsymbol{\epsilon}] \subseteq [X_{\mathbf{n}} \leq x] \cup [|X_{\mathbf{n}} - X| \geq \boldsymbol{\epsilon}]$$

de donde se tiene

$$P(X \leq x - \boldsymbol{\epsilon}) \leq P(X_{\mathbf{n}} \leq x) + P(|X_{\mathbf{n}} - X| \geq \boldsymbol{\epsilon}),$$

o

$$F(x - \boldsymbol{\epsilon}) \leq F_{\mathbf{n}}(x) + P(|X_{\mathbf{n}} - X| \geq \boldsymbol{\epsilon}).$$

Así, si $X_{\mathbf{n}} \xrightarrow[\mathbf{n}\to\infty]{P} X$, cuando se toman límites se obtiene,

$$F(x - \boldsymbol{\epsilon}) \leq \liminf_{\mathbf{n}\to\infty} F_{\mathbf{n}}(x). \tag{15}$$

De una manera similar se tiene,

$$[X_{\mathbf{n}} \leq x] = \Big\{[X_{\mathbf{n}} \leq x] \cap [X \leq x + \boldsymbol{\epsilon}]\Big\} \cup \Big\{[X_{\mathbf{n}} \leq x] \cap [X \geq X + \boldsymbol{\epsilon}]\Big\}$$
$$\subseteq [X \leq x + \boldsymbol{\epsilon}] \cup \Big\{[X_{\mathbf{n}} \leq x] \cap [X \geq X + \boldsymbol{\epsilon}]\Big\}$$
$$[X_{\mathbf{n}} \leq -x] \cap [X \geq X + \boldsymbol{\epsilon}] = [-X_{\mathbf{n}} \geq x] \cap [X \geq x + \boldsymbol{\epsilon}]$$
$$\subseteq [X - X_{\mathbf{n}} \geq \boldsymbol{\epsilon}] \subseteq [|X_{\mathbf{n}} - X| \geq \boldsymbol{\epsilon}].$$

Por lo tanto

$$[X_{\mathbf{n}} \leq x] \subseteq [X \leq x + \boldsymbol{\epsilon}] \cup [|X_{\mathbf{n}} - X| \geq \boldsymbol{\epsilon}]$$

de donde se tiene

$$P(X_{\mathbf{n}} \leq x) \leq P(X \leq x + \boldsymbol{\epsilon}) + P(|X_{\mathbf{n}} - X| \geq \boldsymbol{\epsilon})$$

o equivalentemente

$$F_{\mathbf{n}}(x) \leq F(x + \boldsymbol{\epsilon}) + P(|X_{\mathbf{n}} - X| \geq \boldsymbol{\epsilon}).$$

Así, si $X_{\mathbf{n}} \xrightarrow[\mathbf{n}\to\infty]{P} X$, cuando se toman los límites se obtiene,

$$F(x + \boldsymbol{\epsilon}) \leq \limsup_{\mathbf{n}\to\infty} F_{\mathbf{n}}(x). \tag{16}$$

De la relación (15) y (16) resulta,

$$F(x - \boldsymbol{\epsilon}) \leq \liminf_{\mathbf{n}\to\infty} F_{\mathbf{n}}(x) \leq \limsup_{\mathbf{n}\to\infty} F_{\mathbf{n}}(x) \leq F(x + \boldsymbol{\epsilon}).$$

hacia

Haciendo tender $\boldsymbol{\epsilon}$ hacía 0, se obtiene (por el hecho de que x es un punto de continuidad de F) que

$$F(x) \leq \liminf_{\mathbf{n}\to\infty} F_{\mathbf{n}}(x) \leq \limsup_{\mathbf{n}\to\infty} F_{\mathbf{n}}(x) \leq F(x).$$

En consecuencia $\lim_{\mathbf{n}\to\infty} F_{\mathbf{n}}(x)$ existe y es igual a $F(x)$ □

Proposición 1.5. Si $X_{\mathbf{n}} \xrightarrow{D} c$ constante, entonces $X_{\mathbf{n}} \xrightarrow{P} c$.

