Probabilidades - Primer cuatrimestre 2021

Práctica 6

Desigualdades

1. Un minorista recibe mensualmente galletitas sin sal de 3 fábricas distintas siendo las cantidades recibidas (en kg.) variables aleatorias independientes X, Y y Z con distribuciones: $X \sim \mathcal{N}(100, 20)$, $Y = 97 + W$ con $W \sim \mathcal{E}(1/3)$ y $Z \sim \mathcal{U}(80, 90)$. Acotar la probabilidad de que el peso total recibido en un mes se encuentre entre 275 y 295 kg.

2. Una máquina produce rieles cuya longitud (en metros) es una v.a. con distribución $\mathcal{U}(0{,}8, 1{,}2)$. Se eligen al azar n rieles en forma independientes. Sea $\overline{X}$ el promedio de sus longitudes. Hallar n tal que:
$$P(0{,}99 < \overline{X} < 1{,}01) > 0{,}90.$$

3. Sea p la probabilidad de que una persona elegida al azar apoye una nueva propuesta legislativa (p desconocida). Si se estima p a partir de la frecuencia relativa f_r que resulta al encuestar a n personas:
$$f_r = \frac{\text{número de personas que apoyan la propuesta}}{n},$$
cuánto más cerca esté f_r de p, mejor será la estimación.

   Calcular el mínimo tamaño de muestra requerido para que $P(|f_r - p| \leq 0{,}1) \geq 0{,}95$.

Convergencia

1. Sean $(X_n)_{n>1}$ variables aleatorias independientes e idénticamente distribuídas, con $X_i \sim \mathcal{E}(1)$ para todo $i \in \mathbb{N}$, y sea
$$Y_n = \frac{X_n}{ln(n)}$$
Probar que $Y_n \xrightarrow{p} 0$.

2. Sean $(X_n)_{n\geq 1}$ variables aleatorias $i.i.d$, $X_i \sim U(0, 1)$. Hallar el límite casi seguro de
$$Y_n = (\prod_{i=1}^{n} X_i)^{\frac{1}{n}}$$
.

3. Sean $(X_n)_{n\geq 1}$ variables aleatorias i.i.d. con $E(X_i) = var(X_i) = 1$. Probar que:

$$\frac{\sum_{i=1}^n X_i}{(n\sum_{i=1}^n X_i^2)^{\frac{1}{2}}} \xrightarrow{p} \frac{1}{\sqrt{2}}$$

Convergencia en distribución y Teorema central del límite

1. Sean $X_1, ...X_n$ v.a. i.i.d. con densidad

$$f(x) = \frac{2}{x^3}\, I\{x > 1\}$$

   Calcular aproximadamente

$$P\left(\prod_{i=1}^{100} X_i > e^{55}\right)$$

2. En cierto juego de azar la probabilidad de ganar es 0.3. Para participar en el mismo se paga \$1 y, en caso de ganar, se reciben \$5.

   a) ¿Cuál es la probabilidad aproximada de que en 100 juegos un jugador gane más de \$80? (Suponer que los juegos son independientes entre si).

   b) ¿Cuántas veces tendrá que jugar para ganar más de \$80 con probabilidad mayor o igual que 0.90?

3. Una empresa láctea produce un cierto tipo de queso en unidades cuyo peso (en kg.) es una variable aleatoria con media 2 y varianza 0.04.

   a) Calcular en forma aproximada la probabilidad de que 60 quesos pesen más de 122 kg.

   b) ¿Cuántas unidades serán necesarias para satisfacer un pedido de 5000 kg con probabilidad mayor o igual que 0.95?

4. Sean $(X_n)_{n\geq 1}$ variables aleatorias *i.i.d.*, $X_n \sim \mathcal{U}(0, \theta)$, $\theta > 0$. Probar que

$$\sqrt{n}[ln(2\bar{X}_n) - ln(\theta)] \xrightarrow{D} \mathcal{N}(0, 1/3)$$

   donde $\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$.

