Desigualdad de Chebyshev:

Para calcular la probabilidad de un evento descripto en términos de una v.a. *X* es necesario conocer la distribución de la v.a. La desigualdad de Chebyshev provee una cota que no depende de la distribución sino sólo de la varianza de *X*.

Proposición: Sea *X* una v.a. con E(*X*) = μ y V(*X*)= $\sigma^2 < \infty$, entonces

$$\forall\, \varepsilon > 0, \qquad P\left(\left|X-\mu\right| > \varepsilon\right) \le \frac{\sigma^2}{\varepsilon^2}$$

Dem: Lo haremos para el caso continuo. La demostración para el caso discreto es similar.

$$\sigma^2 = E\left((X-\mu)^2\right) = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\,dx =$$

$$= \int_{\{x/|x-\mu|>\varepsilon\}} (x-\mu)^2 f(x)\,dx + \int_{\{x/|x-\mu|\le\varepsilon\}} (x-\mu)^2 f(x)\,dx \ge$$

$$\ge \int_{\{x/|x-\mu|>\varepsilon\}} (x-\mu)^2 f(x)\,dx \quad \ge \int_{\{x/|x-\mu|>\varepsilon\}} \varepsilon^2 f(x)\,dx = \varepsilon^2 P\left(\left|X-\mu\right| > \varepsilon\right)$$

Entonces,

$$\frac{\sigma^2}{\varepsilon^2} \ge P\left(\left|X-\mu\right| > \varepsilon\right)$$

como queríamos demostrar.

Observación: La cota que provee la desigualdad de Chebyshev puede ser grosera o, peor aún, no informativa, por ejemplo, si $\varepsilon \le \sigma^2$.

Ejemplo: Sea *X* ~ *U*(0,10), entonces E(*X*) = 5 y V(*X*)= 100/12.

Aplicando la desigualdad de Chebyshev,

$$P\left(\left|X-5\right| > 4\right) \le \frac{\sigma^2}{16} = \frac{100/12}{16} \cong 0.52$$

pero, si calculamos en forma exacta esa probabilidad,

$$P\left(\left|X-5\right| > 4\right) = 1 - P\left(\left|X-5\right| \le 4\right) = 1 - P\left(-4 \le X-5 \le 4\right) = 1 - P(1 \le X \le 9) =$$

$$= 1 - F_X(9) + F_X(1) = 1 - \frac{9}{10} + \frac{1}{10} = 0.20$$

Otras formas de la desigualdad de Chebyshev: Otras formas equivalentes de la desigualdad de Chebyshev son las siguientes.

a) $\forall\, \varepsilon > 0, \quad P(|X-\mu| \le \varepsilon) \ge 1 - \dfrac{\sigma^2}{\varepsilon^2}$

b) $\forall\; k \ge 1, \quad P(|X-\mu| > k\sigma) \le \dfrac{1}{k^2}$

(En realidad, es cierto para todo $k > 0$, pero si $k < 1$ es trivial)

c) $\forall\; k \ge 1, \quad P(|X-\mu| \le k\sigma) \ge 1 - \dfrac{1}{k^2}$

Las dos últimas formas muestran como el desvío standard mide el grado de "concentración" de la distribución alrededor de μ = E(X).

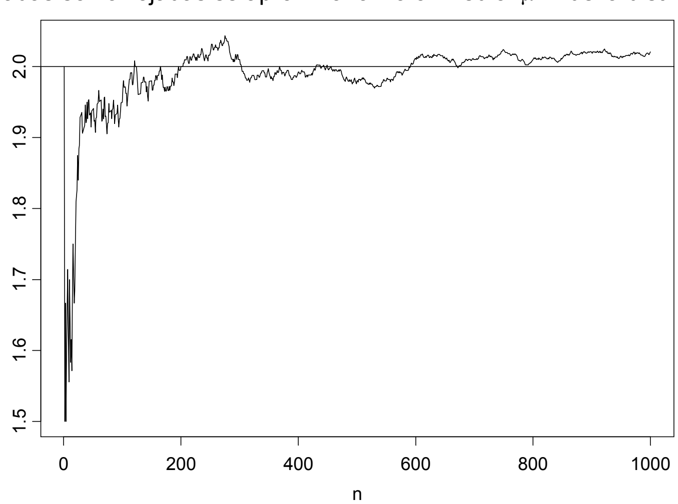
Ley de los Grandes Números:

Sea X una v.a. con función de densidad $f(x)$ o función de probabilidad puntual $p(x)$ y con E(X) = μ. Supongamos que se desea "estimar" μ. Como hemos visto que la esperanza de una v.a. se puede pensar como un promedio de sus valores, parece razonable estimarla mediante el promedio de valores observados de X. Por supuesto que en una situación real sólo tendremos un número finito de observaciones y nos preguntamos: usando sólo un número finito de valores de X, ¿puede hacerse inferencia confiable respecto de E(X)?

