Bloque III(tercera semana)

Contenidos Conceptuales: Teoría asintótica. Sucesiones de variables aleatorias. Modos de convergencia: casi seguro, en probabilidad, en media cuadrada, en distribución. Ley de los grandes números. Teorema Central del Límite. Aplicaciones.

CONVERGENCIA EN PROBABILIDAD

Decimos que una sucesión de variables aleatorias $\{X_n\}$ converge en probabilidad a la variable aleatoria X, y lo expresamos simbólicamente:

$$X_n \xrightarrow{p} X$$

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

A esta convergencia también se la conoce con el nombre de convergencia débil.

Ejemplo 1:

Definamos los siguientes intervalos:

$I_1 = [0, 1]$

$I_2 = [0, ½]$

$I_3 = [½, 1]$

$I_4 = [0, ¼]$

$I_5 = [¼, ½]$

$I_6 = [½, ¾]$

$I_7 = [¾, 1]$

...

y definamos la sucesión de variables aleatorias:

$$X_n = \begin{cases} 1 & si\ x \epsilon I_n \\ 0 & si\ x \notin I_n \end{cases}$$

veamos que $X_n \xrightarrow{p} 0$

en efecto, dado cualquier $\varepsilon>0$ ocurre que

$P(\ |X_n - 0| > \varepsilon\ ) = P(\ X_n > \varepsilon\ )$ ya que las X_n son todas positivas

$= P(\ X_n = 1)$ ya que las X_n adoptan valores 0 o 1 y $\varepsilon>0$

$= P(\ X_n \in I_n\ )$

y $\lim_{n\to\infty} P(\ X_n \in I_n\ ) = 0$

Propiedades de la convergencia en probabilidad:

Sean: $\{X_n\}$ e $\{Y_n\}$ sucesiones de variables aleatorias, $\{x_n\}$ una sucesión numérica, X e Y dos variables aleatorias y k y b dos números reales fijos, entonces:

1.- Si $x_n \to k \Rightarrow x_n \xrightarrow[p]{} k$

2.- Si $X_n \xrightarrow[p]{} X$ y $Y_n \xrightarrow[p]{} Y \Rightarrow X_n + Y_n \xrightarrow[p]{} X + Y$

3.- Si $X_n \xrightarrow[p]{} X$ y b constante cualquiera $\Rightarrow b.X_n \xrightarrow[p]{} b.X$

4.- Si $X_n \xrightarrow[p]{} X$ y $Y_n \xrightarrow[p]{} Y \Rightarrow X_n.Y_n \xrightarrow[p]{} X.Y$

5.- Si $\{Y_n\}$ es acotada y $X_n \xrightarrow[p]{} 0 \Rightarrow X_n.Y_n \xrightarrow[p]{} 0$

6.- Si f es una función acotada y continua en c y $X_n \xrightarrow[p]{} c \Rightarrow E\left(f(X_n)\right) \xrightarrow[n\to+\infty]{} f(c)$

7.- Si $X_n \xrightarrow[p]{} X$ y $Y_n \xrightarrow[p]{} Y \Rightarrow (X_n, Y_n) \xrightarrow[p]{} (X, Y)$

Ley débil de los grandes números (de Kintchin):

Sea $X_1, X_2, \ldots, X_n, \ldots$ una sucesión de variables aleatorias independientes e igualmente distribuidas (*iid*) con media μ y varianza σ^2 ambas finitas.

Sea la sucesión $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ (sucesión de promedios o medias muestrales de una muestra aleatoria de n elementos). Entonces: $\bar{X}_n \xrightarrow[p]{} \mu$.

Demostración:

Recordando que $E\left(\overline{X}_n\right)=\mu$; $Var\left(\overline{X}_n\right) = \dfrac{\sigma^2}{n}$ y usando Chebyschev resulta que:

$$P(\left|\overline{X}_n-\mu\right|>\varepsilon) \leq \frac{\sigma^2}{n.\varepsilon^2}.$$

Como $P(\left|\overline{X}_n-\mu\right|>\varepsilon)\geq 0$ resulta que $\lim\limits_{n\to+\infty} P\left(\left|\overline{X}_n-\mu\right|>\varepsilon\right)\geq 0$

$$\Rightarrow 0 \leq \lim_{n\to+\infty} P\left(\left|\overline{X}_n-\mu\right|>\varepsilon\right) \leq \lim_{n\to\infty}\frac{\sigma^2}{n.\varepsilon^2} = 0 \Leftrightarrow \lim_{n\to+\infty} P\left(\left|\overline{X}_n-\mu\right|>\varepsilon\right) = 0$$

$$\Leftrightarrow \overline{X}_n \xrightarrow[p]{} \mu.$$

Nota:

Este resultado es válido aún cuando las variables $\boldsymbol{X}_1$, $\boldsymbol{X}_2$, ..., $\boldsymbol{X}_n$,… no tengan varianza finita, pero omitimos esa demostración en este curso por su dificultad.

