Introducción a la Probabilidad y la Estadística

Martes y Jueves Aula B17

Dra Ana Georgina Flesia

Desigualdad de Chebyshev

Lema

Sea X una variable aleatoria positiva con media μ, y sea t un número real positivo. Entonces

$$P(X \geq t) \leq \frac{E(X)}{t}$$

Desigualdad de Chebyshev

Sea X una variable aleatoria con media μ y varianza σ^2. Entonces, para cada $t > 0$

$$P(|X - \mu| \geq t) \leq \frac{\sigma^2}{t^2}$$

Desigualdad de Chebyshev

1. La proposición anterior dice que si σ^2 es muy pequeño, existe una probabilidad muy alta de que X no se desvíe mucho de μ.
2. Podemos realizar una cota en función de la varianza.

Lema

Sea X una variable aleatoria con media μ y varianza σ^2. Entonces,

$$\begin{aligned} P(|X-\mu| \geq k\sigma) &\leq \frac{1}{k^2} \\ P(|X-\mu| < k\sigma) &\geq 1-\frac{1}{k^2} \end{aligned}$$

Ejemplo

Un fabricante de tornillos sabe que el 5% de su producción es defectuosa. Por ello garantiza que si una entrega de 10000 piezas tiene más de a defectuosos se reembolsará el dinero pagado por el cargamento. ¿Cuán chico puede el fabricante escoger el valor a para que no deba devolver el dinero del cargamento más del 1% de las veces?

Resolución

1. Supongamos que el lote de 10000 tornillos es una muestra aleatoria de la población (infinita) de tornillos que produce la fábrica.
2. Sea $p = 0.05$ la probabilidad de elegir un tonillo defectuoso de la población. Entonces X, el número de defectuosos del lote de 10000 tornillos, tiene distribución Binomial con media np y varianza $np(1-p)$, y el evento A= *devolver el cargamento* puede ser pensado como

$$A = (X > a) = (X - np > a - np)$$

Resolución

1. Queremos escoger a tal que

$$P(\text{ devolver el cargamento }) = P(X > a) \leq 0.01$$

2. Sabemos por Chebyshev que

$$P(|(X - np| > t) \leq \frac{np(1-p)}{t^2} < 0.01 \text{ si } t \geq \sqrt{\frac{10000 * 0.0475}{0.01}}$$

3. Entonces $t \geq 218$ garantiza que

$$P(A) = P(X - np > a - np) \leq P(|X - np| > t) \leq 0.01$$

por lo cual $a - np = t \geq 218$ implica que
$a = 218 + np = 218 + 500 = 718$ es un valor conservativo para a.

Ley de los grandes números

1. Hemos dicho al comienzo del curso que si una moneda honesta se tira muchas veces, la proporción de caras obtenida estará cerca de 1/2.
2. La ley de los grandes números formaliza esta afirmación empírica.
3. Los resultados de los tiros consecutivos de la moneda son variables aleatorias independientes X_i que toman valores 0 o 1 según la i-esima repetición es cara o número, y la proporción de caras en los n tiros es

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

Ley de los grandes números

Definición

Sea $\{X_n\}$ una sucesión de variables aleatorias. Se dice que $\{X_n\}$ converge en probabilidad a Y si para todo $\varepsilon > 0$

$$\lim_{n\to\infty} P(|X_n - Y| \geq \varepsilon) = 0$$

Ley débil de los grandes números

Sea $X_1, \ldots, X_n$ variables aleatorias independientes con $E(X_i) = \mu$ y $Var(X_i) = \sigma^2$. Sea

$$\overline{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$$

Entonces para cada $t > 0$

$$\begin{aligned} P(|\overline{X}_n - \mu| \geq \varepsilon) &\leq \frac{\sigma^2}{n\varepsilon^2} \\ P(|\overline{X}_n - \mu| < \varepsilon) &\geq 1 - \frac{\sigma^2}{n\varepsilon^2} \end{aligned}$$

Por lo cual

$$\lim_{n\to\infty} P(|\overline{X}_n - \mu| \geq \varepsilon) = 0 \qquad \lim_{n\to\infty} P(|\overline{X}_n - \mu| < \varepsilon) = 1$$

y $\{\overline{X}_n\}$ converge en probabilidad a la variable aleatoria concentrada en μ.

Observación

1. Es importante observar que la noción de convergencia usual en cálculo $a_n \to a$ implica que a_n se transforma cuando n varía y ser acerca a a tanto como uno quiera, siempre que n sea suficientemente grande.
2. En cambio la convergencia enunciada en la ley de los grandes números implica que si $X_n \to X$ en probabilidad, entonces la probabilidad del suceso $(|X_n - X| < \epsilon)$ puede hacerse arbitrariamente próxima a 1 tomando n suficientemente grande.