La respuesta es SI y se demuestra a través de la Ley de los Grandes Números que nos dice que el promedio $\overline{X}$ converge a μ cuando el número de observaciones (o tamaño de la muestra) tiende a infinito.

Observemos lo que sucede en la siguiente figura.

Figura 1:Comportamiento asintótico del promedio muestral. El promedio del múmero observado de caras, $\overline{x}$, cuando 4 monedas equilibradas son arrojadas se aproxima al valor medio μ=2 de la distribución.



¿En qué sentido converge $\overline{X}$ a μ?

Sea (X_n) ($n \geq 1$) una sucesión de variables aleatorias, diremos que X_n *converge en probabilidad* a la v.a. X y lo notaremos $X_n \xrightarrow{p} X$, si

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

<u>Ley de los Grandes Números</u>: Sean X_1, X_2, ... v.a. independientes e idénticamente distribuidas (muestra aleatoria) con $E(X) = \mu$ y $V(X) = \sigma^2 < \infty$, entonces

$$\overline{X}_n \xrightarrow{p} \mu$$

siendo $\overline{X}_n = \dfrac{\sum_{i=1}^{n} X_i}{n}$ el denominado promedio muestral.

Dem: Sabemos que $E(\overline{X}_n) = \mu$ y $V(\overline{X}_n) = \dfrac{\sigma^2}{n}$, entonces aplicando la desigualdad de Chebyshev,

$$P\left(\left|\overline{X}_n - \mu\right| > \varepsilon\right) \leq \frac{V(\overline{X}_n)}{\varepsilon^2} = \frac{\sigma^2}{n\,\varepsilon^2} \qquad \forall\, \varepsilon > 0$$

y, por lo tanto

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

Luego, $\overline{X}_n \xrightarrow{p} \mu$, como queríamos demostrar.

<u>Versión Bernoulli de la Ley de los Grandes Números</u>: Consideremos n repeticiones independientes de un experimento aleatorio y sea A un suceso con probabilidad $P(A) = p$, constante en las n repeticiones. Si llamamos n_A a la frecuencia absoluta de A (número de veces que ocurre A en las n repeticiones) y $f_A = n_A / n$ a la frecuencia relativa, entonces

$$f_A \xrightarrow{p} p$$

es decir,

$$P\left(\left|f_A - p\right| > \varepsilon\right) \xrightarrow[n\to\infty]{} 0 \qquad \forall\, \varepsilon > 0$$

Dem: Como n_A ~ Bi(n,p) con p = P(A), entonces E(n_A) = n p y V(n_A) = n p (*1*-p). Luego

$$E(f_A) = E\left(\frac{n_A}{n}\right) = p \qquad\qquad V(f_A) = V\left(\frac{n_A}{n}\right) = \frac{p(1-p)}{n}$$

y, aplicando la desigualdad de Chebyshev,

$$P\left(\left|f_A - p\right| > \varepsilon\right) \leq \frac{V(f_A)}{\varepsilon^2} = \frac{p(1-p)}{n\,\varepsilon^2} \qquad \forall\, \varepsilon > 0$$

Luego,

$$\lim_{n\to\infty} P\left(\left|f_A - p\right| > \varepsilon\right) \leq \lim_{n\to\infty} \frac{p(1-p)}{n\,\varepsilon^2} = 0 \qquad \forall\, \varepsilon > 0$$

como queríamos demostrar.

<u>Ejemplo</u>: ¿Cuántas repeticiones del experimento deberían hacerse para que la frecuencia relativa difiera de p en menos de 0.01 con probabilidad mayor o igual que 0.95?

En este caso, ε = 0.01 y queremos encontrar n tal que

$$P\left(\left|f_A - p\right| < 0.01\right) \geq 0.95$$

Pero, dado que

$$P\left(\left|f_A - p\right| < 0.01\right) \geq 1 - \frac{p(1-p)}{n\,(0.01)^2}$$

$$1 - \frac{p(1-p)}{n\,(0.01)^2} \geq 0.95 \Leftrightarrow \frac{p(1-p)}{n\,(0.01)^2} \leq 0.05 \Leftrightarrow n \geq \frac{p(1-p)}{(0.01)^2(0.05)}$$

El valor mínimo de n depende de p y es máximo cuando p = 0.50. Por ejemplo,

$p = 0.50 \;\Rightarrow\; n \geq 50000$
$p = 0.10 \;\Rightarrow\; n \geq 18500$
$p = 0.01 \;\Rightarrow\; n \geq 1980$

Distribución de la suma de variables aleatorias independientes: En general es difícil calcular la distribución de la suma o de una combinación lineal de *n* v.a. independientes, aún cuando tengan la misma distribución. Sin embargo, en algunos casos la distribución de la suma o de combinaciones lineales es conocida. Recapitulemos algunos resultados.

a) Si $X_1, X_2, ..., X_n$ son v.a. independientes tales que $X_i \sim Bi(n_i, p)$, entonces $\sum_{i=1}^{n} X_i \sim Bi\left(\sum_{i=1}^{n} n_i, p\right)$.