Ejemplo 2:

Sean $\boldsymbol{X}_1$, $\boldsymbol{X}_2$, ..., $\boldsymbol{X}_n$,… variables aleatorias independientes con distribución binomial de parámetros 10 y ¼. Decidir a qué convergen en probabilidad las siguientes sucesiones:

a.- $\boldsymbol{Y}_n = \dfrac{\sum_{i=1}^{n}\left(X_i-7\right)}{n}$

Observar que $\boldsymbol{Y}_n = \dfrac{\sum_{i=1}^{n} X_i - \sum_{i=1}^{n} 7}{n} = \dfrac{\sum_{i=1}^{n} X_i - n.7}{n} = \overline{X}_n - 7$

Por la ley de los grandes números $\overline{X}_n \xrightarrow[p]{} \mu = E\left(X_1\right)$

Como $E(\boldsymbol{X}_1)$ = 10. ¼ = $5/2$, resulta que: $\overline{X}_n \xrightarrow[p]{} 5/2$.

A su vez la sucesión constante 7 converge a 7 numéricamente, por lo tanto converge en probabilidad. Así resulta que: $\boldsymbol{Y}_n = \overline{X}_n - 7 \xrightarrow[p]{} 5/2 - 7$

Por lo tanto: $Y_n \xrightarrow[p]{} -9/2$

b.- $\boldsymbol{W}_n = \dfrac{(3n-1)\sum_{i=1}^{n} X_i}{n^2}$

Observar que W_n = $\dfrac{(3n-1)\sum_{i=1}^{n} X_i}{n^2}$ = $\dfrac{3n-1}{n}\dfrac{\sum_{i=1}^{n} X_i}{n}$ = $\dfrac{3n-1}{n}.\overline{X}_n$

Por la ley de los grandes números $\overline{X}_n \xrightarrow[p]{} \mu = E(X_1) \Rightarrow \overline{X}_n \xrightarrow[p]{} 5/2$

La sucesión numérica: a_n = $\dfrac{3n-1}{n} \xrightarrow[n\to+\infty]{} 3$; por lo tanto $\dfrac{3n-1}{n} \xrightarrow[p]{} 3$

Resulta así que: $\dfrac{3n-1}{n}.\overline{X}_n \xrightarrow[p]{} 3.\dfrac{5}{2}$.

O sea que: $W_n \xrightarrow[p]{} 15/2$

Proposición

Sean $\boldsymbol{X}_1$, $\boldsymbol{X}_2$, ..., $\boldsymbol{X}_n$,... variables aleatorias independientes, no necesariamente con idéntica distribución, pero con varianzas finitas y uniformemente acotadas, es decir:

$\exists\, c \in \mathbf{R}$ tal que $Var(\boldsymbol{X}_i) \leq c \;\; \forall i$. Entonces: $\dfrac{1}{n}\sum_{i=1}^{n}\left[X_i - E(X_i)\right] \xrightarrow[p]{} 0$.

Demostración

$$E\left(\frac{1}{n}\sum_{i=1}^{n}\left[X_i - E(X_i)\right]\right) = \frac{1}{n}\sum_{i=1}^{n}E\left[X_i - E(X_i)\right] = \frac{1}{n}\sum_{i=1}^{n}\left[E(X_i) - E(E(X_i))\right] =$$

$$= \frac{1}{n}\sum_{i=1}^{n}\left[E(X_i) - E(X_i)\right] = 0$$

Por otro lado:

$$Var\left(\frac{1}{n}\sum_{i=1}^{n}\left[X_i - E(X_i)\right]\right) = \frac{1}{n^2}\sum_{i=1}^{n}V(X_i - E(X_i)) = \text{(por ser las v.a. independientes)}$$

$$= \frac{1}{n^2}\sum_{i=1}^{n}V(X_i) \leq \text{(por ser } E(\boldsymbol{X}_i) \text{ una constante)}$$

(por unif. acotadas) $\leq \dfrac{1}{n^2}\sum_{i=1}^{n} c = \dfrac{1}{n^2}n.c = \dfrac{1}{n}c$

Usando la desigualdad de Chebyschev y que las probabilidades son números positivos:

$$0 \leq P\left(\left|\sum_{i}^{n}\frac{\left(X_i - E(X_i)\right)}{n} - 0\right| > \varepsilon\right) \leq \frac{Var\left(\frac{1}{n}\sum_{i=1}^{n}\left[X_i - E(X_i)\right]\right)}{\varepsilon^2} \leq \frac{\frac{1}{n}c}{\varepsilon^2} \leq \frac{c}{n.\varepsilon^2}$$

Tomando $\lim\limits_{n\to\infty}$ para ε fijo resulta que $P\left(\left|\sum_i^n \frac{(X_i - E(X_i))}{n} - 0\right| > \varepsilon\right)$ = 0. Por lo tanto: $\frac{1}{n}\sum_{i=1}^{n}\left[X_i - E(X_i)\right] \xrightarrow[p]{} 0$.

