Unidad 5: Distribuciones de funciones de variables aleatorias
Teorema de Combinaciones lineales de vs as normales independientes y
Teorema Central del Límite!!!!!

Probabilidad y Estadística
Año 2020

Muestra aleatoria (m.a.):

Sea X una v. a. una m.a. de X es un conjunto de vs. as. X_1, $X_2,\ldots,X_n$. independientes e idénticamente distribuidas, iid, como X

Ejemplo: Sea X_1, $X_2,\ldots,X_n$. iid con media μ y varianza σ^2

$$E(\overline{X}) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}E\left(\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n}\sum_{i=1}^{n}\mu = \frac{1}{n}n\mu = \mu$$

$$V(\overline{X}) = V\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}V\left(\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} V(X_i) = \frac{1}{n^2}\sum_{i=1}^{n}\sigma^2 = \frac{1}{n^2}n\sigma^2 = \frac{\sigma^2}{n}$$

Ejemplo: Sea $X_1, X_2,\ldots,X_n$. iid con media μ y varianza σ^2

$$E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} E(X_i) = \sum_{i=1}^{n} \mu = n\mu$$

$$V\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} V(X_i) = \sum_{i=1}^{n} \sigma^2 = n\sigma^2$$

Teorema de las Combinaciones lineales de variables aleatorias con distribución Normal independientes

Sean $X_1, X_2, \ldots, X_n$ vs. as. independientes, tales que $\forall$ i = 1,...,n $X_i \sim N(\mu_i, \sigma_i)$, $\sigma_i < \infty$,

y sean $a_1, \ldots, a_n$ constantes reales, entonces la variable aleatoria

$$\sum_{i=1}^{n} a_i X_i \quad \sim \quad N\left(\sum_{i=1}^{n} a_i \mu_i,\ \sqrt{\sum_{i=1}^{n} a^2 \sigma_i^2}\right)$$

Ejemplo

Supongamos que el peso de un paquete lleno de yerba es una v.a. con distribución Normal con media 1000 gr. y varianza 120 gr². y que el peso del envase (paquete vacío) es una v.a. con distribución Normal N(10,1). ¿Cuál es la probabilidad de que el peso neto de yerba este entre 990 y 1000 gr.? Suponga T y E independientes

T = peso del paquete lleno ~ $N(1000, \sqrt{120})$

E = peso del envase ~ N(10, 1)

Y = peso neto   Y=T-E

E(Y) = E(T-E) = E(T) - E(E) = 1000 – 10 = 990

V(Y) = V(T-E) = V(T) + V(E) = 120 +1 =121

Por el Teorema de Combinaciones Lineales de las variables aleatorias normales independientes

$$Y \sim N(990, \sqrt{121}) = N(990{,}11)$$

Continuación ejemplo

$$P(990 < Y < 1000) = P\left(\frac{990 - 990}{\sqrt{121}} < \frac{Y - 990}{\sqrt{121}} < \frac{1000 - 990}{\sqrt{121}}\right)$$

$= \Phi(0.91) - \Phi(0) = 0.8186 - 0.5 = 0.3186$

Ejemplo

Suponga que el peso de ciertas piezas de acero sigue una distribución N(800, 20) en gr. Se eligen 4 piezas al azar. ¿Cuál es probabilidad de que el peso promedio sea menor que 810 gr.?