En particular, si $X_i \sim Bi(1, p) \quad \forall\, i$, entonces $\sum_{i=1}^{n} X_i \sim Bi(n, p)$.

b) Si $X_1, X_2, ..., X_n$ son v.a. independientes tales que $X_i \sim P(\lambda_i)$, entonces $\sum_{i=1}^{n} X_i \sim P\left(\sum_{i=1}^{n} \lambda_i\right)$.

c) Si $X_1, X_2, ..., X_n$ son v.a. i.i.d. tales que $X_i \sim G(p)$, entonces $\sum_{i=1}^{n} X_i \sim BN(n, p)$.

d) Si $X_1, X_2, ..., X_n$ son v.a. i.i.d. tales que $X_i \sim \mathcal{E}(\lambda)$, entonces $\sum_{i=1}^{n} X_i \sim \Gamma(n, \lambda)$.

e) Si $X_1, X_2, ..., X_n$ son v.a. independientes tales que $X_i \sim \Gamma(n_i, \lambda)$, entonces $\sum_{i=1}^{n} X_i \sim \Gamma\left(\sum_{i=1}^{n} n_i, \lambda\right)$.

f) Si $X_1, X_2, ..., X_n$ son v.a. independientes tales que $X_i \sim N(\mu_i, \sigma_i^2)$ y $a_1, a_2, ..., a_n$ son números reales, entonces $\sum_{i=1}^{n} a_i X_i \sim N\left(\sum_{i=1}^{n} a_i \mu_i, \sum_{i=1}^{n} a_i^2 \sigma_i^2\right)$.

En particular, si $X_1, X_2, ..., X_n$ son v.a. i.i.d. tales que $X_i \sim N(\mu, \sigma^2)$, entonces

$$\overline{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \quad \text{y} \quad T = \sum_{i=1}^{n} X_i \sim N(n\mu, n\sigma^2).$$

Dem: Todos estos resultados pueden demostrarse fácilmente usando funciones generadoras de momentos. Como ejemplo, demostremos la propiedad e), es decir que si

$X_1, X_2, ..., X_n$ son v.a. independientes tales que $X_i \sim \Gamma(n_i, \lambda)$, entonces $\sum_{i=1}^{n} X_i \sim \Gamma\left(\sum_{i=1}^{n} n_i, \lambda\right)$.

Por ser las X_i v.a. independientes, la función generadora de la suma es el producto de las funciones generadoras, entonces

$$M_{\sum_{i=1}^{n} X_i}(t) = \prod_{i=1}^{n} M_{X_i}(t) = \prod_{i=1}^{n} \left(\frac{\lambda}{\lambda - t}\right)^{n_i} = \left(\frac{\lambda}{\lambda - t}\right)^{\sum_{i=1}^{n} n_i} \Rightarrow \sum_{i=1}^{n} X_i \sim \Gamma\left[\sum_{i=1}^{n} n_i, \lambda\right]$$

como queríamos demostrar.

Veremos ahora que, cuando las v.a. no son normales, la distribución normal resulta una buena aproximación para la distribución de $\overline{X}$ y T.

Teorema Central del Límite: Sean $X_1, X_2, ...$ v.a. i.i.d con $E(X_i) = \mu$ y $V(X_i) = \sigma^2 < \infty$, entonces si *n* es suficientemente grande,

$$\frac{T - n\mu}{\sqrt{n}\,\sigma} \overset{(a)}{\sim} N(0,1) \qquad \frac{\sqrt{n}\left(\overline{X} - \mu\right)}{\sigma} \overset{(a)}{\sim} N(0,1)$$

o, dicho de otro modo,

$$\frac{T - n\mu}{\sqrt{n}\,\sigma} \xrightarrow{d} Z \sim N(0,1) \qquad \frac{\sqrt{n}\,(\overline{X} - \mu)}{\sigma} \xrightarrow{d} Z \sim N(0,1)$$

donde la convergencia en distribución ($\xrightarrow{d}$) se interpreta en el siguiente sentido:

$$P\left(\frac{T - n\mu}{\sqrt{n}\,\sigma} \le a\right) \cong \Phi(a) \qquad P\left(\frac{\sqrt{n}\,(\overline{X} - \mu)}{\sigma} \le a\right) \cong \Phi(a)$$

es decir, que las funciones de distribución convergen a la función de distribución normal standard.