Ejemplo 3:

Sean X_1, X_2, ..., X_n ,... variables aleatorias independientes, con esperanzas finitas y $Var(X_i) \leq 5/i$,$\forall i$. Decidir a qué convergen en probabilidad la sucesión $\bar{X}_n + \frac{1}{10^n}$.

Si $Var(X_i) \leq 5/i$,$\forall i$ $\Rightarrow Var(X_i) \leq 5$ (por ser $i \in \mathbf{N}$) $\Rightarrow$

Por la proposición anterior; $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n}\left[X_i - E(X_i)\right] \xrightarrow[p]{} 0$

La sucesión numérica $\left\{\frac{1}{10^n}\right\}$ verifica que $\frac{1}{10^n} \to 0 \Rightarrow \frac{1}{10^n} \xrightarrow[p]{} 0$

Por las propiedades de la convergencia débil: $\bar{X}_n + \frac{1}{10^n} \xrightarrow[p]{} 0 + 0 = 0$

Convergencia en distribución:

Sean X_1, X_2, ..., X_n,... variables aleatorias con funciones de distribución F_1, F_2,. ..., F_n ,... y sea X una variable aleatoria con función de distribución F.

Diremos que $\{X_n\}$ converge en distribución a X si y sólo si $F_n(x) \xrightarrow[n\to+\infty]{} F(x)$ para todo punto x de continuidad de F.

Es decir la función F_n "se parece" cada vez más a la función F a medida que n crece.
En otras palabras a medida que n crece F es una mejor aproximación de F_n. Por lo tanto se pueden aproximar las probabilidades de X_n utilizando la función de distribución F en lugar de calcularlas exactamente usando la función de distribución F_n.

Notación:

$$X_n \xrightarrow[d]{} X \text{ ó bien } X_n \xrightarrow{d} X$$

Se dice también que F_{X_n} converge débilmente a F_X .

Algunos autores escriben (*w*: weak): $X_n \xrightarrow[w]{} X$

¿Por qué se pide solamente convergencia en los puntos de continuidad de la función?

Sea { X_n } la sucesión de variables aleatorias $X_n = 1/n$, es decir $X_n(w) = 1/n \quad \forall\, w \in \Omega$

y sea la variable aleatoria con valor constante 0 ($X = 0$).

Intuitivamente X_n debería converger de alguna manera a X, de hecho:

$X_n \xrightarrow[p]{} X$ puesto que $\lim\limits_{n\to+\infty} P\left(\left|X_n - 0\right| > \varepsilon\right) = \lim\limits_{n\to+\infty} P\left(\left|1/n\right| > \varepsilon\right) = 0$

Pero observemos que: $F_{X_n}(x) = \begin{cases} 0 & si\ x < 1/n \\ 1 & si\ x \geq 1/n \end{cases}$ y $F_X(x) = \begin{cases} 0 & si\ x < 0 \\ 1 & si\ x \geq 0 \end{cases}$

Observemos que F_X tiene un único punto de discontinuidad que es $x = 0$ y que:

i) si $x > 0$ se cumple que $F_{X_n}(x) \xrightarrow[n\to+\infty]{} 1 = F_X(x)$

ii) si $x < 0$ se cumple que $F_{X_n}(x) \xrightarrow[n\to+\infty]{} 0 = F_X(x)$

iii) si $x = 0$ se cumple que $F_{X_n}(0) \xrightarrow[n\to+\infty]{} 0 \neq F_X(0)$

Además:

Dos variables aleatorias X e Y pueden tener la misma frecuencia de distribución y no solamente no ser iguales sino aún más no coincidir en ningún caso.

Ejemplo 4:

Consideremos el experimento aleatorio de arrojar una moneda equilibrada y mirar la cara superior. Se definen las variables:

$X = \begin{cases} 1 & si\ sale\ cara \\ 0 & si\ sale\ ceca \end{cases}$ $\qquad Y = \begin{cases} 1 & si\ sale\ ceca \\ 0 & si\ sale\ cara \end{cases}$ (son variables distintas)

Las dos variables tienen la misma función de distribución: $F_X = F_Y = \begin{cases} 0 & si\ x < 0 \\ 0.5 & si\ 0 \leq x < 1 \\ 1 & si\ x \geq 1 \end{cases}$

En síntesis: ***la coincidencia en distribución no implica la coincidencia de las variables***

Esto no es así en la convergencia en probabilidad.

Veremos luego que la convergencia en probabilidad implica convergencia en distribución pero no es cierta la recíproca.

