TEMA CERO

SUCESIONES DE VARIABLES ALEATORIAS Y CONVERGENCIA

ESTADISTICA II - TEMA 0

* Sucesiones de variables aleatorias

Una sucesión de variables aleatorias $\{X_i\}_{i\in\mathbb{N}}$ es un conjunto infinito numerable $\{X_1, X_2, \dots X_n, \dots\}$ de variables aleatorias.

Por ejemplo, la sucesión formada por la media muestral:

$$\overline{X}_n = \frac{\sum_{i=1}^{n} X_i}{n}$$ es una sucesión de variables aleatorias

$$\overline{X}_1 = X_1 \,;\; \overline{X}_2 = \frac{X_1+X_2}{2}, \dots$$

* Convergencia de sucesiones de variables aleatorias

Queremos ver el comportamiento que tiene la sucesión de variables aleatorias $X_1, X_2, \dots X_n, \dots$ cuando n tiende a $+\infty$, en esto consiste la convergencia.

Hay tres tipos importantes de convergencia

* Convergencia en ley o distribución

Sea $\{X_n\}_{n\in\mathbb{N}}$ una sucesión de variables aleatorias con función de distribución $\{F_n(x)\}_{n\in\mathbb{N}}$ (esto quiere decir que la función de distribución de X_n es $F_n(x)$)

Sea X una variable aleatoria con función de distribución $F(x)$. Decimos que $\{X_n\}_{n\in\mathbb{N}}$ converge en ley o distribución a X y se escribe $\{X_n\}_{n\in\mathbb{N}} \xrightarrow{L} X$ si y sólo si $\lim_{n\to\infty} F_n(x) = F(x)$

El hecho de que exista el límite no asegura la convergencia, hay que comprobar que efectivamente $F(x)$ es función de distribución.

Veamos un ejemplo.

Sea $X_1, X_2, \dots X_n, \dots$ la sucesión de variables aleatorias cuya función de probabilidad es:

X_n	0	n
$P(X_n)$	$1-\frac{1}{n}$	$\frac{1}{n}$

, para que se entienda mejor:

X_1	0	1
$P(X_1)$	0	1

(esta variable es degenerada. Si recordamos, una variable aleatoria es degenerada en $x=a$ si $P(a)=1$)

X_2	0	2
$P(X_2)$	$\frac{1}{2}$	$\frac{1}{2}$

aquí, al igual que en las distribuciones anteriores de probabilidad, tan solo se mencionan estos dos valores porque en el resto la probabilidad es cero, por ejemplo $P(X_2=5)=0$

X_3	0	3
$P(X_3)$	$\frac{2}{3}$	$\frac{1}{3}$

Si tomamos la variable degenerada X con función de probabilidad

X	0	$\neq 0$
$P(X)$	1	0

Entonces $\{X_n\}_{n\in\mathbb{N}} \xrightarrow{L} X$. Veamos esto.

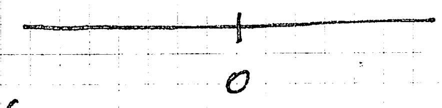
La función de distribución de X_n es $F_n(x)$

$F_n(x) = P(X_n \leq x) = \sum_{x_i \leq x} P(X_n = x_i)$ por tanto

$$\underset{}{0}\quad \underset{0}{|}\quad 1-\frac{1}{n}\quad \underset{n}{|}\quad 1-\frac{1}{n}+\frac{1}{n}=1$$

$$F_n(x) = \begin{cases} 0 & \text{si } x<0 \\ 1-\frac{1}{n} & \text{si } 0\leq x<n \\ 1 & \text{si } n\leq x \end{cases}$$

La función de distribución de X es:

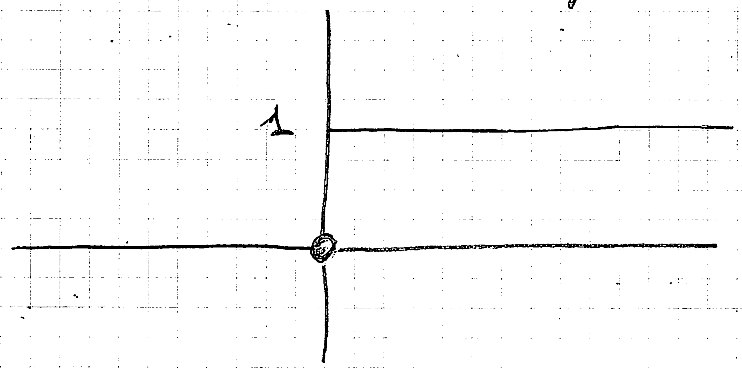


$$F(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 & \text{si } 0 \leq x \end{cases}$$

Se verifica evidentemente que

$\lim_{n \to \infty} F_n(x) = F(x)$. Tan sólo tendríamos que ver que $F(x)$ es función de distribución, pero esto es casi obvio:

* $F(-\infty) = 0$
* $F(+\infty) = 1$
* F es no decreciente ya que su gráfica es:



* F es continua a la derecha en cada punto

Cuando $\{X_n\} \xrightarrow{L} X$ se dice que X_n se distribuye asintóticamente como la variable aleatoria X.

Proposición 1:

Sea $\{X_n\}$ una sucesión de variables aleatorias discretas y sea X una variable aleatoria discreta. Sea $P_{X_n}(x_n^0)$ y $P_X(x_n^0)$ las funciones de probabilidad de X_n y X valoradas en el mismo punto x_n^0. Entonces:

$$\{X_n\} \xrightarrow{L} X \iff \lim_{n \to \infty} P_{X_n}(x_n^0) = P_X(x_n^0)$$

Proposición 2

Sea $\{X_n\}_{n\in\mathbb{N}}$ una sucesión de variables aleatorias continuas y X una variable aleatoria continua con funciones de densidad respectivas $f_{X_n}(x)$ y $f_X(x)$. Entonces:

si $\lim\limits_{n\to\infty} f_{X_n}(x) = f_X(x) \Rightarrow \{X_n\} \xrightarrow{L} X$

Proposición 3

Sea $\{X_n\}_{n\in\mathbb{N}}$ una sucesión de variables aleatorias y X una variable aleatoria, con funciones generatriz de momentos respectivas $\phi_{X_n}(t)$ y $\phi_X(t)$ que suponemos que existen en el intervalo $(-t_0, t_0)$. Entonces:

Si $\lim\limits_{n\to\infty} \phi_{X_n}(t) = \phi_X(t) \Rightarrow \{X_n\} \xrightarrow{L} X$

Otras propiedades de la convergencia en ley

* Sea $\{X_n\}_{n\in\mathbb{N}}$ una sucesión de v.a. y X una variable aleatoria tal que $\{X_n\} \xrightarrow{L} X$. Entonces $\forall K$ constante se verifica:

a) $\{X_n + C\} \longrightarrow X + C$

$\{C \cdot X_n\} \longrightarrow C \cdot X$

* Si g es una función continua entonces:

$\{g(X_n)\}_{n\in\mathbb{N}} \xrightarrow{L} g(X)$

Ejemplo

La función de probabilidad de la distribución de Poisson de parámetro λ es:

$P(X = k) = \dfrac{e^{-\lambda} \cdot \lambda^k}{k!}$. Si X se distribuye según una distribución de Poisson se escribe $X \rightsquigarrow P(\lambda)$

Si $X \rightsquigarrow P(\lambda)$ entonces:

$\mu = E[X] = \lambda$
$\sigma^2 = Var[X] = \lambda \Rightarrow \sigma = \sqrt{\lambda}$

Proposición

Sea $\{X_n\}_{n\in\mathbb{N}}$ una sucesión de variables aleatorias discretas que se distribuyen según una distribución de Poisson de parámetro $\{\lambda_n\}_{n\in\mathbb{N}}$, es decir que $X_n \rightsquigarrow P(\lambda_n)$. Entonces la variable aleatoria tipificada:

$$Y_n = \frac{X_n - \lambda_n}{\sqrt{\lambda_n}} \quad \text{verifica:}$$

$$\{Y_n\}_{n\in\mathbb{N}} \xrightarrow{L} Y \rightsquigarrow N(0,1)$$

¿Para qué nos sirve todo esto?