Dem: Lo demostraremos bajo la hipótesis de que la función generadora de momentos de X_i, $M_{X_i}(t)$ existe y es finita.

Supongamos inicialmente que μ = 0 y σ^2 = 1. En este caso, la función generadora de momentos de $\frac{T}{\sqrt{n}}$ está dada por

$$M_{T/\sqrt{n}}(t) = M_T(t/\sqrt{n}) = M_{\sum_{i=1}^{n} X_i}(t/\sqrt{n}) = \prod_{i=1}^{n} M_{X_i}(t/\sqrt{n}) = \left(M_{X_i}(t/\sqrt{n})\right)^n$$

por ser las X_i independientes. Sea $L(t) = \ln\left(M_{X_i}(t)\right)$, entonces

$$L(0) = \ln\left(M_{X_i}(0)\right) = \ln(1) = 0$$

$$L'(0) = \left.\frac{\partial \ln M_{X_i}(t)}{\partial t}\right|_{t=0} = \left.\frac{M'_{X_i}(t)}{M_{X_i}(t)}\right|_{t=0} = \frac{M'_{X_i}(0)}{M_{X_i}(0)} = \frac{\mu}{1} = \mu = 0$$

$$L''(0) = \left.\frac{\partial^2 \ln M_{X_i}(t)}{\partial t^2}\right|_{t=0} = \left.\frac{M''_{X_i}(t)\cdot M_{X_i}(t) - \left[M'_{X_i}(t)\right]^2}{\left[M_{X_i}(t)\right]^2}\right|_{t=0} = \frac{M''_{X_i}(0)M_{X_i}(0) - \left[M'_{X_i}(0)\right]^2}{\left[M_{X_i}(0)\right]^2} = E(X_i^2) = 1$$

Ahora, para probar el teorema, demostraremos que $M_{T/\sqrt{n}}(t) \rightarrow e^{t^2/2}$ o equivalentemente, que $nL(t/\sqrt{n}) \rightarrow t^2/2$. Aplicando la regla de L'Hospital dos veces,

$$\lim_{n\to\infty} \frac{L(t/\sqrt{n})}{1/n} = \lim_{n\to\infty} \frac{-L'(t/\sqrt{n})t\, n^{-3/2}}{-2\, n^{-2}} = \lim_{n\to\infty} \frac{L'(t/\sqrt{n})t}{2\, n^{-1/2}} =$$

$$= \lim_{n\to\infty} \frac{-L''(t/\sqrt{n})\, t^2\, n^{-3/2}}{-2\, n^{-3/2}} = \lim_{n\to\infty} \frac{L''(t/\sqrt{n})\, t^2}{2} = \frac{t^2}{2}.$$

por lo tanto hemos probado el Teorema Central del Límite para μ = 0 y σ^2 = 1. El caso general resulta considerando las v.a. standarizadas $\frac{X_i - \mu}{\sigma} = X_i^*$ y teniendo en cuenta las propiedades de las funciones generadoras de momentos.

<u>Observación</u>: ¿Qué significa n suficientemente grande? ¿Cómo sabemos si la aproximación es buena? El tamaño de muestra requerido para que la aproximación sea

razonable depende de la forma de la distribución de las X_i . Mientras más simétrica y acampanada sea, más rápidamente se obtiene una buena aproximación.