Ejemplo: Frecuencia relativa

- Sea A un suceso con $P(A) = p$, que es uno de los resultados posibles de un experimento.
- Supongamos que repetimos el experimento n veces en forma independiente y consideramos las variables aleatorias indicadoras X_i que valen 1 si el i-esimo experimento resulta en A y cero si no.
- Estas variables resultan Bernoulli independientes y el promedio

$$\hat{p}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$$

es la frecuencia relativa del evento A en las n repeticiones.

Ejemplo: Frecuencia relativa

-
- Como $E(X_i) = p$, y $Var(X_i) = p(1-p)$, por la ley de los grandes números

$$\hat{p}_n \underset{n \to \infty}{\longrightarrow} p$$

esto es, la frecuencia relativa del evento A converge débilmente (en probabilidad) a la probabilidad del evento A.

Ejemplo

Tiramos un dado 20 veces ¿Cuál es la probabilidad de que la frecuencia relativa $\hat{p}_n$ de veces que aparece el cuatro difiere de la probabilidad real en menos de 0.1?

Resolución

- La ley de los grandes números afirma que

$$P(|\hat{p}_n - p| < \varepsilon) \geq 1 - \frac{p(1-p)}{n\varepsilon^2}$$

- En este caso sabemos que la verdadera probabilidad p es $1/6$, por lo cual $p(1-p) = 1/6(1-1/6) = 0.14$. Como $\varepsilon = 0.01$ y $n = 20$ resulta

$$P(|\hat{p}_n - (1/6)| < 0.01) \geq 1 - \frac{0.14}{20(0.01^2} = 1 - 0.7 = 0.3$$

Ejemplo

¿Cuantas veces tenemos que tirar el dado para que estemos un 90% veces seguros de que la frecuencia relativa $\hat{p}_n$ de veces que aparece el cuatro difiere de la probabilidad real en menos de 0.1?

Resolución

- La ley de los grandes números afirma que

$$P(|\hat{p}_n - p| < \varepsilon) \geq 1 - \frac{p(1-p)}{n\varepsilon^2}$$

- en este caso,

$$P(|\hat{p}_n - p| < 0.1) \geq 1 - \frac{0.14}{n(0.1^2)} = 0.9$$

por lo cual $\frac{0.14}{n(0.1^2)} = 0.1$ y $n = \frac{0.14}{(0.1)^3} = 140$.

Aproximaciones Normales: Binomial

- La ley de los grandes números tiene como corolario fundamental dar una estimación de la probabilidad del cometer **por lo menos un error** ε al aproximar p, la probabilidad de un evento, usando $\hat{p}_n$, la frecuencia relativa de este.
- Dicha probabilidad se achica cuando el número de repeticiones independientes del experimento aumenta.

$$P(|\hat{p}_n - p| \geq \varepsilon) \leq \frac{p(1-p)}{n\varepsilon^2} \leq \frac{1}{4n\varepsilon^2} = \tau$$

- Si deseo calcular cualquier otra probabilidad que involucre la variable $\hat{p}_n$ debo estudiar qué distribución tiene este promedio.

Teorema de De Moivre-Laplace

Si X tiene distribución binomial de parámetros n y p entonces

$$Y = \frac{X - np}{\sqrt{np(1-p)}}$$

tiene distribución normal estándar, si n es suficientemente grande.
Para esta aproximación se recomienda $np \geq 10$ y $n(1-p) \geq 10$

Ejemplo

Sea p la proporción real de votantes a favor del candidato Marquez en la población, y sea $\hat{p}_n$ la proporción muestral de votantes a favor de Marquez, calculado so9bre una muestra aleatoria de tamaño n.

(a) Acotar $P\left(|\hat{p}_{900} - p| \geq 0.025\right)$ y comparar con la cota encontrada usando la ley de los grandes números.

(b) Calcular un valor de n que garantice que $P\left(|\hat{p}_n - p| \geq 0.025\right) \leq 0.01$. y comparar con el n encontrado usando la ley de los grandes números.

Resolución

- Como la variable X que mide el número de votantes de la muestra a favor de Marquez tiene distribución binomial de parámetros n y p, usando la aproximación normal a la binomial resulta

$$
\begin{aligned}
P\left(|\hat{p}_{900} - p| \geq 0.025\right) &= P\left(|X - np| \geq n0.025\right) \\
&= P\left(\left|\frac{X - np}{\sqrt{np(1-p)}}\right| \geq \frac{n0.025}{\sqrt{np(1-p)}}\right) \\
&= P\left(|Z| \geq \frac{n0.025}{\sqrt{np(1-p)}}\right) \\
&= 2\Phi(-\frac{900 * 0.025}{\sqrt{900\frac{1}{4}}}) \\
&= 2\Phi(-1.5) = 2 * 0.0668 = 0.1336
\end{aligned}
$$

Resolución

- Debemos observar que la ley de los grandes números solo nos permite decir que

$$P\left(|\hat{p}_{900} - p| \geq 0.025\right) \leq 0.44$$

mientras que la verdadera probabilidad es tres veces menor.