El cálculo de probabilidades con variables discretas a veces se vuelve muy complicado. Por ejemplo, supongamos que $X \rightsquigarrow P(44)$ y queremos calcular $P(X \leq 987)$. Si hacemos uso de la definición de la distribución de Poisson

$$P[X \leq 987] = \sum_{k=0}^{987} P(X=k) = \sum_{k=0}^{987} \frac{e^{-44} \cdot 44^k}{k!}, \text{ el}$$

hallar el valor de este sumatorio es bastante penoso. Si se pudiera aproximar X a una normal el calculo de $P(X \leq 987)$ sería mucho más sencillo. Veamos esto con el siguiente ejercicio.

Ejercicio:

Dos aparcamientos de una zona comercial pueden absorber a la hora durante los periodos de máxima actividad un maximo de 499 vehiculos. El número de vehiculos que llegan por hora buscando aparcamiento sigue una distribución de Poisson de parámetro 475. ¿Cuál es la probabilidad de que a una determinada hora la zona comercial pierda clientes que no encuentren aparcamiento en ella?

Solución:

Sea X la variable aleatoria que nos da el número de vehículos que llega en una hora determinada. Entonces X se distribuye como

$X \rightsquigarrow P(475)$ y nos piden $P(X > 499)$

ya que si $X > 499$ habrá vehículos que no puedan aparcar (hay más vehículos que plazas de aparcamiento).

Observación: corrección por continuidad

En una distribución discreta tiene sentido calcular la probabilidad en un punto. Por ejemplo si $X \rightsquigarrow P(4)$ tiene sentido calcular $P(X=9)$. Por el contrario, en una distribución continua la probabilidad en un punto es cero. Por ejemplo si $X \rightsquigarrow N(4,2)$ entonces $P(X=9)=0$. En las distribuciones continuas tiene sentido calcular la probabilidad en un intervalo. Por ejemplo si $X \rightsquigarrow N(4,2)$ tiene sentido calcular

- $P(X<9) = P(-\infty < X < 9)$
- $P(X \geq 9) = P(9 \leq X < +\infty)$
- $P(3 \leq X < 9)$

Si aproximamos una distribución discreta como por ejemplo una distribución de Poisson a una distribución continua como por ejemplo $N(0,1)$ estaremos mezclando el concepto de probabilidad en un punto con el de probabilidad en un intervalo.

Si $X \rightsquigarrow P(\lambda) \Rightarrow \mu = \lambda \quad \sigma = \sqrt{\lambda}$

Si aproximamos, por ejemplo, $X \approx N(\mu, \sigma)$ entonces:

- Considerando que $X \rightsquigarrow P(\lambda) \Rightarrow P[X=5] \neq 0$
- Considerando que $X \rightsquigarrow N(\mu, \lambda) \Rightarrow P[X=5] = 0$

Para remediar este inconveniente hacemos corresponder el valor de la probabilidad puntual $P(X=5)$ considerando $X \rightsquigarrow P(\lambda)$ con el valor de la probabilidad por intervalo $P(5-0'5 \leq X \leq 5+0'5)$ considerando que $X \rightsquigarrow N(\mu, \sigma)$

Tenemos la siguiente equivalencia entre el valor puntual considerando que la variable es discreta y el intervalo considerando que la variable es continua

Valor puntual		Intervalo equivalente
$X = 5$	$\longrightarrow$	$4'5 \leq X \leq 5'5$ (dos colas)
$X \geq 5$	$\longrightarrow$	$5'5 \leq X < +\infty$ (una cola)
$X \leq 5$	$\longrightarrow$	$-\infty < X \leq 5'5$ (una cola)