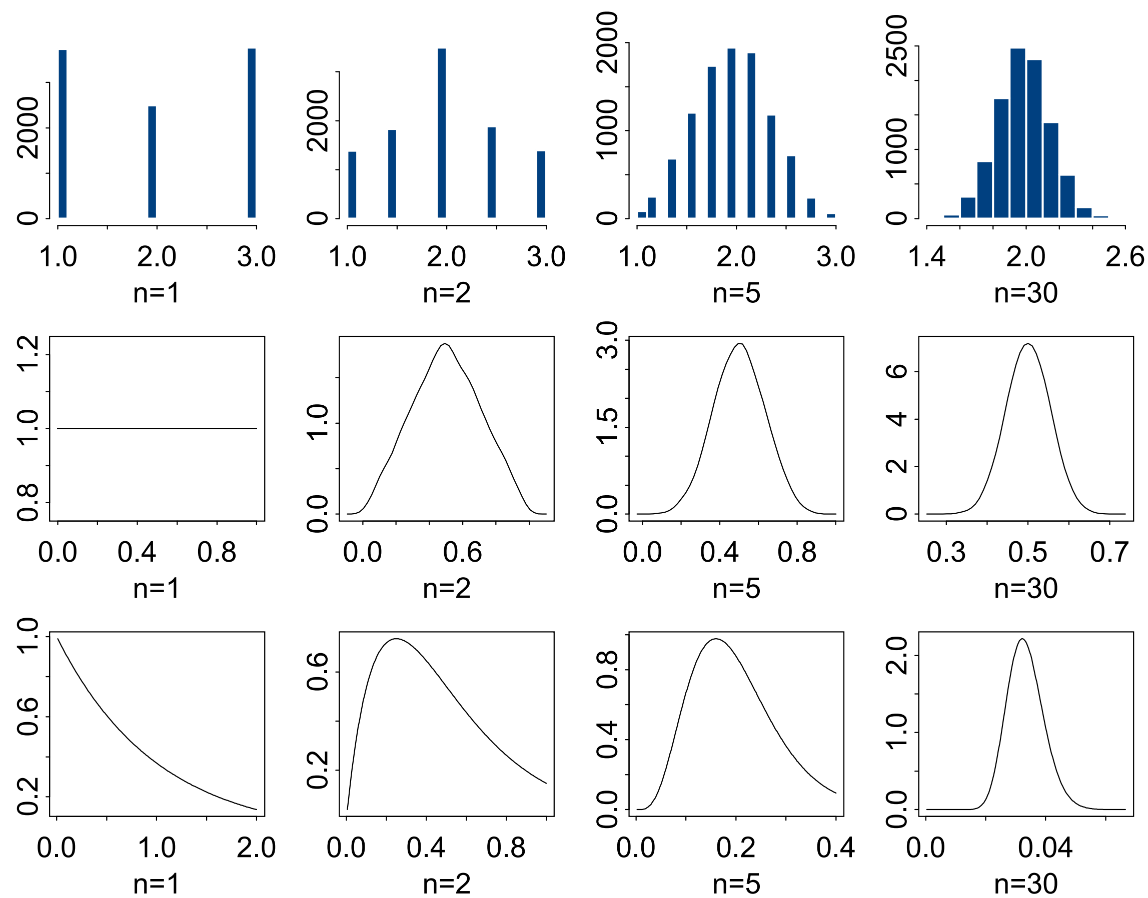


Figura 2: Distribución de $\overline{x}$ para distintas distribuciones cuando n=2, 5 y 30.
a) Distribución discreta, b) Distribución Uniforme, c) Distribución Exponencial

<u>Ejemplo</u>: Al sumar números, una calculadora aproxima cada número al entero más próximo. Los errores de aproximación se suponen independientes y con distribución $U(-0.5,0.5)$.

a) Si se suman 1500 números, ¿cuál es la probabilidad de que el valor absoluto del error total exceda 15?

Si llamamos X_i al error correspondiente al i-ésimo sumando, el error total es $T_{1500} = \sum_{i=1}^{1500} X_i$.

Entonces,

$$P(|T_{1500}| > 15) = 1 - P(|T_{1500}| \leq 15) = 1 - P(-15 \leq T_{1500} \leq 15) =$$

$$= 1 - P\left(\frac{-15}{\sqrt{1500/12}} \le \frac{T_{1500}}{\sqrt{1500/12}} \le \frac{15}{\sqrt{1500/12}} \right) \cong 1 - \Phi\left(\frac{15}{\sqrt{1500/12}} \right) + \Phi\left(\frac{-15}{\sqrt{1500/12}} \right) =$$

$$= 1 - \Phi(1.34) + \Phi(-1.34) = 0.18$$

Hemos usado que $E(X_i) = 0$ y $V(X_i) = \dfrac{1}{12}$ y por lo tanto $E(T_{1500}) = 0$ y $V(T_{1500}) = \dfrac{1500}{12}$.

b) ¿Cuántos números pueden sumarse a fin de que el valor absoluto del error total sea menor o igual que 10 con probabilidad mayor o igual que 0.90?

Buscamos el valor de *n* tal que

$$P(|T_n| \le 10) \ge 0.90$$

$$P(|T_n| \le 10) \ge 0.90 \Leftrightarrow P(-10 \le T_n \le 10) \ge 0.90 \Leftrightarrow P\left(\frac{-10}{\sqrt{n/12}} \le T_n \le \frac{10}{\sqrt{n/12}} \right) \ge 0.90$$

Aplicando la aproximación normal, debemos hallar *n* tal que

$$\Phi\left(\frac{10}{\sqrt{n/12}} \right) - \Phi\left(\frac{-10}{\sqrt{n/12}} \right) \ge 0.90 \Leftrightarrow 2\Phi\left(\frac{10}{\sqrt{n/12}} \right) - 1 \ge 0.90 \Leftrightarrow \Phi\left(\frac{10}{\sqrt{n/12}} \right) \ge 0.95$$

$$\Leftrightarrow \frac{10}{\sqrt{n/12}} \ge 1.64 \Leftrightarrow \sqrt{n} \le 21.12 \Leftrightarrow n \le 446$$

es decir, que se pueden sumar a lo sumo 446 números para que el valor absoluto del error total sea menor o igual que 10 con probabilidad mayor o igual que 0.90.