Ejemplo 5:

Sean $X_n \sim Bi(n, ½)$ e $Y_n = \dfrac{X_n - n/2}{\sqrt{n/4}}$ (notar que $E(X_n) = n/2$ y $Var(X_n) = n/4$)

Se puede probar que $Y_n \xrightarrow{d} Z$ con Z ~ N(0;1).

Esto indica que a medida que *n* crece, la distribución normal es cada vez una mejor aproximación para calcular probabilidades de distribuciones binomiales de parámetros *n* y $p = ½$.

Caso particular: *Convergencia en distribución a una constante*

Sea { X_n } una sucesión de variables aleatorias y sea $c \in$ **R**. Diremos que:

$$X_n \xrightarrow[d]{} c \iff \lim_{n\to+\infty} F_{X_n}(x) = \begin{cases} 0 & si\ x < c \\ 1 & si\ x \geq c \end{cases} = I_{[c;+\infty)}(x) \quad \forall x \neq c.$$

(si bien en c podría converger, esto no es obligatorio)

Teorema:

(a) Sea { X_n } una sucesión de variables aleatorias discretas con probabilidad puntual $p_n(x)$ y sea X una variable aleatoria con probabilidad puntual $p(x)$. Si se verifica que $p_n(x) \to p(x) \quad \forall x \Rightarrow X_n \xrightarrow{d} X$.

(b) Sea { X_n } una sucesión de variables aleatorias continuas con función de densidad de probabilidad f_n. Si existe una variable aleatoria X con función de densidad f, tal que $f_n(x) \to f(x) \quad \forall x \Rightarrow X_n \xrightarrow{d} X$.

Proposición:

Sea { X_n } una sucesión de variables aleatorias y X una variable aleatoria:

(a) Si $X_n \xrightarrow[p]{} X \Rightarrow X_n \xrightarrow{d} X$

(b) Si $X_n \xrightarrow{d} X$ y $P(X = c) = 1 \Rightarrow X_n \xrightarrow[p]{} X$

Dicho de otra forma: $X_n \xrightarrow{d} c \Rightarrow X_n \xrightarrow[p]{} c$

Consecuencia de (a):

Si la sucesión numérica $\{a_n\}$ verifica que $a_n \xrightarrow[n\to+\infty]{} a$ entonces $a_n \xrightarrow[d]{} a$

<u>Demostración</u>

Si $a_n \xrightarrow[n\to+\infty]{} a \Rightarrow a_n \xrightarrow[p]{} a \Rightarrow a_n \xrightarrow[d]{} a$.

Teorema Central del límite:

Las distintas versiones del teorema central del límite establecen que una sucesión de variables aleatorias converge en distribución a la distribución normal.

Esto nos permite, para valores de *n* grande, aproximar mediante la distribución normal, las probabilidades de la variable que constituye el *n*-ésimo término de la sucesión.

Teorema Central del límite para variables aleatorias independientes:

Sean $X_1, X_2, \ldots, X_n$ variables aleatorias *iid* tales que $E(X_i) = \mu$ y $V(X_i) = \sigma^2$ (finita).

Sea $T_n = \sum_{i=1}^{n} X_i$, entonces:

$$\frac{T_n - E(T_n)}{Var(T_n)} = \frac{T_n - n\mu}{\sqrt{n}.\sigma} \xrightarrow[d]{} Z \quad \text{con } Z \sim N(0;1).$$

O sea que para *n* suficientemente grande la variable $\frac{T_n - E(T_n)}{Var(T_n)}$ se comporta como una variable normal estándar.

En símbolos:

$$\frac{T_n - n\mu}{\sqrt{n}\sigma} \sim N(0;1) \Leftrightarrow T_n \sim N(n\mu; n\sigma^2)$$

Otras formas del teorema central del límite:

1.- $\sqrt{n}\,\frac{\overline{X}_n - \mu}{\sigma} \xrightarrow[d]{} N(0,1)$

2.- $\sqrt{n}\frac{\overrightarrow{X_n} - \mu}{\sigma} \sim N(0;1)$

equivalente a $\overrightarrow{X_n} \sim N(\mu;\ \sigma 2/n)$

<u>Observaciones:</u>

1.- La convergencia en distribución nos dice que si $\frac{T_n - n\mu}{\sqrt{n}\,\sigma} \xrightarrow[d]{} N(0,1)$ entonces

$$P\left(\frac{T_n - n\mu}{\sqrt{n}\,\sigma} \le a\right) \approx \phi_Z(a)$$ con $Z \sim N(0;1)$ con tal de tomar *n* suficientemente grande.

2.- ¿Qué significa *n* suficientemente grande? ¿Cómo sabemos si la aproximación es buena?

El tamaño *n* de la muestra necesario para que la aproximación sea razonable depende la distribución de la variable original. Mientras más acampanada y simétrica es la distribución de las observaciones, más rápidamente se logra una buena aproximación.