Visto esto volvemos al ejercicio

$\longrightarrow$

La variable X = nº vehículos que llegan a una hora determinada al aparcamiento, que se distribuye como $X \rightsquigarrow P(475)$. Nos piden que calculemos la probabilidad de que en una hora determinada los coches que llegan no encuentren aparcamiento, es decir $P(X > 499)$

Si pretendiéramos calcular este valor "a la brava" procederíamos así:

$$P(X>499) = 1 - P[X \leq 499] = 1 - \sum_{K=0}^{499} P(X=K) =$$

$$1-\sum_{k=0}^{499} \frac{e^{-475} \cdot 475^{k}}{k!}$$ ($475 = \lambda$) (por tratarse de una distribución de Poisson $P(475)$). Este cálculo es penoso y largo, por ello hemos de aproximar a una normal.

corrección por continuidad

$$P[X>499] \doteq P[X \geq 499'5]$$

consideramos que $X \rightsquigarrow P[475]$

consideramos que $X \rightsquigarrow N(475, \sqrt{475})$ ($\mu = 475$, $\sigma = \sqrt{475}$)

tipificación

$$P[X>499] = P[X \geq 499'5] \doteq P\left[\underbrace{\frac{X-475}{\sqrt{475}}}_{Z} \geq \frac{499'5-475}{\sqrt{475}}\right] =$$