Prueba. Asumamos ahora que $P(X = c) = 1$. Entonces, la función de distribución de una variable aleatoria constante c es

$$F_c(x) = \begin{cases} 1 & \text{si} \quad x \geq c \\ 0 & \text{si} \quad x < c \end{cases}$$

y por hipótesis, $F_{\mathbf{n}}(x) \xrightarrow[\mathbf{n}\to\infty]{} F(x), x \neq c$. Se debe demostrar que $X_{\mathbf{n}} \xrightarrow[\mathbf{n}\to\infty]{P} c$. Se tiene

$$\begin{aligned}
\mathbf{P}(|X_{\mathbf{n}} - c| \leq c) &= \mathbf{P}(-\epsilon \leq X_{\mathbf{n}} - c \leq \epsilon) \\
&= \mathbf{P}(c - \epsilon \leq X_{\mathbf{n}} \leq c + \epsilon) \\
&= \mathbf{P}(X_{\mathbf{n}} \leq c + \epsilon) - \mathbf{P}(X_{\mathbf{n}} < c - \epsilon) \\
&= F_{\mathbf{n}}(c + \epsilon) - F_{\mathbf{n}}(c - \epsilon)
\end{aligned}$$

Como $c - \epsilon, c + \epsilon$ son puntos de continuidad de F, tomando límite se obtiene

$$\begin{aligned}
\lim_{\mathbf{n}\to\infty} \mathbf{P}(|X_{\mathbf{n}} - c| \leq \epsilon) &= \lim_{\mathbf{n}\to\infty} F_{\mathbf{n}}(c + \epsilon) - \lim_{\mathbf{n}\to\infty} F_{\mathbf{n}}(c - \epsilon) \\
&= 1 - 0 = 1
\end{aligned}$$

Por loa tanto

$$\lim_{\mathbf{n}\to\infty} \mathbf{P}(|X_{\mathbf{n}} - c| \leq \epsilon) \xrightarrow[\mathbf{n}\to\infty]{} 1. \quad \text{o} \quad X_{\mathbf{n}} \xrightarrow{D} c.$$

□

Antes de establecer las relaciones entre los diferentes tipos de convergencia, damos un resultado pequeño sorprendentemente útil que caracteriza la convergencia en probabilidad.

Teorema 1.18.

$$X_{\mathbf{n}} \xrightarrow{P} X \Leftrightarrow \lim_{\mathbf{n}\to\infty} \mathbf{E}\left(\frac{|X_{\mathbf{n}} - X|}{1 + |X_{\mathbf{n}} - X|}\right) = 0$$

Prueba. Sin pérdida de generalidad asumimos $X = 0$. Por consiguiente se desea demostrar que $X_{\mathbf{n}} \xrightarrow{P} X \Leftrightarrow \lim_{\mathbf{n}\to\infty} \mathbf{E}\left(\frac{|X_{\mathbf{n}}|}{1+|X_{\mathbf{n}}|}\right) = 0$.

Primero supongamos que $X_{\mathbf{n}} \xrightarrow{P} 0$. Entonces por definición de convergencia en probabilidad se tiene, $\forall \epsilon > 0$, $\lim_{\mathbf{n}\to\infty} \mathbf{P}(|X_{\mathbf{n}}| > \epsilon) = 0$. definamos la función indicadora como sigue:

Definamos

$$I(|X_{\mathbf{n}}| > \epsilon) = \begin{cases} 1 & \text{si} \quad |X_{\mathbf{n}}| > \epsilon \\ 0 & \text{si} \quad |X_{\mathbf{n}}| \leq \epsilon \end{cases}$$

Luego se tiene

$$\frac{|X_{\mathbf{n}}|}{1 + |X_{\mathbf{n}}|} \leq \frac{|X_{\mathbf{n}}|}{1 + |X_{\mathbf{n}}|} I(|X_{\mathbf{n}}| > \epsilon) + \epsilon I(|X_{\mathbf{n}}| \leq \epsilon) \leq I(|X_{\mathbf{n}}| > \epsilon) + \epsilon$$