<u>**Ejemplo 6:**</u>

Al sumar números una calculadora aproxima cada número al entero más próximo.
Los errores de aproximación se suponen independientes y con distribución uniforme **U**(-0.5,0.5).

a.- Si se suman 1500 números, ¿cuál es la probabilidad de que la magnitud total del error exceda 15 unidades?

Definamos:

$\boldsymbol{X_i}$ = "magnitud del error cometido al aproximar el *i*-ésimo número" ($1 \le i \le 1500$)

La magnitud total del error al sumar los 1500 números es: $\boldsymbol{T}_{1500} = \sum_{i=1}^{1500} X_i$

Recordemos que si $\boldsymbol{X}_i \sim U(-0.5,0.5)$ entonces $E(\boldsymbol{X}_i) = 0$ y $V(\boldsymbol{X}_i) = 1/12$. Como se puede suponer que el error cometido al aproximar un número no tiene porqué influir en el error que se comete en la aproximación del próximo, se trata de variables independientes e igualmente distribuidas. Por lo tanto las variables $\boldsymbol{X}_1,\ldots, \boldsymbol{X}_n$ cumplen con las hipótesis del teorema central del límite y en consecuencia:

$$\frac{T_n - 0}{\sqrt{n}\sqrt{1/12}} \xrightarrow[d]{} Z \text{ con } Z \sim N(0;1) \iff T_n \sim N(n\mu; n\sigma^2)$$

Así resulta que $\boldsymbol{T}_{1500}$ se distribuye en forma aproximada como una normal de parámetros: $n\mu = 0$ y $n.\sigma^2 = 1500/12$.

En consecuencia:

$$P(|T_{1500}| > 15) = 1 - P(|T_{1500}| < 15) = 1 - P(-15 < \boldsymbol{T}_{1500} < 15) =$$

$$= 1 - P\left(\frac{-15}{\sqrt{1500/12}} < \frac{T_{1500} - 0}{\sqrt{1500/12}} < \frac{15}{\sqrt{1500/12}}\right) \cong 1 - (\Phi_Z(1.34) - \Phi_Z(-1.34)) = 0.18.$$

Por lo tanto: $P(|T_{1500}| > 15) \cong 0.18$.

b.- ¿Cuántos números pueden sumarse a fin de que la magnitud del error total sea menor que 10 con probabilidad superior a 0.9?

Deseamos hallar el valor de *n* para que $P(|T_n| < 10) \ge 0.9$. Usando el TCL:

$$P(|T_n| < 10) = P(-10 < T_n < 10) = P\left(\frac{-10}{\sqrt{n/12}} < \frac{T_n - 0}{\sqrt{n/12}} < \frac{10}{\sqrt{n/12}}\right) \cong$$

$$\cong \Phi_Z\left(\frac{10}{\sqrt{n/12}}\right) - \Phi_Z\left(\frac{-10}{\sqrt{n/12}}\right) = 2\Phi_Z\left(\frac{10}{\sqrt{n/12}}\right) - 1 \ge 0.9 \iff \Phi_Z\left(\frac{10}{\sqrt{n/12}}\right) \ge 0.95$$

$$\Leftrightarrow \frac{10}{\sqrt{n/12}} \geq 1.64 \Leftrightarrow \sqrt{n} \leq \frac{10\sqrt{12}}{1.64} \Rightarrow n \leq 446.162 \Rightarrow n \text{ es como máximo } 446.$$

Ejemplo 7:

El contenido de nicotina de un cigarrillo de cierta marca es una variable aleatoria ***X*** con media 0.8 mg y varianza 0.01 (mg)^2.

Si una persona fuma 5 paquetes de 20 cigarrillos cada uno por semana. ¿Cuál es la probabilidad de que en una semana consuma más de 81 mg de nicotina?

Sea X_i = "cantidad de nicotina del cigarrillo i"

Sea T_n = "cantidad de nicotina consumida en la semana" ; $n = 100$

Es razonable suponer que las variables son independientes, pues la cantidad de nicotina de un cigarrillo no tiene porqué influir en la cantidad de nicotina que contenga el siguiente que la persona fumará.

¿Por qué se puede considerar que tienen la misma distribución? Porque son todos cigarrillos de la misma marca.

En virtud de que se cumplen los supuestos del teorema central del límite (TCL)

$\Rightarrow E(T_{100}) = 100\text{x}0.8 = 80$ y $V(T_{100}) = 100\text{x } 0.01 = 1$

$$\Rightarrow P(T \geq 81) = P\left(\frac{T-n\mu}{\sqrt{n}\,\sigma} \geq \frac{81-n\mu}{\sqrt{n}\,\sigma}\right) = P\left(\frac{T-n\mu}{\sqrt{n}\,\sigma} \geq \frac{81-80}{\sqrt{100.0,1}}\right) = 1- \phi_Z(1) = 0.1587$$

Corolario del TCL: Teorema Central del Límite de De-Moivre- Laplace

Este corolario permite aproximar la distribución binomial por la normal:

Sea X ~ Bi(n, p) y sea $\frac{X}{n}$ = "la proporción muestral de éxitos".