$$= P[Z \geq 1'2] = 1-P[Z \leq 1'2] = 1-0'8686 = 0'1314$$

No lo podemos buscar en la tabla $N(0,1)$

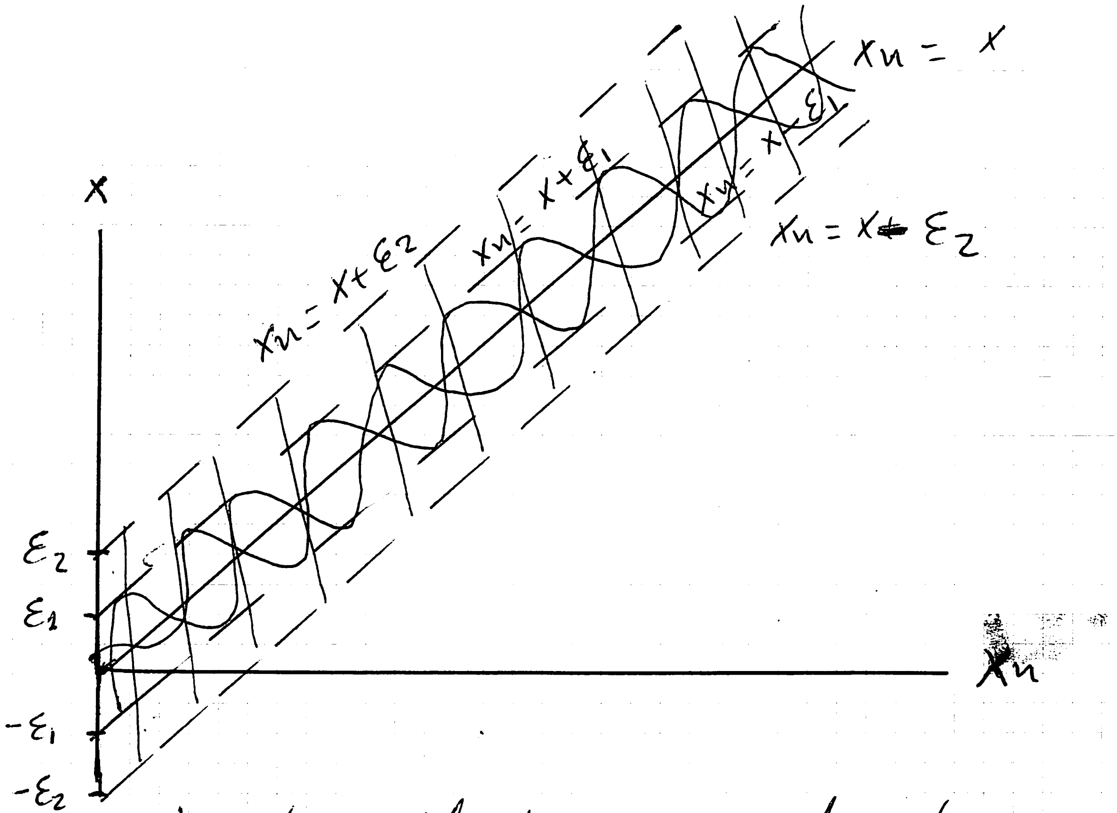
* Convergencia en Probabilidad

Sea $\{X_n\}_{n\in\mathbb{N}}$ una sucesión de variables aleatorias y sea X una variable aleatoria. Se dice que la sucesión $\{X_n\}_{n\in\mathbb{N}}$ converge en probabilidad a X y se escribe $\{X_n\}_{n\in\mathbb{N}} \xrightarrow{P} X$ si se verifica que $\forall \varepsilon > 0$

$\lim_{n\to\infty} P[|X_n - X| \leq \varepsilon] = 1$. Esta definición es equivalente a poner $\forall \varepsilon > 0$ $\lim_{n\to\infty} P[|X_n - X| > \varepsilon] = 0$.

Vamos a dar una interpretación a la convergencia en probabilidad.

Consideremos la variable bidimensional (X_n, X) que representaremos en el plano, situando X_n en el eje de abcisas y X en el de ordenadas.



La bisectriz del primer cuadrante es $x_n = x$

Que $|x_n - x| \le \varepsilon$ es equivalente a poner que $x - \varepsilon \le x_n \le x + \varepsilon$. Lo que viene a decir el concepto de convergencia en probabilidad es que con independencia en que x_n se encuentre en un tubo más ancho (en la figura el tubo en rojo $(x - \varepsilon_2, x + \varepsilon_2)$) o que x_n se encuentre en el tubo más estrecho (en la figura el tubo en negro $(x - \varepsilon_1, x + \varepsilon_1)$) se verifica que

$$\lim_{n \to \infty} P[\, |x_n - x| \le \varepsilon \,] = 1.$$

Ejemplo.

Sea $\{x_n\}_{n \in \mathbb{N}}$ una sucesión de variables aleatorias con función de probabilidad:

x_n	$\frac{1}{n}$	n
$P(x_n)$	$1 - \frac{1}{n^2}$	$\frac{1}{n^2}$

La variable aleatoria nula es aquella que toma el valor cero para cualquier suceso. Se representa por 0 y verifica que $0(w) = 0 \quad \forall w \in \Omega$

Vamos a ver que $\{X_n\}_{n\in\mathbb{N}} \xrightarrow{P} 0$. Para ello tendremos que demostrar que:

$\forall \varepsilon > 0 \quad P[|X_n - 0| \leq \varepsilon] = 1 \iff \forall \varepsilon > 0 \; P[|X_n| \leq \varepsilon] = 1$

Si $|X_n| \leq \varepsilon \Rightarrow X_n$ ha de tomar el valor $X_n = \frac{1}{n}$ ya que si X_n tomara el otro valor, es decir $X_n = n$ no podríamos asegurar que $\forall \varepsilon \; X_n \leq \varepsilon$ si por ejemplo $\varepsilon = 0'01 > 0$ y fuera $X_n = n$, n es un número entero positivo $n \geq 1 \Rightarrow$ $\Rightarrow X_n = n > \varepsilon$. Luego si $|X_n| \leq \varepsilon \quad \forall \varepsilon > 0$ $\Rightarrow X_n = \frac{1}{n} \Rightarrow P[|X_n| \leq \varepsilon] = P[X_n = \frac{1}{n}] =$ $= 1 - \frac{1}{n^2}$. Por tanto $\forall \varepsilon > 0$

$$\lim_{n\to\infty} P[|X_n - 0| \leq \varepsilon] = \lim_{n\to\infty} 1 - \frac{1}{n^2} = 1 \text{ y}$$

ya hemos demostrado que $\{X_n\}_{n\in\mathbb{N}} \xrightarrow{P} 0$

Proposición

Sea $\{X_n\}_{n\in\mathbb{N}}$ una sucesión de variables aleatorias y sea X una variable aleatoria. Entonces si $\{X_n\}_{n\in\mathbb{N}} \xrightarrow{P} X \Rightarrow \{X_n\}_{n\in\mathbb{N}} \xrightarrow{L} X$