Por tanto, tomando la esperanza a los extremos de la expresión anterior se tiene

$$\mathbf{E}\left[\frac{|X_\mathbf{n}|}{1+|X_\mathbf{n}|}\right] \leq \mathbf{E}\left[I(|X_\mathbf{n}| > \epsilon)\right] + \epsilon$$

Tomando los límites resulta

$$\lim_{\mathbf{n}\to\infty} \mathbf{E}\left[\frac{|X_\mathbf{n}|}{1+|X_\mathbf{n}|}\right] \leq \lim_{\mathbf{n}\to\infty} \mathbf{P}(|X_\mathbf{n}| > \epsilon) + \epsilon$$

$$\lim_{\mathbf{n}\to\infty} \mathbf{E}\left[\frac{|X_\mathbf{n}|}{1+|X_\mathbf{n}|}\right] \leq \epsilon;$$

como ϵ es arbitrario resulta que $\lim_{\mathbf{n}\to\infty} \mathbf{E}\left[\frac{|X_\mathbf{n}|}{1+|X_\mathbf{n}|}\right] = 0$.

Ahora, supongamos que $\lim_{\mathbf{n}\to\infty} \mathbf{E}\left[\frac{|X_\mathbf{n}|}{1+|X_\mathbf{n}|}\right] = 0$. La función $f(x) = \frac{x}{1+x}$ tiene derivada $f'(x) = \frac{1}{(1+x)^2} > 0$ y por tanto es una función estrictamente creciente. Luego

$$\frac{\epsilon}{1+\epsilon} I(|X_\mathbf{n}| > \epsilon) \leq \frac{|X_\mathbf{n}|}{1+|X_\mathbf{n}|} I(|X_\mathbf{n}| > \epsilon) \leq \frac{|X_\mathbf{n}|}{1+|X_\mathbf{n}|}$$

Tomando la esperanza y luego el límite resulta

$$\frac{\epsilon}{1+\epsilon} \lim_{\mathbf{n}\to\infty} \mathbf{P}(|X_\mathbf{n}| > \epsilon) \leq \lim_{\mathbf{n}\to\infty} \mathbf{E}\left[\frac{|X_\mathbf{n}|}{1+|X_\mathbf{n}|}\right] = 0.$$

Como $\epsilon > 0$ es fijo y arbitrario, se concluye que $\lim_{\mathbf{n}\to\infty} \mathbf{P}(|X_\mathbf{n}| > \epsilon) = 0$. □

Observación 8. Este Teorema afirma $X_\mathbf{n} \xrightarrow{P} X \Leftrightarrow \mathbf{E}(f(|X_\mathbf{n} - X|)) \to 0$ para la función $f(x) = \frac{x}{1+x}$. Un examen cuidadoso de la prueba demuestra que la misma equivalencia se cumple para cualquier función f en $\mathbb{R}^+$ que sea limitada, estrictamente creciente en el intervalo $(0, 1)$, continua y con $f(0) = 0$. Por ejemplo, se tiene $X_\mathbf{n} \xrightarrow{P} X \Leftrightarrow \mathbf{E}(|X_\mathbf{n} - X| \wedge 1) \to 0$ y también si y solo si $\mathbf{E}[\arctan(|X_\mathbf{n} - X|)] \to 0$.

Teorema 1.19. Sea $\{X_\mathbf{n}\}_{\mathbf{n}\geq 1}$ una sucesión de variables aleatorias.

(a) Si $\mathbf{X_n} \xrightarrow{L^p} X$, entonces $X_\mathbf{n} \xrightarrow{P} X$.

(b) Si $\mathbf{X_n} \xrightarrow{c.s.} X$, entonces $X_\mathbf{n} \xrightarrow{P} X$.