5. Sean $(X_n)_{n\geq 1}$ variables aleatorias *i.i.d.*, $E(X_i) = 0$, $E(X_i^2) = 2$. Hallar el límite en distribución de

a) $Y_n = (\sqrt{n}\sum_{i=1}^n X_i)/\sum_{i=1}^n X_i^2$.

b) $U_n = (\sum_{i=1}^n X_i)/\sqrt{\sum_{i=1}^n X_i^2}$.

Simulación

A través de los correspondientes histogramas y otras formas de estimación de la densidad, analizaremos el comportamiento de la distribución de la suma de n variables aleatorias independientes, $S_n = \sum_{i=1}^n X_i$, a medida que sumamos un numero creciente de variables aleatorias (n aumentando). Para ello, fijado n, generaremos datos correspondientes a una muestra de n réplicas de una variable aleatoria X, $(X_1, \ldots, X_n \sim X$, i.i.d), para una distribución dada y luego calcularemos la suma de cada conjunto de datos. Repetimos este procedimiento Nrep = 1000 veces. A partir de las Nrep = 1000 replicaciones realizaremos un histograma con las sumas generadas, para obtener una aproximación de la densidad o la función de probabilidad de S_n.

Para las diferentes distribuciones de la variable aleatoria X, desarrollaremos el siguiente ejercicio:

i. Simular 1000 realizaciones de la variable $X_1 \sim X$.

ii. Realizar un histograma para la variable X_1.

iii. Simular 1000 realizaciones de una variable $X_2 \sim X$ independiente de X_1.

iv. Defina $S_2 = X_1 + X_2$. Realizar un histograma para S_2 ¿qué observa?

v. Repita el mismo procedimiento, agregando de a una variable aleatoria, sumando hasta 10 variables independientes con la misma distribución que X, definiendo así las variables $S_1, \ldots, S_{10}$.

vi. Utilizando las funciones de R `mean()` y `sd()`, crear una función que dado un vector, le reste a cada elemento su promedio y luego lo divida por la estimación de su desviación estándar. Llamar a esa función *estandarizar*.

vii. Aplicar la función *estandarizar* a los vectores creados $S_1, \ldots, S_{10}$, y luego a los nuevos vectores estandarizados graficarlos en un mismo plot utilizando la función *density*, con diferentes colores. Concluir acerca de lo que observa.

viii. Observar graficamente que ocurre si sigue sumando variables aleatorias. ¿Para todas las distribuciones propuestas obtiene resultados similares?

Desarrollar los 8 pasos para las siguientes distribuciones:

a. $X \sim \mathcal{U}(0,1)$.

b. $X \sim \mathcal{B}(10, 0{,}5)$.

c. $X \sim \mathcal{B}(10, 0{,}1)$.

d. $X \sim \mathcal{N}(5, 9)$.

e. $X \sim \mathcal{E}(0{,}5)$.

f. $X \sim \mathcal{C}(0{,}5)$.

g. $X \sim \mathcal{LN}(0{,}5, 1)$.

Ejercicios complementarios

1. Sean $X_1, ...X_n$ variables aleatorias *i.i.d.*, $X_n \sim \mathcal{U}(0,1)$. Sean $X_{(1)} = \min(X_1, ..., X_n)$, $X_{(n)} = \max(X_1, ..., X_n)$, $U_n = nX_{(1)}$, $V_n = n(1 - X_{(n)})$. Probar que:

   a) $X_{(1)} \xrightarrow{p} 0, X_{(n)} \xrightarrow{p} 1$.

   b) $U_n \xrightarrow{D} W, V_n \xrightarrow{D} W$, donde W tiene distribución exponencial de parámetro 1.

2. Sean $U_1, ...U_n$, variables aleatorias independientes con distribución $\mathcal{U}(0,1)$ y sea h una función continua.

   a) Si se define $I_1 = \frac{\sum_{i=1}^n h(U_i)}{n}$, verificar que

   $$E(I_1) = I = \int_0^1 h(x)dx$$

   b) Proponer un método basado en generación de números al azar, para calcular en forma aproximada el valor de la integral I.

   c) Proponer un método para el cálculo aproximado de

   $$I = \int_a^b h(x)dx$$

   siendo a y b número reales tales que $a < b$.