Resolución

- Si queremos encontrar un valor de n que garantice que

$$P\left(|\hat{p}_n - p| \geq 0.025\right) \leq 0.01$$

entonces

$$\begin{aligned} P\left(|\hat{p}_n - p| \geq 0.025\right) &= P\left(\left|\frac{X - np}{\sqrt{np(1-p)}}\right| \geq \frac{n0.025}{\sqrt{np(1-p)}}\right) \\ &= P\left(|Z| \geq \frac{n0.025}{\sqrt{np(1-p)}}\right) \\ &= 2\Phi(-\frac{n * 0.025}{\sqrt{n\frac{1}{4}}}) = 0.01 \end{aligned}$$

por lo cual $-2.575 = z_{-0.005} = -\frac{n*0.025}{\sqrt{n\frac{1}{4}}} = -n^{1/2}0.05$ y $n = 2653$.

- Este es un tamaño de muestra grande pero 20 veces más chico que el calculado con la ley de los grandes números.

Muestra aleatoria

Una muestra aleatoria de tamaño n es un vector de variables $X_1, \ldots, X_n$ independientes e idénticamente distribuidas.

Teorema central del límite

Sea $X_1, \ldots X_n$ una muestra aleatoria de una distribución con media μ y varianza σ^2. Entonces si $S_n = X_1 + \cdots + X_n$, la función de distribución de la variable

$$\frac{S_n - E(S_n)}{\sqrt{Var(S_n)}} = \frac{S_n - n\mu}{\sqrt{n}\sigma}$$

converge a la la función de distribución de una variable normal standard. Esto es

$$P\left(\frac{S_n - n\mu}{\sqrt{n}\sigma} \leq t\right) \longrightarrow \Phi(t) \qquad \forall t$$

Por lo cual, si n es suficientemente grande, puede considerarse a S_n con distribución normal con media $n\mu$ y varianza $n\sigma^2$.

Corolario Teorema central del límite

Sea $X_1, \ldots X_n$ una muestra aleatoria de una distribución con media μ y varianza σ^2. Entonces si $\overline{X}_n = \frac{X_1+\cdots+X_n}{n}$, la función de distribución de la variable

$$\frac{\overline{X}_n - E(\overline{X}_n)}{\sqrt{Var(\overline{X}_n)}} = \sqrt{n}\frac{\overline{X}_n - \mu}{\sigma}$$

converge a la la función de distribución de una variable normal estándar. Esto es

$$P\left(\sqrt{n}\frac{\overline{X}_n - \mu}{\sigma} \leq t\right) \longrightarrow \Phi(t) \qquad \forall t$$

Por lo cual, si n es suficientemente grande, puede considerarse a $\overline{X}_n$ con distribución normal con media μ y varianza $\frac{\sigma^2}{n}$.

Observación

- Toda variable que puede escribirse como una suma de variables independientes idénticamente distribuidas puede ser aproximada por una distribución Normal.
- El teorema de De Moivre -Laplace muestra esto para la variable Binomial, pero la Poisson y la Binomial Negativa tambien pueden ser aproximadas por una Normal
- Y la distribución Gamma de tasa λ, cuando su parámetro de forma es un número natural, pues puede considerarse una suma de exponenciales independientes con tasa λ.

Lema

a) Sea X con distribución Poisson de parámetro λ. Si pensemos a X como suma de variables Poisson independientes $X_1, \ldots, X_n$ todas con el mismo parámetro λ/n, cuando n es suficientemente grande

$$P(X \leq x) \sim \Phi(x - E(X)/\sqrt{Var(X)}) = \Phi(x - \lambda/\sqrt{\lambda})$$

b) Sea X una variable con distribución binomial negativa de parámetros n y p, entonces X puede ser vista como una suma de n variables geométricas independientes de parámetro p. Por lo tanto, cuando n es suficientemente grande

$$P(X \leq x) \sim \Phi(x - E(X)/\sqrt{Var(X)}) = \Phi(x - (n/p)/\sqrt{n(1-p)/p})$$

Ejemplo

Supongamos que un programa suma números aproximando cada sumando al entero más próximo. Si todos los errores cometidos son independientes entre sí y están distribuidos uniformemente entre -0.5 y 0.5 y se suman 1500 números,

1. ¿cuál es la probabilidad de que la magnitud del error total exceda 15?
2. ¿A lo sumo cuántos números pueden sumarse juntos para que la magnitud del error total se mantenga menor que 10 con probabilidad 0.9?