<u>Aproximación de la distribución binomial por la normal</u>: Sea $X \sim Bi(n,p)$, entonces X es el número de éxitos en *n* repeticiones de un experimento binomial con probabilidad de éxito igual a *p*, y X / n es la proporción muestral de éxitos.

Definamos las siguientes variables aleatorias

$$X_i = \begin{cases} 1 & \text{si se obtuvo Éxito en la repetición } i \\ 0 & \text{si se obtuvo Fracaso en la repetición } i \end{cases}$$

para $i = 1, ..., n$. Estas v.a. son independientes, $X_i \sim Bi(1, p)$ $\forall\, i$ y $X = \sum_{i=1}^{n} X_i$.

Aplicando el Teorema Central del Límite, si *n* es suficientemente grande,

$$X \overset{(a)}{\sim} N\big(np, np(1-p)\big) \qquad\qquad \frac{X}{n} \overset{(a)}{\sim} N\left(p, \frac{p(1-p)}{n}\right)$$

Se considera que la aproximación es buena si $n\,p \geq 5$ y $n\,(1\text{-}p) \geq 5$.

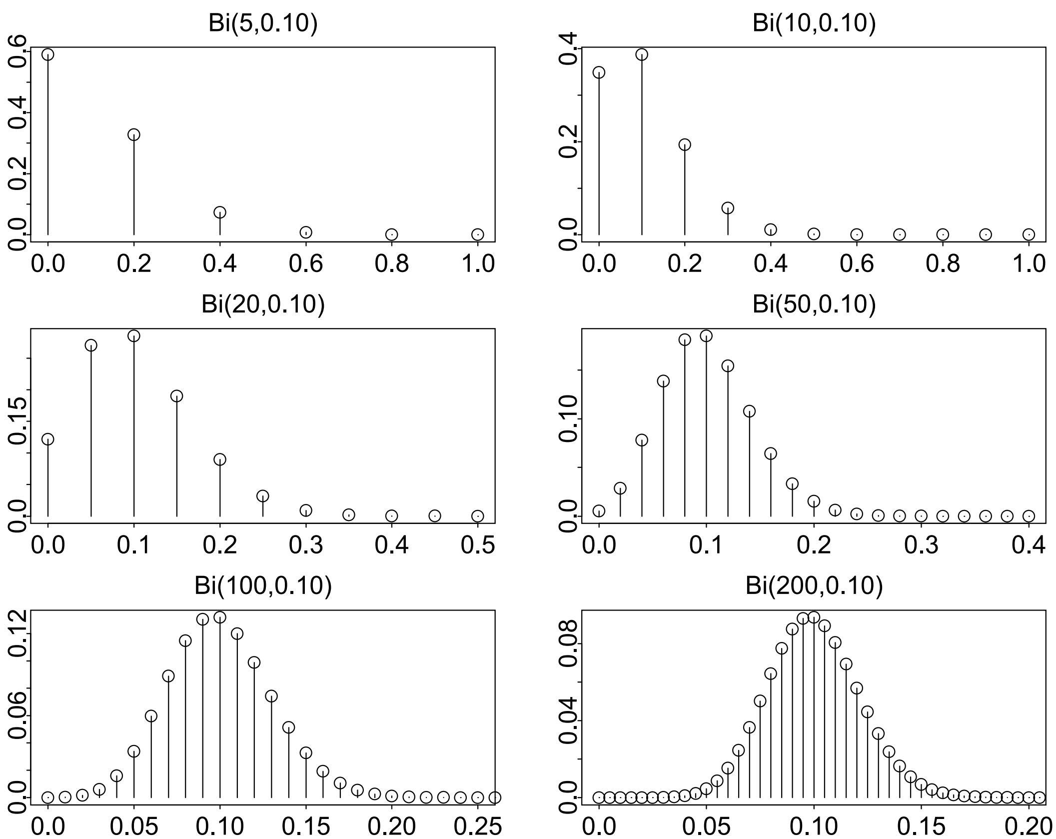


Figura 3: Distribución de $\dfrac{X}{n}$

<u>Corrección por continuidad</u>: Cuando se aproxima una distribución discreta por una continua, como es el caso de la aproximación de la distribución binomial por la normal, es necesario efectuar una corrección. Consideremos el siguiente ejemplo:

Sea $X \sim$ Bi (100, 0.6) y calculemos en forma aproximada P($X \leq 50$) y P($X \geq 51$).