X_i = Peso de cada pieza, i = 1, 2, 3, 4. $X_i \sim N(800, 20)$

$$\overline{X} = \frac{1}{4}\sum_{i=1}^{4} X_i = \text{Pesopromedio}$$

Por el Teorema de Combinaciones Lineales

$$\overline{X} \sim N\left(800, {}^{20}\!\big/\!_{\sqrt{4}}\right)$$

$$P(\overline{X} < 810) = P\left(\frac{\overline{X} - 800}{20/\sqrt{4}} \leq \frac{810 - 800}{10}\right) \simeq P(Z \leq 1) = \Phi(1) = 0.8413.$$

Teorema Central del Límite (TCL)

Sean $X_1, \ldots, X_n$ vs. as. iid. como X tal que $EX = \mu$ y $V(X) = \sigma^2$. Luego $\mu = EX_i$ y $\sigma^2 = V(X_i)$, $\forall i = 1, 2, \ldots, n$ y supongamos que $0 < \sigma^2 < \infty$. Entonces:

Para todo $x \in \mathbf{R}$

$$\lim_{n \to \infty} P\left[\frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\sigma} \leq x \right] = \phi(x)$$

donde ϕ es la Función de distribución acumulada de la distribución Normal estándar.

Formas equivalentes de la Tesis del TCL

$$\sum X_i \approx N(n\mu, \sqrt{n}\sigma)$$

Cuando n tiene a infinito

$$\frac{\sum X_i - n\mu}{\sqrt{n}\sigma} \approx N(0, 1)$$

Cuando n tiene a infinito

$$\frac{\overline{X} - \mu}{\sigma / \sqrt{n}} \approx N(0,1)$$

Cuando n tiene a infinito

Ejemplo TCL

Un nuevo ascensor está diseñado para cargar como máximo un peso total de hasta 1900 kg. Superado este valor el ascensor no funciona. El peso de las personas alojadas en un hotel es una variable aleatoria con media 68 kg y desviación estándar de 25 kg. Suponga que 25 huéspedes toman un ascensor de este tipo y que los pesos de los mismos son variables aleatorias independientes. ¿Cuál es la probabilidad de que el ascensor no funcione?

X_i = Peso del i-ésimo huésped, i = 1, ..., 25

$E(X_i)=68$ $\sigma_{X_i} = 25$

X_i v. a. independientes X = Peso total

Por TCL

$$X = \sum_{i=1}^{25} X_i \simeq N(1700, 125)$$

Continuación ejemplo

$$P(X > 1900) = 1 - P(X \leq 1900) = 1 - P\left(\frac{X - 1700}{125} \leq \frac{1900 - 1700}{125}\right)$$

$$\cong 1 - \Phi(1.6) = 1 - 0.9452 = 0.055$$

Comparación entre Teoremas

	T.C.L.	Teorema de las Combinaciones Lineales
Hipótesis	$X_1, X_2, \ldots, X_n$ independientes	$X_1, X_2, \ldots, X_n$ independientes
	$X_1, X_2, \ldots, X_n$ idéntica distribución	$X_1, X_2, \ldots, X_n$ Normales no necesariamente idénticas
	$n \to \infty$	$n \geq 2$
Tesis	$\sum X_i$ ó $\overline{X}$	$\sum_{i=1}^{n} a_i X_i$
	Normal Aproximada	Normal Exacta

Aproximación de la distribución Binomial por una distribución Normal (np(1-p)>5)

Sea $X \sim b(n, p)$ entonces para $n \to \infty$ se sigue que:

$$X \simeq N(np, \sqrt{np(1-p)}) \quad \text{por TCL}$$

Estandarizando:

$$Z = \frac{X - np}{\sqrt{np(1-p)}} \sim N(0, 1)$$

Justificación: $X \sim b(n, p)$, sea $Y_i \sim B(p)$ con $i = 01,2, \ldots, n$, luego

$$X = \sum_{i=1}^{n} Y_i$$

Observaciones

- En la práctica si $np(1 - p) > 5$ la aproximación es aceptable.
- Siempre que se realiza esta aproximación (de discreta a continua) se debe corregir por continuidad. Por ejemplo:
  - $P(a \leq X \leq b) \equiv P(a - 0.5 \leq X \leq b + 0.5)$
  - $P(a \leq X < b) \equiv P(a - 0.5 \leq X < b - 0.5)$
  - $P(a < X \leq b) \equiv P(a + 0.5 < X \leq b + 0.5)$
  - $P(a < X < b) \equiv P(a + 0.5 < X < b - 0.5)$

Ejemplo: Aproximación de la distribución binomial a la Normal

El 20% de los chips de memoria fabricados en cierta planta son defectuosos. Se eligen 100 chips al azar.

a) ¿Cuál es la probabilidad de que a lo sumo 15 sean defectuosos?
b) ¿Cuál es la probabilidad de que exactamente 15 sean defectuosos?