Convergencia en media cuadrática

Sea $\{X_n\}_{n\in\mathbb{N}}$ una sucesión de variales aleatorias y sea X una variable aleatoria. Se dice que X_n converge en media cuadrática a X y se escribe $\{X_n\} \xrightarrow{2} X$ o bien $\{X_n\} \xrightarrow{m.c.} X$ si se verifica que $\lim_{n\to\infty} E[(X_n - X)^2] = 0$.

La convergencia en media cuadrática es la convergencia más fuerte ya que engloba a las dos anteriores.

Lo que quiere decir la convergencia en media cuadrática es que conforme aumenta n hay poca dispersión entre el valor de la variable aleatoria X_n y el valor de X.

Ejemplo

Sea $\{X_n\}_{n\in\mathbb{N}}$ la sucesión de variables aleatorias con función de probabilidad:

X_n	$\frac{1}{n}$	n
$P(X_n)$	$1-\frac{1}{n^2}$	$\frac{1}{n^2}$

¿Se verifica que $\{X_n\} \xrightarrow{m.c} 0$?

Solución:

Si fuera cierto que $\{X_n\}_{n\in\mathbb{N}} \xrightarrow{m.c} 0$ entonces $\lim_{n\to\infty} E[(X_n - 0)^2] = 0 \Rightarrow \lim_{n\to\infty} E[(X_n)^2] = 0$

$E[X_n^2] = \sum x_i^2 \cdot P[X = x_i) = \sum \frac{1}{n^2} \cdot P[X_n = \frac{1}{n}] +$

$+ n^2 \cdot P[X_n = n] = \sum \frac{1}{n^2} \cdot \left(1 - \frac{1}{n^2}\right) + n^2 \cdot \frac{1}{n^2}$

$= -\frac{1}{n^4} + \frac{1}{n^2} + 1 \Rightarrow \lim_{n\to\infty} E[(X_n - 0)^2] = \lim_{n\to\infty}\left(-\frac{1}{n^4} + \frac{1}{n^2} + 1\right)$

$\Rightarrow \lim_{n\to\infty} E[(X_n - 0)^2] = 1$ por tanto:

$$\{X_n\} \not\xrightarrow{m.c} 0$$

Proposición

Sea $\{X_n\}_{n\in\mathbb{N}}$ una sucesión de variables aleatorias y C una variable aleatoria. Entonces

$$\{X_n\}_{n\in\mathbb{N}} \xrightarrow{m.c.} C \iff \begin{cases} \lim_{n\to\infty} E[X_n] = C \\ \lim_{n\to\infty} Var[X_n] = 0 \end{cases}$$

Demostración:

Recordemos que $E[aX+bY] = aE[X] + bE[Y]$

y que por definición $Var[X_n] = E[(X_n - E[X])^2]$

Veamos en primer lugar que $E[(X_n - C)^2] = Var[X_n] + (E[X_n] - C)^2$

Si sumamos $E[X_n]$ y restamos $E[X_n]$ en la expresión $E[(X_n - C)^2]$ y desarrollamos obtenemos:

$$E[(X_n - C)^2] = E[(X_n - E[X_n] + E[X_n] - C)^2] =$$

$$= E[((X_n - E[X_n]) + (E[X_n] - C))^2] = E[(X_n - E[X_n])^2 +$$

$$+ 2\cdot(X_n - E[X_n])\cdot(E[X_n] - C) + (E[X_n] - C)^2] =$$

Propiedad de E

$$= E[(X_n - E[X_n])^2] + 2\cdot E[(X_n - E[X_n])\cdot(E(X_n) - C)] +$$

$+E[(E[X_n] - C)^2]$. Tenemos que tener en cuenta que $E[X_n] - C$ es constante, por tanto su esperanza es este mismo valor, es decir $E[(E[X_n] - C)^2] = (E[X_n] - C)^2$ y que $E[X_n - E[X_n]] = E[X_n] - E[X_n] = 0$, al ser constante $E[X_n]$ su esperanza es este mismo valor. Si tenemos en cuenta todo esto la expresión anterior se reduce a:

$$E[(X_n - C)^2] = E[(X_n - E[X_n])^2] + (E[X_n] - C)^2 \Rightarrow$$

$$\Rightarrow E[(X_n - C)^2] = Var[X_n] + (E[X_n] - C)^2.$$

Se tiene que:

$$\{X_n\} \xrightarrow{m.c.} C \iff \lim_{n\to\infty} E[(X_n - C)^2] = 0 \iff$$

$\Leftrightarrow \lim_{n\to\infty} Var[X_n] + (E[X_n]-c)^2 = 0$ y como se trata del valor de dos sumandos positivos $\Leftrightarrow$

$$\Leftrightarrow \begin{cases} \lim_{n\to\infty} Var[X_n] = 0 \\ \lim_{n\to\infty} (E[X_n]-c)^2 = 0 \Leftrightarrow \lim_{n\to\infty} E[X_n]-c=0 \Leftrightarrow \lim_{n\to\infty} E[X_n] = c \end{cases}$$

Proposición:

Sea $\{X_n\}$ una sucesión de variables aleatorias, sea X una variable aleatoria tal que:

$\{X_n\}_{n\in\mathbb{N}} \xrightarrow{m.c.} X$. Entonces:

a) $\{X_n\}_{n\in\mathbb{N}} \xrightarrow{P} X$

b) $\{X_n\}_{n\in\mathbb{N}} \xrightarrow{L} X$

Teorema Central del límite

Se dice que la sucesión de variables aleatorias $\{X_n\}_{n\in\mathbb{N}}$ verifica el teorema central del límite si la correspondiente sucesión de variables tipificadas convergen en ley a la variable $N(0,1)$, o lo que es lo mismo la sucesión $\{X_n\}$ converge en ley a una normal con media y varianza las de X_n

Si X_n tiene por media μ_n y por desviación típica σ_n:

$\{X_n\}_{n\in\mathbb{N}}$ cumple el teorema central del límite si:

$\{X_n\}_{n\in\mathbb{N}} \xrightarrow{L} X \rightsquigarrow N(\mu_n, \sigma_n)$ o bien si

$y_n = \dfrac{X_n - \mu_n}{\sigma_n}$ entonces $\{y_n\}_{n\in\mathbb{N}} \xrightarrow{L} y \rightsquigarrow N(0,1)$

Teorema Central del límite de Lindeberg-Levy

Sea $\{x_n\}_{n\in\mathbb{N}}$ una sucesión de variables aleatorias independientes, todas con la misma distribución de probabilidad, con la misma media μ y la misma varianza σ^2. Sea $\{y_n\}_{n\in\mathbb{N}}$ la sucesión de variables aleatorias:

$$y_1 = x_1 \ , \quad y_2 = x_1 + x_2 \ , \quad y_3 = x_1 + x_2 + x_3 \ , \ \dots \quad y_n = \sum_{i=1}^{n} x_i$$

Entonces la sucesión $\{y_n\}_{n\in\mathbb{N}}$ verifica el teorema central del límite, es decir:

$$\{y_n\}_{n\in\mathbb{N}} \xrightarrow{L} N(\mu, \sigma) \quad \text{o bien:}$$

$$\left\{\frac{y_n - \mu}{\sigma}\right\}_{n\in\mathbb{N}} \xrightarrow{L} N(0,1)$$

Hemos de hacer notar que en estadística, "grande" quiere decir que $n \geq 30$

Para el caso de aproximar una binomial a una normal existe otro teorema central del límite.