Prueba. (a) Como $X_\mathbf{n} \xrightarrow{L^p} X$, entonces $\lim_{\mathbf{n}\to\infty} \mathbf{E}(|X_\mathbf{n} - X|^p) = 0$. Por otra parte se sabe que para un evento A, $\mathbf{P}(A) = \mathbf{E}[I(A)]$, donde $I(A)$ es la función indicadora del evento A. Por tanto,

$$\mathbf{P}(|X_{\mathbf{n}} - X| > \boldsymbol{\epsilon}) = \mathbf{E}[I(|X_{\mathbf{n}} - X| > \boldsymbol{\epsilon})].$$

Note que $\frac{|X_{\mathbf{n}} - X|^p}{\boldsymbol{\epsilon}^p} > 1$ en el evento $[|X_{\mathbf{n}} - X| > \boldsymbol{\epsilon}]$, en consecuencia

$$I(|X_{\mathbf{n}} - X| > \boldsymbol{\epsilon}) \leq \frac{|X_{\mathbf{n}} - X|^p}{\boldsymbol{\epsilon}^p} I(|X_{\mathbf{n}} - X| > \boldsymbol{\epsilon})$$

Tomando la esperanza a los dos miembros de la expresión anterior se obtiene

$$\mathbf{E}[I(|X_{\mathbf{n}} - X| > \boldsymbol{\epsilon})] \leq \mathbf{E}\left[\frac{|X_{\mathbf{n}} - X|^p}{\boldsymbol{\epsilon}^p} I(|X_{\mathbf{n}} - X| > \boldsymbol{\epsilon})\right]$$

y como $|X_{\mathbf{n}} - X|^p \geq 0$ siempre, se puede omitir la función indicadora en la expresión anterior, lo que resulta

$$\mathbf{P}(|X_{\mathbf{n}} - X| > \boldsymbol{\epsilon}) \leq \frac{1}{\boldsymbol{\epsilon}^p} \mathbf{E}\left[|X_{\mathbf{n}} - X|^p\right].$$

Luego, tomando el límite a ambos lados de la desigualdad se obtiene

$$\lim_{\mathbf{n} \to \infty} \mathbf{P}(|X_{\mathbf{n}} - X| > \boldsymbol{\epsilon}) \leq \frac{1}{\boldsymbol{\epsilon}^p} \lim_{\mathbf{n} \to \infty} \mathbf{E}\left[|X_{\mathbf{n}} - X|^p\right]$$
$$\lim_{\mathbf{n} \to \infty} \mathbf{P}(|X_{\mathbf{n}} - X| > \boldsymbol{\epsilon}) = 0$$

puesto que $\boldsymbol{\epsilon} > 0$ fijo y arbitrario.

(b) Como $\frac{|X_{\mathbf{n}} - X|}{1 + |X_{\mathbf{n}} - X|} \leq 1$ siempre, se tiene

$$\lim_{\mathbf{n} \to \infty} \mathbf{E}\left[\frac{|X_{\mathbf{n}} - X|}{1 + |X_{\mathbf{n}} - X|}\right] = \mathbf{E}\left[\lim_{\mathbf{n} \to \infty} \frac{|X_{\mathbf{n}} - X|}{1 + |X_{\mathbf{n}} - X|}\right] = \mathbf{E}(0) = 0$$

por la convergencia dominada de Lebesgue. Luego, aplicando el Teorema 1.18 se completa la prueba.

$\square$

Teorema 1.20. Sea $1 \leq r \leq s$. Si $X_{\mathbf{n}} \xrightarrow{L^s} X$, entonces $X_{\mathbf{n}} \xrightarrow{L^r} X$.

Prueba. Usamos la desigualdad de Hölder. La desigualdad de Hölder afirma que

$$\mathbf{E}|XY| \leq \left[\mathbf{E}|X|^p\right]^{\frac{1}{p}} \left[|Y|^q\right]^{\frac{1}{q}},$$

donde $1 \leq p, q < \infty$ y $\frac{1}{p} + \frac{1}{q} = 1$. En la desigualdad de Hölder, escogemos

$$X = |X_{\mathbf{n}} - X|^r, \qquad Y = 1, \qquad p = \tfrac{s}{r} > 1.$$

Con estas elecciones obtenemos

$$\mathbf{E}|X_{\mathbf{n}} - X|^r \leq \mathbf{E}\left(|X_{\mathbf{n}} - X|^r\right)^{\frac{1}{p}} = \mathbf{E}\left(|X_{\mathbf{n}} - X|^r\right)^{\frac{s}{r}} = \mathbf{E}\left(|X_{\mathbf{n}} - X|\right)^s$$