Sea $X_i = \begin{cases} 1 & si\ hubo\ exito\ en\ el\ intento\ i \\ 0 & si\ hubo\ fracaso\ en\ el\ intento\ i \end{cases}$

Entonces: X ~ Bi(n, p) $\Rightarrow$ $\boldsymbol{X} = \sum_{i=1}^{n} X_i$

Por el teorema central del límite: X~N(np; np(1-p)) y : X/n ~N(p; p(1-p)/n) º

Se considera que la aproximación es buena cuando: $np \geq 5$ y además $np(1-p) \geq 5$

Ejemplo 8:

a) ¿Cuál es la probabilidad de obtener 198 caras si se lanza una moneda equilibrada 400 veces?

La variable ***T*** = "cantidad de caras en 400 tiradas" es una *Bi*(400 ; 0.5) como se cumple: *np* = 400.0.5 = 200 ≥ 5 y además *np(1-p)* = 400. 0.25 ≥ 100 podemos utilizar la aproximación normal considerando que: T~ N(200; 100)

$$P(T = 198) = P\,(197.5 \le T \le 198.5) = P\left(\frac{197.5-200}{10} \le \frac{T-200}{10} \le \frac{198.5-200}{10}\right) =$$

$$= P\left(-0.25 \le \frac{T-200}{10} \le -0.15\right) \cong \Phi_Z(-0.15) - \Phi_Z(-0.25)$$

= 0.4404 – 0.4013 = 0.0391

b) ¿Cuál es la probabilidad de obtener un número de caras superior a 198 pero que no exceda de 220 si se lanza una moneda equilibrada 400 veces?

Al igual que en la parte a) : $T \sim N(200;100)$

$$P\,(198 < T \le 220) = P\,(199 \le T \le 220) = P\,(198.5 \le T \le 220.5) =$$

$$= P\left(\frac{198.5-200}{10} \le \frac{T-200}{10} \le \frac{220.5-200}{10}\right) \cong \Phi_Z(2.05) - \Phi_Z(-0.15)$$

= 0.9798 – 0.4404 = 0.5394

c) Un productor vende focos en cartones de 1000. Hallar la probabilidad de que cualquier caja dada contenga no más del 1% de focos defectuosos, suponiendo que el proceso de producción sea un experimento Bernoulli con parámetro 0.01.

1% de focos defectuosos es: 10. Por lo tanto se pide que ***T*** ≤ 10 cuando la variable ***T*** = "cantidad de focos defectuosos en una caja de 1000".

Como *np* = 1000.0.01 = 10 ≥ 5 y además *np(1-p)* = 1000. 0.0099= 9.9 ≥ 5 podemos utilizar la aproximación normal considerando que: : T~N(10; 9,9) .

$$P(T \le 10) = P(T \le 10.5) = P\left(\frac{T-10}{\sqrt{9.9}} \le \frac{10.5-10}{\sqrt{9.9}}\right) \cong \Phi_Z(0.158910) = \Phi_Z(0.16) =$$

= 0.5636

Ejemplo 9:

Supongamos que se tira una moneda 100 veces. Si se obtienen 60 caras, ¿es razonable pensar que la moneda es equilibrada?

Deseamos ver si *P*(cara) = *P*(ceca) = *p* = ½ .

Para ello definimos T_{100}= "cantidad de caras en las 100 tiradas". Si fuese p = ½ resultaría que $E(T_{100})$ = 50 y $Var(T_{100})$ = 100.½.½ = 25, entonces por el TCL:

$$\frac{T_{100}-50}{\sqrt{25}} \sim N(0;1) \Rightarrow P(T_{100}=60) \cong P(59.5 \le T_{100} \le 60.5) =$$

$$= P\left(\frac{59.5-50}{5} \le \frac{T_{100}-50}{5} \le \frac{60.5-50}{5}\right) = P(4.75 \le Z \le 5.25) =$$

$$= \Phi_Z(5.25) - \Phi_Z(4.75) \cong 0$$

Por lo tanto es razonable dudar de que la moneda no se encuentre cargada.