Si aplicamos directamente el TCL, obtenemos:

$$P(X \leq 50) = P\left(\frac{X-60}{\sqrt{24}} \leq \frac{50-60}{\sqrt{24}}\right) \cong \Phi(-2.04) = 0.021$$

$$P(X \geq 51) = P\left(\frac{X-60}{\sqrt{24}} \geq \frac{51-60}{\sqrt{24}}\right) \cong 1 - \Phi(-1.84) = 0.967$$

Si bien, P($X \leq 50$) + P($X \geq 51$) = 1, los valores aproximados no satisfacen esta restricción. Para evitar este problema, se efectúa la siguiente corrección, denominada *corrección por continuidad*,

$$P(X \leq 50) = P(X \leq 50.5) = P\left(\frac{X-60}{\sqrt{24}} \leq \frac{50.5-60}{\sqrt{24}}\right) \cong \Phi(-1.94) = 0.026$$

$$P(X \geq 51) = P(X \geq 50.5) = P\left(\frac{X-60}{\sqrt{24}} \geq \frac{50.5-60}{\sqrt{24}}\right) \cong 1-\Phi(-1.94) = 0.974$$

En general, cuando la v.a. es discreta y $x_i - x_{i-1} = 1$, la corrección se realiza en la forma:

$$P(X \leq a) = P(X \leq a+0.5)$$
$$P(X \geq a) = P(X \geq a-0.5)$$

Si la distancia entre dos valores sucesivos de X es $k > 1$, ¿cómo aplicaría la corrección por continuidad?

Ejemplo: Sea $X \sim Bi(60,1/3)$. Calcular en forma aproximada la probabilidad de que X sea mayor o igual que 25.

$$P(X \geq 25) = P(X \geq 24.5) = P\left(\frac{X-60\cdot\frac{1}{3}}{\sqrt{60\cdot\frac{1}{3}\cdot\frac{2}{3}}} \geq \frac{24.5-60\cdot\frac{1}{3}}{\sqrt{60\cdot\frac{1}{3}\cdot\frac{2}{3}}}\right) \cong 1-\Phi(1.23) = 0.11$$

Otras aplicaciones del Teorema Central del Límite:

a) Sean $X_1, X_2, ..., X_n$ v.a. i.i.d. con distribución Poisson de parámetro λ, entonces

$$\sum_{i=1}^{n} X_i \sim P(n\lambda)$$

Por lo tanto, cualquier v.a. con distribución de Poisson con parámetro suficientemente grande puede ser aproximada por la distribución normal.

b) Sean $X_1, X_2, ..., X_n$ v.a. independientes con distribución Gamma de parámetros n_i y λ, o sea $X_i \sim \Gamma(n_i, \lambda)$ entonces

$$\sum_{i=1}^{n} X_i \sim \Gamma\left(\sum_{i=1}^{n} n_i, \lambda\right)$$

Por lo tanto, cualquier v.a. con distribución $\Gamma(m, \lambda)$ con parámetro *m* suficientemente grande puede ser aproximada por la distribución normal.

Una aplicación de suma de v.a. independientes y generación de números al azar:

Recordemos que un **proceso de Poisson** permite modelar una situación en la que los eventos ocurren a lo largo del tiempo (o espacio, volumen, etc.).

Hemos visto, que bajo ciertos supuestos, si definimos la variable

X_t = cantidad de eventos que ocurren en el intervalo [0,t]

entonces $X_t \sim P(\lambda\, t)$, donde λ es la tasa media de ocurrencias o intensidad del proceso.

También hemos mencionado que, si denotamos

- T_1 = tiempo que transcurre entre que empezamos a medir y el momento en que ocurre el primer evento

- T_2 = tiempo que transcurre entre el primer evento y el segundo evento.

y, en general,

- T_i = tiempo que transcurre entre el (*i-1*)- ésimo evento y el *i*-ésimo evento ($i \in N$)

las T_i son variables aleatorias independientes y con distribución exponencial, todas con el mismo parámetro λ.

Es claro que, si a uno le interesara el tiempo que transcurre desde el inicio hasta la k-ésima ocurrencia, esta variable aleatoria podría expresarse como

$$\sum_{i=1}^{k} T_i$$

Veamos la recíproca, es decir, veamos como podemos construir un proceso de Poisson a partir de v.a. i.i.d. con distribución exponencial.

<u>Proposición</u>: Sean $W_1, W_2, ..., W_k, ...$ v.a. independientes con distribución *E*(1). Consideremos el siguiente proceso. Comenzamos a medir el tiempo en t = 0 y consideramos que ocurre el primer evento en el instante W_1, el segundo en el instante $W_1 + W_2$, y en general el k-ésimo evento en el instante $W_1 + W_2 + ... + W_k$. Si para t > 0, definimos la variable aleatoria

X_t = cantidad de eventos que ocurren en el intervalo [0,t]

entonces X_t es una variable discreta y su distribución es $P(t)$.