Ahora, como por hipótesis $X_{\mathbf{n}} \xrightarrow{L^s} X$, lo que significa que

$$\lim_{\mathbf{n}\to\infty} \mathbf{E}\left(|X_{\mathbf{n}} - X|^s\right) = 0.$$

Entonces se tiene que,

$$\lim_{\mathbf{n}\to\infty} \mathbf{E}\left(|X_{\mathbf{n}} - X|^r\right) \leq \lim_{\mathbf{n}\to\infty} \left(\mathbf{E}|X_{\mathbf{n}} - X|^s\right) = 0.$$

Por lo tanto, $X_{\mathbf{n}} \xrightarrow{L^r} X$ $\square$

Tal

Ejemplo 1.35. Sea una sucesión,$X_1, X_2, \ldots$ tal que

$$X_{\mathbf{n}} = \begin{cases} \mathbf{n}^2 & \text{con} \quad \mathbf{P}(X_{\mathbf{n}} = \mathbf{n}^2) = \frac{1}{\mathbf{n}} \\ 0 & \text{con} \quad \mathbf{P}(X_{\mathbf{n}} = 0) = 1 - \frac{1}{\mathbf{n}} \end{cases}$$

Pruebe que:

(a) $X_{\mathbf{n}} \xrightarrow{P} 0$.

(b) $X_{\mathbf{n}}$ no converge en la r-ésima media para cualquier $r \geq 1$.

Solución. (a) Para probar que $X_{\mathbf{n}} \xrightarrow{P} 0$, utilizamos la definición,

$$\begin{aligned} \forall \boldsymbol{\epsilon} > 0, \lim_{\mathbf{n}\to\infty} \mathbf{P}(|X_{\mathbf{n}} - 0| \geq \boldsymbol{\epsilon}) &= \lim_{\mathbf{n}\to\infty} \mathbf{P}(X_{\mathbf{n}} \geq \boldsymbol{\epsilon}) \\ &= \lim_{\mathbf{n}\to\infty} \mathbf{P}(X_{\mathbf{n}} = \mathbf{n}^2) \\ &= \lim_{\mathbf{n}\to\infty} \frac{1}{\mathbf{n}} = 0 \end{aligned}$$

Por lo tanto, $X_{\mathbf{n}} \xrightarrow{P} 0$.

(b) Para cualquier $r \geq 1$, se tiene

$$\begin{aligned}
\lim_{\mathbf{n}\to\infty} \mathbf{E}\left(|X_{\mathbf{n}}|^r\right) &= \lim_{\mathbf{n}\to\infty} \left[\mathbf{n}^{2r}\mathbf{P}(X_{\mathbf{n}} = \mathbf{n}^2) + 0\mathbf{P}(X_{\mathbf{n}} = 0)\right] \\
&= \lim_{\mathbf{n}\to\infty} \left[\mathbf{n}^{2r}\frac{1}{\mathbf{n}}\right] \\
&= \lim_{\mathbf{n}\to\infty} \mathbf{n}^{2r-1} = \infty \quad (\text{ puesto que } r \geq 1).
\end{aligned}$$

Por lo tanto, $X_{\mathbf{n}}$ no converge en la r-ésima media para cualquier $r \geq 1$. En particular, es interesante observar que, aunque $X_{\mathbf{n}} \xrightarrow{P} 0$, el valor esperado de $X_{\mathbf{n}}$ no converge a 0; esto es, aunque $X_{\mathbf{n}} \xrightarrow{P} 0 \nrightarrow \mathbf{E}(X_n) \to 0$. □