Proposición:

Sean X_1, X_2,...,X_n,.. variables aleatorias y sea g: **R**$\rightarrow$**R** una función continua, entonces:

a.- Si $X_n \xrightarrow[p]{} X \Rightarrow g(X_n) \xrightarrow[p]{} g(X)$

b.- Si $X_n \xrightarrow[d]{} X \Rightarrow g(X_n) \xrightarrow[d]{} g(X)$

c.- Sea $\vec{X}_n$ una sucesión de vectores aleatorios $\vec{X}_n = \begin{pmatrix} X_{1n} \\ X_{2n} \end{pmatrix}$; sean $\vec{a} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$; $\vec{b} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$

entonces vale que:

Si $\vec{X}_n \xrightarrow[d]{} \vec{X}$ con $\vec{X} \sim N(\vec{\mu}; \Sigma) \Rightarrow \vec{a}^t \vec{X}_n + \vec{b} \xrightarrow[d]{} \vec{Z}$ con $\vec{Z} \sim N(\vec{a}\vec{\mu}; \vec{a}' \Sigma \vec{a})$

Ejemplo 10:

Sean X_1, X_2,...,X_n,.. una sucesión de variables aleatorias tales que $X_n \xrightarrow[d]{} Z$ con Z una variable con distribución normal estándar. Hallar a qué convergen en distribución la sucesión de los cuadrados de las variables X_i con $i \in$ **N**.

Debemos analizar a qué converge la sucesión $\{X_i^2\}$.

Como $X_n \xrightarrow[d]{} Z$ considerando la función $g(x) = x^2$ que es continua resulta por la proposición anterior que: $g(X_n) \xrightarrow[d]{} g(Z)$ o sea $X_n^2 \xrightarrow[d]{} Z^2$.

Si Z ~ N(0;1) $\Rightarrow$ Z2 ~$\chi^2_{(1)}$.

Por lo tanto se concluye que la sucesión $\{X_i^2\}$ converge en distribución a una variable que tiene distribución chi- cuadrado con 1 grado de libertad.

Ejemplo 11:

Sean $\boldsymbol{X}_1$, $\boldsymbol{X}_2$,...,$\boldsymbol{X}_n$,.. una sucesión de variables aleatorias independientes tales que todas tienen distribución exponencial de parámetro 4. Decidir a qué converge en probabilidad la sucesión $\boldsymbol{W}_n = \dfrac{n}{\sum_{i=1}^{n} X_i}$.

Notar que $\boldsymbol{W}_n = \left(\overline{X}_n\right)^{-1}$ y que por la ley de los grandes números $\overline{X}_n \xrightarrow[p]{} E(X_1)$.

$E(\boldsymbol{X}_1)$ = ¼ pues $X_i \sim \varepsilon(4) \Rightarrow \overline{X}_n \xrightarrow[p]{} 1/4$.

Consideremos la función $g(x) = 1/x$ continua en las reales mayores que cero. Por lo tanto

$$g\left(\overline{X}_n\right) \xrightarrow[p]{} g\left(1/4\right) \Rightarrow \frac{1}{\overline{X}_n} \xrightarrow[p]{} \frac{1}{1/4} .$$

Concluimos que $W_n \xrightarrow[p]{} 4$

Teorema de Slutzky:

Sean $\boldsymbol{X}_1$, $\boldsymbol{X}_2$,...,$\boldsymbol{X}_n$,.. e $\boldsymbol{Y}_1$, $\boldsymbol{Y}_2$,...,$\boldsymbol{Y}_n$,.. dos sucesiones de variables aleatorias y sea $c \in \mathbf{R}$.
Si se verifica que $X_n \xrightarrow[d]{} X$ y $Y_n \xrightarrow[d]{} c$ entonces:

a.- $X_n + Y_n \xrightarrow[d]{} X + c$

b.- $X_n - Y_n \xrightarrow[d]{} X - c$

c.- $X_n . Y_n \xrightarrow[d]{} c.X$

d.- Si $c \neq 0$ y $P\left(Y_n \neq 0\right) = 1 \Rightarrow X_n / Y_n \xrightarrow[d]{} X / c$

<u>Ejemplo 12:</u>

Sea $\{\boldsymbol{Z}_n\}$ una sucesión de variables aleatorias normales estándar. Hallar a qué converge en distribución la sucesión $\{\boldsymbol{X}_n\}$ de variables definidas por: $\boldsymbol{X}_n = \dfrac{n+1}{n} Z_n + \dfrac{2n}{n-1}$.

Si las variables $Z_n \sim N(0;1) \Rightarrow Z_n \xrightarrow[d]{} Z$ con $Z \sim N(0;1)$

Recordando que si $\dfrac{n+1}{n} \xrightarrow[n \to +\infty]{} 1 \Rightarrow \dfrac{n+1}{n} \xrightarrow[d]{} 1$ resulta por las propiedades anteriores que: $\dfrac{n+1}{n} Z_n \xrightarrow[d]{} 1.Z \Rightarrow \dfrac{n+1}{n} Z_n \xrightarrow[d]{} Z$

Como, por otra parte sabemos que la sucesión numérica $\left\{\dfrac{2n}{n-1}\right\}$ verifica que

$$\frac{2n}{n-1} \xrightarrow[n \to +\infty]{} 2 \Rightarrow \frac{2n}{n-1} \xrightarrow[d]{} 2 \Rightarrow \boldsymbol{X}_n = \frac{n+1}{n} Z_n + \frac{2n}{n-1} \xrightarrow[d]{} Z + 2 .$$

Si Z ~ N(0;1) $\Rightarrow$ Z +2 ~ N(2;1).