- $X \sim b(100, 0.2)$ como $np(1-p) = 16 > 5$
- $X \simeq N(20, 4)$ (por Teorema Central del Límite)

Continuación ejemplo

a) $P(X \leq 15) = P(X \leq 15.5)$

$$= P\left(\frac{X-20}{4} \leq \frac{15.5-20}{4}\right) \cong \Phi(-1.13) = 1-\Phi(1.13) = 1-0.8708 = 0.1292$$

b) $P(X = 15) = P(14.5 \leq X \leq 15.5)$

$$= P\left(\frac{14.5-20}{4} \leq \frac{X-20}{4} \leq \frac{15.5-20}{4}\right) \cong \Phi(-1.13) - \Phi(-1.38)$$

$$= [1-\Phi(1.13)] - [1-\Phi(1.38)] = \Phi(1.38) - \Phi(1.13) = 0.9162 - 0.8708 = 0.0454$$

Aproximación de la distribución Poisson por una distribución Normal ($\lambda > 5$)

Sea $X \sim P(\lambda)$ entonces, para λ suficientemente grande

$$X \quad \simeq X(\lambda, \sqrt{\lambda}) \qquad \text{por TCL}$$

Estandarizando:

$$Z = \frac{X - \lambda}{\sqrt{\lambda}} \cong N(0, 1)$$

Justificación: $X \sim P(\lambda)$, sea $Y_i \sim P(1)$ con $i = 1, 2, \ldots, \lambda$ ($\lambda \in \mathbb{N}$ grande) luego

$$X = \sum_{i=1}^{\lambda} Y_i$$

Observaciones

- En la práctica si $\lambda > 5$ la aproximación es aceptable.
- También estamos aproximando una v. a. discreta por una continua, por lo tanto debe aplicarse corrección por continuidad. Por ejemplo:
  - $P(a \leq X \leq b) \equiv P(a - 0.5 \leq X \leq b + 0.5)$
  - $P(a \leq X < b) \equiv P(a - 0.5 \leq X < b - 0.5)$
  - $P(a < X \leq b) \equiv P(a + 0.5 < X \leq b + 0.5)$
  - $P(a < X < b) \equiv P(a + 0.5 < X < b - 0.5)$

Ejemplo: Aproximación de la distribución Poisson a la Normal

El número de rayos Gamma emitidos, por segundos, por cierta sustancia radiactiva es una variable aleatoria con distribución de Poisson con λ = 5.8. Si un instrumento queda fuera de registro cuando se emiten 12 o más rayos por segundo, ¿cuál es la probabilidad de que este instrumento quede fuera de operación, en cualquier segundo dado?

- $X \sim P(5.8)$
- $\lambda > 5 \Rightarrow X \stackrel{\sim}{-} N(5.8, \sqrt{5.8})$ por TCL

$P(X \geq 12) = 1 - P\ (X < 12) = 1 - P(X < 11.5)$

$$= 1 - P\left(\frac{X - 5.8}{\sqrt{5.8}} < \frac{11.5 - 5.8}{\sqrt{5.8}}\right) \cong 1 - \Phi(2.37) = 1 - 0.9911 = 0.0009$$

Resumen de las Aproximaciones

$b(n, p)$

$np(1-p) > 5$

$\mu = np$

$\sigma = \sqrt{np(1-p)}$

$P(\lambda)$

$\lambda > 5$

$\mu = \lambda$

$\sigma = \sqrt{\lambda}$

$N(\mu, \sigma)$