Observación 9. El hecho de que si una sucesión, $\{X_{\mathbf{n}}\}_{\mathbf{n}\in\mathbb{N}}$ de variables aleatorias converge en distribución, entonces la sucesión debe ser $O_p(1)$, nos dice que debe haber sucesiones de variables aleatorias que no convergen en distribución a nada. Por ejemplo, tomemos la variable aleatoria $X\mathbf{n}$, tal que $X_{\mathbf{n}} \sim \mathscr{N}(\mathbf{n}, 1), \mathbf{n} \geq 1$. Esta sucesión $X_{\mathbf{n}}$ no es $O_p(1)$, y por tanto no puede converger en distribución. La cuestión que se plantea es si la propiedad $O_p(1)$ es suficiente para la convergencia. Incluso eso, evidentemente, no es cierto; basta con considerar que $X_{2\mathbf{n}-1} \sim \mathscr{N}(0, 1)$, y $X_{2\mathbf{n}} \sim \mathscr{N}(1, 1)$. Sin embargo, subsucesiones por separado, las sub-sucesiones pares e impares sí convergen. Es decir, puede haber una inversión parcial del hecho de que si una sucesión $X\mathbf{n}$ converge en distribución, entonces debe ser $O_p(1)$. Este es un famoso teorema sobre la convergencia en distribución, y se enuncia a continuación.

Ejemplo 1.36. Sea $\{X_{\mathbf{n}}\}_{\mathbf{n}\in\mathbb{N}}$ una sucesión de variables aleatorias tal que,

$$X_{\mathbf{n}} \sim \mathbf{Binom}\left(\mathbf{n}, \tfrac{\lambda}{\mathbf{n}}\right), \text{para } \mathbf{n} \in N,\ \mathbf{n} > \lambda,$$

donde $\lambda > 0$ es una constante. Demuestre que $X_{\mathbf{n}}$ converge en distribución a Poisson(λ).

Solución. Se debe probar, $\lim_{\mathbf{n}\to\infty} \mathbf{P}_{X_\mathbf{n}}(k) = \mathbf{P}_X(k)$, $\forall\ k = 0, 1, 2, \ldots$. Se tiene

$$\begin{aligned}
\lim_{\mathbf{n}\to\infty} \mathbf{P}_{X_\mathbf{n}}(k) &= \lim_{\mathbf{n}\to\infty} \binom{\mathbf{n}}{k} \left(\frac{\lambda}{\mathbf{n}}\right)^k \left(1-\frac{\lambda}{\mathbf{n}}\right)^{\mathbf{n}-k} \\
&= \lambda^k \lim_{\mathbf{n}\to\infty} \binom{\mathbf{n}}{k} \left(\frac{1}{\mathbf{n}^k}\right) \left(1-\frac{\lambda}{\mathbf{n}}\right)^{\mathbf{n}-k} \\
&= \lambda^k \lim_{\mathbf{n}\to\infty} \frac{\mathbf{n}!}{k!(\mathbf{n}-k)!} \left(\frac{1}{\mathbf{n}^k}\right) \left(1-\frac{\lambda}{\mathbf{n}}\right)^{\mathbf{n}-k} \\
&= \frac{\lambda^k}{k!} \lim_{\mathbf{n}\to\infty} \left(\left[\frac{\mathbf{n}(\mathbf{n}-1)\ldots(\mathbf{n}-k+1)}{\mathbf{n}^k}\right] \left(1-\frac{\lambda}{\mathbf{n}}\right)^{\mathbf{n}} \left(1-\frac{\lambda}{\mathbf{n}}\right)^{-k} \right).
\end{aligned}$$

Tenga en cuenta que para un k fijo, resulta

$$\begin{aligned}
\lim_{\mathbf{n}\to\infty} \frac{\mathbf{n}(\mathbf{n}-1)\ldots(\mathbf{n}-k+1)}{\mathbf{n}^k} &= 1 \\
\lim_{\mathbf{n}\to\infty} \left(1-\frac{\lambda}{\mathbf{n}}\right)^{-k} &= 1 \\
\lim_{\mathbf{n}\to\infty} \left(1-\frac{\lambda}{\mathbf{n}}\right)^{\mathbf{n}} &= e^{-\lambda}
\end{aligned}$$

Así, concluimos que

$$\lim_{\mathbf{n}\to\infty} \mathbf{P}_{X_\mathbf{n}}(k) = \mathbf{P}_X(k) = \frac{e^{-\lambda}\lambda^k}{k!}$$