Por lo tanto concluimos que {X_n} converge en distribución a una normal con media 2 y varianza 1.

Proposicón:

Sean Y_1, Y_2,...,Y_n,...variables aleatorias tales que $\sqrt{n}\left(Y_n-\mu\right)\xrightarrow[d]{}Z$ con **Z ~ N(0;σ^2)**

y sea una función g: **A** $\subseteq$ **R**→**R** que es derivable en $x = \mu$. Se verifica que:

1.- si $g'(\mu) \neq 0$ entonces $\sqrt{n}\left(g\left(Y_n\right)-g(\mu)\right)\xrightarrow[d]{}Z_1$ con $Z_1 \sim N\left(0;\sigma^2\left(g'(\mu)\right)^2\right)$

2.- si $g'(\mu) = 0$ entonces $\sqrt{n}\left(g\left(Y_n\right)-g(\mu)\right)\xrightarrow[d]{}0$

<u>Ejemplo 13:</u>

Sean X_1, X_2, ..., X_n variables aleatorias *iid* tales que $E(X_i) = \mu \neq 0$ y $V(X_i) = \sigma^2$ (finita).

Hallar a qué converge en distribución la sucesión { $\sqrt{n}\left(\overline{X}_n{}^2-\mu\right)$ } siendo $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$.

Si $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$, entonces por el teorema central del límite (TCL):

$$\frac{\overline{X}_n - E(\overline{X}_n)}{Var(\overline{X}_n)} = \sqrt{n}\,\frac{\overline{X}_n-\mu}{\sigma} \xrightarrow[d]{} Z \quad \text{con Z ~ N(0;1)}$$

$\Rightarrow \sqrt{n}\left(\overline{X}_n-\mu\right)\xrightarrow[d]{}Z_2$ con Z^2~ N(0;σ^2).

Sean Y_n = $\overline{X}_n \Rightarrow Y_n^2 = \overline{X}_n{}^2$.

Consideremos la función g: **R**→**R** definida por $g(x) = x^2$ derivable en $x = \mu$ con derivada $g'(\mu) = 2\mu \neq 0$.

Como $\left(g'(\mu)\right)^2 = 4\mu^2$, aplicando la proposición anterior resulta que:

$$\sqrt{n}\left(g\left(\overline{X}_n\right)-g(\mu)\right) = \sqrt{n}\left(\overline{X}_n{}^2-\mu^2\right)\xrightarrow[d]{}Z_1 \quad \text{donde } Z_1\text{~ N(0;}4\mu^2\sigma^2)$$

Convergencia en distribucón para vectores aleatorios:

Deseamos conocer cómo se comporta la convergencia en distribución de una sucesión de vectores aleatorios. Este problema puede reducirse al de convergencia de sucesiones de variables aleatorias, aplicando el teorema de Cramer- Wold que se enuncia a continuación

Teorema de Cramer- Wold:

La sucesión de vectores aleatorios k-dimensionales $\{\vec{X}_n\}$ converge en distribución al vector aleatorio $\{\vec{X}\}$ si y sólo si $\vec{a}^t\vec{X}_n \xrightarrow[d]{} \vec{a}^t\vec{X} \quad \forall\ \vec{a} \in \mathbf{R}^k$.

Corolario:

Si $\vec{X}_n \xrightarrow[d]{} \vec{X}$ con $\vec{X} \sim N(\vec{\mu};\Sigma) \Rightarrow \forall\ \vec{a} \in \mathbf{R}^k$ se cumple que:

$\vec{a}^t.\vec{X}_n \xrightarrow[d]{} \vec{Y}$ con $\vec{Y} \sim N\left(\vec{a}^t\vec{\mu};\vec{a}^t.\Sigma.\vec{a}\right)$.

Teorema Central del límite para vectores aleatorios:

Sea $\{\vec{X}_n\}$ una sucesión de vectores aleatorios independientes e igualmente distribuidos que verifican: $E\left(\vec{X}_i\right) = \vec{\mu}$ y $Var\left(\vec{X}_i\right) = \Sigma$ para todo $i \in \mathbf{N}$ (con Σ una matriz definida positiva).

Entonces:

$$\sqrt{n}\left(\frac{\sum_{i=1}^{n}\vec{X}_i}{n} - \vec{\mu}\right) \xrightarrow[d]{} N\left(\vec{0};\Sigma\right)$$