Desigualdad de Chebycheff y Ley de los grandes números

June 11, 2020

Tchebycheff

Sea W una v.a. , y sea $\varepsilon > 0$

$$\mathbb{P}(|W - \mathbb{E}[W]| \geq \varepsilon) \leq \frac{V[W]}{\varepsilon^2}$$

Tchebycheff aplicado al promedio

Sean $(X_i)_{i\geq 1}$ iid, $\mathbb{E}\left(X_i\right) = \mu$, $V\left(X_i\right) = \sigma^2$, y sea $\varepsilon > 0$

$$\mathbb{P}\left(\left|\bar{X}_n - \mu\right| \geq \varepsilon\right) \leq \frac{\sigma^2/n}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2}$$

Tchebycheff aplicado al promedio de v.a. Bernoulli

Si $X_i \sim \mathcal{B}(1, p)$

$$\mathbb{P}\left(\left|\bar{X}_n - p\right| \geq \varepsilon\right) \leq \frac{p(1-p)}{n\varepsilon^2} \leq \frac{1}{4n\epsilon^2}$$

Convergencia en probabilidad

Sean $(Y_n)_{n\geq 1}$, Y variables aleatorias. Diremos que $(Y_n)_{n\geq 1}$ converge a Y en probabilidad si para todo $\epsilon > 0$

$$\lim_{n\to\infty} \mathbb{P}(|Y_n - Y| > \varepsilon) = 0$$

Notación: $Y_n \to Y$ en probabilidad

Ley de los Grandes Números

$\text{Sean}\,(X_i)_{i\geq 1}$ i.i.d., $\text{con}\,\mathbb{E}\,(X_i) = \mu \text{y}\, V\,(X_i) = \sigma^2$, para todo i
Entonces, el promedio de $X_1 \ldots X_n$ converge a μ en probabilidad:
es decir para todo $\varepsilon > 0$ vale que

$$\lim_{n\to\infty} \mathbb{P}\left(\left|\bar{X}_n - \mu\right| > \varepsilon\right) = 0$$

$\bar{X}_n \to \mu$ en probabilidad

Ejercicio

Consideremos las duraciones de lámparas en días. Sea X_i la duración de la i-ésima lámpara, para $i = 1, \dots, n$. Asumiremos que $X_1, \dots, X_n$ son v.a.i.i.d. Haga click aquí para obtener datos simulados, introduciendo su número de libreta.

1. Indique qué cuenta debe hacer con la muestra $(X_1, \dots, X_n)$ para estimar $\mu = E(X_1)$ y $\sigma^2 = V(X_1)$. Es decir, proponga un estimador $\widehat{\mu}_n$ para μ y un estimador $\hat{\sigma}_n^2$ para σ^2. Justifique.
2. Considere $n = 5$ datos de duraciones de lámparas y calcule las estimaciónes de μ y σ^2 correspondientes a estos datos. Repita considerando $n = 30$ y $n = 100$. Introducir los resultados en la página y chequear que sean correctos.

Ejercicio

Consideremos las duraciones de lámparas en días. Sea X_i la duración de la i-ésima lámpara, para $i = 1, \ldots, n$. Asumiremos que $X_1, \ldots, X_n$ son v.a.i.i.d. Haga click aquí para obtener datos simulados, introduciendo su número de libreta.

3. Sea $p = P(X \leq 12)$ con $X \sim X_1$. Indique qué cuenta debe hacer con la muestra $(X_1, \ldots, X_n)$ para estimar p. Es decir, proponga un estimador $\widehat{p}_n$ para p. Justifique.
4. Considere $n = 5$ datos de duraciones de lámparas y calcule la estimación de p correspondiente a estos datos. Repita considerando $n = 30$ y $n = 100$.
5. Acotar la probabilidad de que el estimador propuesto en c) diste de la verdadera probabilidad p, en menos de 0.01, para $n = 5, 30$, 100 y 10000 .

Ejercicio

Consideremos las duraciones de lámparas en días. Sea X_i la duración de la i-ésima lámpara, para $i = 1, \ldots, n$. Asumiremos que $X_1, \ldots, X_n$ son v.a.i.i.d. Haga click aquí para obtener datos simulados, introduciendo su número de libreta.

6. Supongamos ahora que el parámetro de la exponencial es conocido: $\lambda = 1/4$. Se tienen 10 lámparas producidas por esta compañía conectadas en serie, de manera de que cuando una se rompe comienza a funcionar la siguiente. Acotar la probabilidad de que la duración de las 10 lámparas esté entre 20 y 60 días. Comparar con el valor exacto

Resolución

1 -

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

Por la LGN

$$\hat{\mu}_n \to \mu \text{ en probabilidad}$$

$$V(X_1) = E(X_1^2) - E(X_1)^2$$

$$\hat{\sigma}_n^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)^2$$

$Y_1 = 1$ si $X_i \leq 12$
$Y_1 = 0$ si no

$$\hat{p}_n = \overline{Y}_n$$

3 -

$$P(|\overline{Y}_n - p| \geq 0.01) < V(\overline{Y}_n)/\epsilon^2$$

$$P(|\overline{Y}_n - p| < 0.01) > 1 - V(\overline{Y}_n)/\epsilon^2 = 1 - p(1-p)/(n/0.01^2)$$

$$P(|\overline{Y}_n - p| < 0.01) > 1 - \frac{1}{4n0.01^2}$$

6- S es la duracion de las 10 lámparas sumadas.

$$S = X_1 + X_2 + \dots X_{10}$$

$$P(20 < S < 60) = ???$$

$E(S) = \sum_{i=1}^{10} E(X_i) = 10 * 4 = 40$
$V(S) = \sum_{i=1}^{10} V(X_i) = 10 * 16 = 160$

$$P(20 - 40 < S - 40 < 60 - 40) = P(-20 < S - 40 < 20)$$

$$P(|S - 40| < 20) > 1 - \frac{V(S)}{400} = 1 - \frac{160}{400} = 0.6$$

Sabemos que la suma de exponenciales es gamma

$$S \sim \Gamma(10, 1/4)$$

$$P(20 < S < 60) = pgamma(60, shape = 10, rate = 1/4) -$$

$$pgamma(20, shape = 10, rate = 1/4) = 0.898$$