$\square$

Teorema 1.21. (Teorema de Helly). Sea $\{X_\mathbf{n}\}_{\mathbf{n}\in\mathbb{N}}$ una sucesión de variables aleatorias definidas en un espacio muestral común $\mathbf{\Omega}$, y supongamos que $X\mathbf{n}$ es $O_p(1)$. Entonces existe una sub-sucesión $X_{\mathbf{n}_j}, j \geq 1$, y una variable aleatoria
subsucesión
X (en el mismo espacio muestral $\mathbf{\Omega}$), tal que $X_{\mathbf{n}_j} \xrightarrow{D} X$. Además, $X_\mathbf{n} \xrightarrow{D} X$ si y sólo si toda sub-sucesión convergente $X_{\mathbf{n}_j}$, converge en distribución a esta
si y solo si subsucesión
misma variable aleatoria X.

Véase la demostración en Port (1994, p. 625). Se conocen importantes generalizaciones del teorema de Helly a espacios mucho más generales. Por lo general, en estos resultados se supone algún tipo de estructura métrica. E *en estos resultados; véase van der Vaart y Wellner (2000) para tales resultados generales sobre la compacidad débil. resultados generales sobre la compacidad débil.*
R

Ejemplo 1.37. (Son posibles varios fenómenos de convergencia:Anirban DasGupta). Este ejemplo rápido demuestra que una sucesión de distribuciones discretas puede converger en distribución a una distribución discreta, o a una distribución continua, y una secesión de distribuciones continuas puede converger en distribución a una continua, o a una discreta.

Un buen ejemplo de variables aleatorias discretas que convergen en distribución a una variable aleatoria discreta es la sucesión $X\mathbf{n} \sim \mathbf{Binom}\left(\mathbf{n}, \frac{\lambda}{\mathbf{n}}\right)$. Hemos visto en el 1.36 $X_\mathbf{n}$ converge a una variable aleatoria X que se distribuye en forma de Poisson con media λ. Un ejemplo familiar de variables aleatorias discretas que convergen en distribución a una variable aleatoria continua es el teorema central de límite de Moivre-Laplace, que dice que si $X_\mathbf{n} \sim \mathbf{Binom}\,(\mathbf{n}, \theta)$, entonces $\frac{X_\mathbf{n}-\mathbf{n}\theta}{\sqrt{\mathbf{n}\theta(1-\theta)}}$ converge a una variable normal estándar.

Los ejemplos de variables aleatorias continuas que convergen a una variable aleatoria continua están inmediatamente disponibles usando el teorema central del límite. Por ejemplo, si X_i son variables aleatorias independientes distribuidas en forma uniforma en el intervalo $[-1, 1]$, entonces $\frac{\sqrt{\mathbf{n}}\bar{X}}{\sigma}$ donde $\sigma^2 = \frac{1}{3}$, converge a una variable aleatoria normal estándar. Finalmente, como ejemplo de variables aleatorias continuas que convergen en una variable aleatoria discreta, considere que la sucesión de variables aleatorias $X_\mathbf{n} \sim \mathbf{Beta}\left(\frac{\mathbf{1}}{\mathbf{n}}, \frac{\mathbf{1}}{\mathbf{n}}\right)$. Visualmente, la función de densidad densidad de la variable aleatoria $X_\mathbf{n}$ para $\mathbf{n}$ grande es una densidad en forma de U simétrica, ilimitada tanto en 0 como en 1. La variable $X_\mathbf{n}$ converge en distribución a X, donde X es una variable aleatoria de Bernoulli con parámetro $\frac{1}{2}$.▲

Por definición de convergencia en la distribución, si $X_\mathbf{n} \xrightarrow{D} X$, y si X tiene una función de distribución (fd) continua F (continua en todas partes), entonces $F_\mathbf{n}(x) \to F(x)\forall x$ donde $F_\mathbf{n}(x)$ es la fd de $\mathbf{n}$. El siguiente teorema dice que mucho más es cierto, que la convergencia es realmente uniforme; ver pág. 265 en Chow y Teicher (1988).

Teorema 1.22. (Teorema de Pólya). Sea $\{X_\mathbf{n}\}_{\mathbf{n}\in\mathbb{N}}$ una sucesión de variables aleatorias con función de distribución $F_\mathbf{n}$ y X tiene función de distribución F.