Dem: Sea $k \in N \cup \{0\}$ y consideremos el evento $[X_t \geq k]$. Observemos que

$[X_t \geq k]$ $\Longleftrightarrow$ hubo k ó más eventos en el intervalo [0,t] $\Longleftrightarrow$ hubo por lo menos k eventos en el intervalo [0,t] $\Longleftrightarrow$ $\sum_{i=1}^{k} W_i \leq t$

Calculemos la probabilidad de dicho evento:

$$P(X_t \geq k) = P\left(\sum_{i=1}^{k} W_i \leq t\right)$$

Como las $W_1, W_2, ..., W_k, ...$ son variables aleatorias independientes y con distribución E(1)=Γ(1,1) , entonces

$$\sum_{i=1}^{k} W_i \sim \Gamma(k,1)$$

y por lo tanto

$$P\left(\sum_{i=1}^{k} W_i \leq t\right) = \int_{-\infty}^{t} f_S(s)ds$$

con $S = \sum_{i=1}^{k} W_i \sim \Gamma$ (k,1) y en consecuencia $f_S(s) = \frac{1}{(k-1)!} s^{k-1} e^{-s} I_{(0,+\infty)}(s)$. Entonces,

$$P\left(\sum_{i=1}^{k} W_i \leq t\right) = \int_{0}^{t} \frac{1}{(k-1)!} s^{k-1} e^{-s} ds$$

Llamemos

$$A_k(t) = \int_{0}^{t} \frac{1}{(k-1)!} s^{k-1} e^{-s} ds$$

a la función de distribución acumulada de una Γ(k,1). Integrando por partes una vez, si consideramos

$$u = \frac{s^{k-1}}{(k-1)!} \quad u' = \frac{(k-1)s^{k-2}}{(k-1)!} = \frac{s^{k-2}}{(k-2)!}$$

y

$$v' = e^{-s} \qquad v = -\mathrm{e}^{-s}$$

obtenemos

$$A_k(t) = \int_0^t \frac{1}{(k-1)!} s^{k-1} e^{-s} ds = \frac{-1}{(k-1)!} s^{k-1} e^{-s} \Big|_0^t + \int_0^t \frac{1}{(k-2)!} s^{k-2} e^{-s} ds$$

$$= \frac{-1}{(k-1)!} t^{k-1} e^{-t} + A_{k-1}(t)$$

$$= \frac{-1}{(k-1)!} t^{k-1} e^{-t} + \frac{-1}{(k-2)!} t^{k-2} e^{-t} + A_{k-2}(t)$$

Finalmente, por inducción, después de M pasos obtenemos

$$A_k(t) = ... = -e^{-t} \sum_{i=k-M}^{k-1} \frac{t^i}{i!} + A_{k-M}(t)$$

Como

$$A_1(t) = \int_0^t e^{-s} ds = -e^{-t} + 1$$

resulta

$$A_k(t) = -e^{-t} \sum_{i=1}^{k-1} \frac{t^i}{i!} + A_1(t) = -e^{-t} \sum_{i=0}^{k-1} \frac{t^i}{i!} + 1$$

y por lo tanto

$$P(X_t \geq k) = P\left(\sum_{i=1}^{k} W_i \leq t \right) = A_k(t) = 1 - \sum_{i=0}^{k-1} \frac{t^i}{i!} e^{-t}$$

Si tomamos el complemento resulta

$$P(X_t < k) = P(X_t \leq k-1) = \sum_{i=0}^{k-1} \frac{t^i}{i!} e^{-t}$$

que corresponde a la función de distribución acumulada de una variable con distribución *P*(t), tal como queríamos demostrar.

Este resultado es muy útil para generar variables aleatorias con distribución de Poisson a partir de exponenciales, a las que podemos generar fácilmente a partir de U(0,1).

Supongamos que deseamos generar una variable aleatoria *X* con distribución P(λ). Para ello basta utilizar la proposición anterior tomando t = λ. Podemos describir el algoritmo de la siguiente forma:

Paso 1: generamos una v.a. W_1 con distribución $E(1)$.

Paso 2: chequeamos si $W_1 \leq t$. Si ésto ocurre, continuamos con el paso siguiente. Si, en cambio, $W_1 \geq t$ terminamos y $X = 0$.

Paso 3: generamos una v.a. W_2 con distribución $E(1)$, independiente de W_1.

Paso 4: chequeamos si $W_1 + W_2 \leq t$. Si ésto ocurre, continuamos con el paso siguiente. Si no, terminamos y $X = 1$.

Paso 2k-1: generamos una v.a. W_k con distribución $E(1)$, independiente de $W_1, W_2, \ldots, W_{k-1}$.

Paso 2k: chequeamos si $W_1 + W_2 + \ldots + W_k \leq t$. Si ésto ocurre seguimos, si no terminamos y $X = k$.