Teorema Central del límite de Moivre

Sea $\{X_n\}_{n\in\mathbb{N}}$ una sucesión de variables aleatorias independientes, todas con distribución de Bernouilli de parámetro p, $X_n \rightsquigarrow b(p)$. Entonces la sucesión $\{y_n\}_{n\in\mathbb{N}}$ definida por $y_n = \sum_{i=1}^{n} x_i$ se distribuye según una binomial $B(n,p)$:

$y_1 \rightsquigarrow B(1,p) \quad y_2 \rightsquigarrow B(2,p) \ \dots \ y_n \rightsquigarrow B(n,p)$

Entonces la media de y_n es $E[y_n] = n \cdot$ [illegible]

y la varianza es $Var(y_n) = n \cdot p \cdot q$ donde $q = 1 - p$

Además $\{y_n\}_{n\in\mathbb{N}}$ verifica que:

$$\{y_n\}_{n\in\mathbb{N}} \xrightarrow{L} N(\mu, \sigma) = N(n\cdot p, \sqrt{npq})$$

y la variable tipificada $z_n = \dfrac{y_n - np}{\sqrt{npq}}$

cumple que:

$$\{z_n\}_{n\in\mathbb{N}} \xrightarrow{L} N(0,1)$$

Ejemplo:

Se sabe que las ventas de un producto en un determinado comercio oscilan entre 20 y 30 unidades diariamente. Obtener la probabilidad de vender más de 4800 unidades en 200 días

Solución

Si x_i es la variable aleatoria: "ventas en el día i" entonces x_i se distribuye según una distribución uniforme en el intervalo $[20, 30]$

x_1 = ventas en el primer día

x_2 = ventas en el 2º día

$\vdots$

x_n = ventas en el nº día

Si $S_{200} = x_1 + x_2 + \cdots + x_{200}$ queremos calcular

$P(S_{200} \geq 4800)$. Como la distribución uniforme no es reproductiva no tenemos más remedio que aplicar el teorema central del límite para hallar la solución. Las variables son indep. e idénticamente distribuidas

$$\mu = E[x_i] = \frac{a+b}{2} = \frac{20+30}{2} = 25$$

$$\sigma^2 = Var[x] = \frac{(b-a)^2}{12} = \frac{(30-20)^2}{12} = \frac{25}{3}$$

Si aplicamos el teorema de Lindeberg - Levy

$S_n \approx N[E[S_n], \sqrt{Var[S_n]})$

$S_{200} \approx N[200 \cdot 25, \sqrt{200 \cdot \frac{25}{3}}]$

$S_{200} \approx N[5000, 40'82]$

$$P[S_{200} \geq 4800] = P\left[\frac{S_{200} - 5000}{40'82} \geq \frac{4800 - 5000}{40'82}\right]$$

$$P[S_{200} \geq 4800] = P[Z \geq -4'89] = P[Z \leq 4'89] = 1$$

(el valor se sale de los que hay en la tabla)

OTRA versión del teorema central del límite: el teorema de Moivre (aproximación de una binomial a una normal).

Sea $X_1, X_2, \ldots X_n, \ldots$ una sucesión de variables de Bernouilli independientes

$$X_i \rightsquigarrow b(p) \quad i = 1, 2, \ldots n, \ldots$$

Entonces la variable aleatoria $Y_n = \sum_{i=1}^{n} X_i \rightsquigarrow B(n, p)$

y verifica que $\{Y_n\}_{n \in \mathbb{N}} \xrightarrow{\mathcal{L}} N(n \cdot p, \sqrt{npq})$

(Recordemos que en la $B(n,p)$. $\mu = np$ $\sigma = \sqrt{npq}$ siendo $q = 1 - p$)

Por tanto:

$$\frac{Y_n - np}{\sqrt{npq}} \approx N(0,1)$$