Resolución

- Cada error cometido es una variable aleatoria ε_k con distribución $\mathcal{U}[-0.5, 0.5]$, media $E(\varepsilon_k) = [0.5 + (-0.5)]/2 = 0$ y varianza $Var(\varepsilon_k) = (0.5 - (0.5))^2/12 = 1/12$.
- Definamos $S_n = \sum_{k=1}^{n} \varepsilon_k$, con $n = 1500$. Entonces por el teorema central del límite, $[S_{1500} - nE(\varepsilon)]/\sqrt{nVar(\varepsilon)} \sim N(0,1)$ y

$$\begin{aligned}
P(|S_{1500}| > 15) &= 1 - P(|S_{1500}| \leq 15) = 1 - P(-15 \geq S_{1500} \leq 15) \\
&= 1 - P\left(\frac{-15 - nE(\varepsilon)}{\sqrt{nVar(\varepsilon)}} \leq \frac{S_{1500} - nE(\varepsilon)}{\sqrt{nVar(\varepsilon)}} \leq \frac{15 - nE(\varepsilon)}{\sqrt{nVar(\varepsilon)}}\right) \\
&= 1 - P(-1.34 \leq Z \leq 1.34) \\
&= 2\Phi(-1.34) = 0.1802
\end{aligned}$$

Resolución

- Ahora, deseamos encontrar el n más grande para el cual

$$0.9 = P(|S_n| < 10)$$

Usando el teorema central del límite $[S_{1500} - nE(\varepsilon)]/\sqrt{nVar(\varepsilon)} \sim N(0,1)$ y

$$\begin{aligned}
P(|S_n| < 10) &= P(-10 < S_n < 10) \\
&= P\left(\frac{-10 - nE(\varepsilon)}{\sqrt{nVar(\varepsilon)}} \leq \frac{S_n - nE(\varepsilon)}{\sqrt{nVar(\varepsilon)}} \leq \frac{10 - nE(\varepsilon)}{\sqrt{nVar(\varepsilon)}}\right) \\
&= P\left(\frac{-10}{\sqrt{\frac{n}{12}}} \leq Z \leq \frac{10}{\sqrt{\frac{n}{12}}}\right) \\
&= 1 - 2P\left(Z \leq \frac{-10}{\sqrt{\frac{n}{12}}}\right)
\end{aligned}$$

Resolución

- por lo cual

$$0.9 = 1 - 2P\left(Z \leq \frac{-10}{\sqrt{\frac{n}{12}}}\right) \Longrightarrow P\left(Z \leq \frac{-10}{\sqrt{\frac{n}{12}}}\right) = 0.05$$

y $\frac{-10}{\sqrt{\frac{n}{12}}} = -1.65$. Entonces, despejando resulta

$$n = \frac{10^2 12}{1.65^2} = 440.7$$

Ejemplo

Suponga que se tienen 100 lámparas de un cierto tipo, cuya duración puede modelarse como una variable exponencial de parámetro $\lambda = 0.002$. Si la duración de cada lámpara es independiente de la duración de las otras, encuentre la probabilidad de que el promedio muestral $\overline{T} = (1/100)(T_1 + \cdots + T_{100})$ se encuentre entre 400 y 550 horas.

Resolución

Como n es 100, podemos suponerlo suficientemente grande y aproximar la distribución del promedio por una normal. Entonces la esperanza y varianza de $S_n = T_1 + \cdots + T_n$ son

$$E(S_n) = E(T_1 + \cdots + T_{100}) = 100.E(T_1) = \frac{100}{0.002} = 50000$$

$$Var(S_n) = Var(T_1 + \cdots + T_{100}) = 100.Var(T_1) = \frac{100}{0.002^2}$$

$$\begin{aligned}
P(400 \leq \frac{T_1 + \cdots + T_{100}}{100} \leq 550) &= P(40000 \leq T_1 + \cdots + T_{100} \leq 55000) \\
&\sim \Phi(\frac{55000 - E(S_n)}{\sqrt{Var(S_n)}}) - \Phi(\frac{40000 - E(S_n)}{\sqrt{Var(S_n)}}) \\
&= \Phi(\frac{55000 - 50000}{5000}) - \Phi(\frac{40000 - 50000}{5000}) \\
&= \Phi(1) - \Phi(-2) \\
&= 0.8413 - 0.0228 = 0.8185
\end{aligned}$$