Práctica 6
Convergencia y desigualdades

1. Ejercicio 1

Sea $X_1, \dots, X_n$ una muestra aleatoria simple (m.a.s.) de una población normal con media $\mu = 50$ y varianza $\sigma^2 = 16$. Para $n = 1, 10, 100, 1000$ se pide:

(a) Como se distribuye la variable aleatoria media muestral?

(b) ¿Cómo cambia tu respuesta si la distribución de X_i para $i = 1, \dots, n$ fuera desconocida?

(c) Calcula $P(45 \leq \bar{X} \leq 55)$ para los valores de n anteriores.

(d) Discutí que relación hay entre los resultados anteriores y la ley de los grandes números.

2. Ejercicio 2

Sabemos que la v.a. X, que representa una característica de interés de una población, tiene una distribución normal de parámetros $\mu = 0$ y $\sigma^2 = 1$. Si tomamos una muestra aleatoria de tamaño $n = 10$ de la población, responde a las siguientes preguntas: (Superponer las distribuciones de X y $\bar{X}$ para tener intuición sobre tu respuesta)

(a) $P(X \leq \mu)$ es mayor, menor o igual a $P(\bar{X} \leq \mu)$?

(b) $P(|X - \mu| \geq 1)$ es mayor, menor o igual a $P(|\bar{X} - \mu| \geq 1)$?

(c) $P(|X - \mu| \leq 1)$ es mayor, menor o igual a $P(|\bar{X} - \mu| \leq 1)$?

(d) ¿Qué ocurre en todos los puntos anteriores a medida que n se hace cada vez más grande? ¿En términos prácticos, que relevancia tiene?

3. Ejercicio 3

Una encuesta de la AFIP encontró que el 82 % de los contribuyentes manifestaron la importancia de que el organismo se asegurara de que los contribuyentes de mayores ingresos no evadieran impuestos.

(a) Si tomás 8 contribuyentes de manera aleatoria, ¿cuál es la probabilidad de que al menos 6 de ellos consideren importante que el organismo se asegure que los contribuyentes de mayores ingresos no evadan impuestos?

(b) Si tomás 80 contribuyentes de manera aleatoria, ¿cuál es la probabilidad de que al menos 60 de ellos consideren importante que el organismo se asegure que los contribuyentes de mayores ingresos no evadan impuestos?

Resolvé ambos incisos utilizando la distribución Binomial y la distribución Normal. Compará los resultados.

4. Ejercicio 4

Un videojuego fue descargado por 192 personas. El desarrollador hizo un análisis, determinando que la cantidad de horas diarias que cada usuario utiliza la app se puede suponer $U(1;7)$ y que representan cantidades aleatorias simultáneamente independientes.

(a) Para un día específico, cuál es la probilidad aproximada de que el tiempo acumulado que juegan los 192 usuarios sea de al menos 800 hs?

(b) El desarrollador quiere establecer una tarifa proporcional al tiempo diario de uso de la app, cual debe ser esa tarifa por hora si pretende ganar al menos 2000 USD diarios con una confianza del 95 %?

5. Ejercicio 5

Los pesos de ciertos paquetes de café de una compañía cafetera son variables aleatorias independientes se distribuyen normalmente con media 500 grs. y desvío σ. ¿Cuál debe ser el valor de σ para tener una confianza del 99 % de que el peso promedio de 25 paquetes no se desviará en más de 10 grs. de la media?

6. Ejercicio 6

Un vuelo de Aerolíneas Argentinas dispone de 100 asientos. La experiencia indica que cada reserva tiene una probabilidad de 0,1 de ser cancelada a último momento. Suponé que no hay lista de espera y que los pasajeros hacen sus reservas individualmente, es decir, de forma independiente. Si la compañía desea que la probabilidad de que queden clientes indignados por haber hecho su reserva y no poder viajar sea menor que 0,01 ¿cuál debería ser el número máximo de reservas que deberían aceptar?

7. Ejercicio 7

Un minorista recibe mensualmente galletitas de salvado de 3 fábricas distintas, siendo las cantidades recibidas (en kilos) variables aleatorias independientes X, Y y Z, cuya distribución está dada por $X \sim \mathcal{N}(100, 20)$, $Y = 97 + W$ con $W \sim Exp(1/3)$ y $Z \sim U[80, 90]$. Acotar la probabilidad que el total recibido en un mes se encuentre entre 275 y 295 kilos.

8. Ejercicio 8

(a) Sea X una variable aleatoria con distribución desconocida tal que $\mathbb{E}[X] = 5$ y $Var(X) = 0{,}1$. Acote la probabilidad de que X esté entre 4.5 y 5.5.

(b) Sea $X_1, \ldots, X_{10}$ una muestra aleatoria, es decir, varibles aleatorias independientes e idénticamente distribuidas con $\mathbb{E}[X_i] = 5$ y $Var(X_i) = 0{,}1$ para todo i. Acote la probabilidad de que la media muestral se sitúe entre 4.5 y 5.5.

(c) Si se contase con una muestra aleatoria de tamaño n: $X_1, \ldots, X_n$ ¿Qué pasaría con la cota hallada en (b) cuando n tendiese a infinito?

(d) Sean $X_1, \ldots, X_n$ variables aleatorias independientes que verifican $\mathbb{E}[X_i] = 4$ y $Var(X_i) = 9 + \dfrac{2^i + 1}{2^i}$. Hallar a qué converge en probabilidad $Y_n = \overline{X}_n + e^{5-\frac{1}{n}}$.

9. Ejercicio 9

Suponga usted que $X_1, \ldots, X_n$ es una muestra aleatoria de tamaño n de variables independientes e idénticamente distribuidas como $U(0, \theta)$. Sea $X_{\text{máx}} = \text{máx}\,\{X_1, \ldots, X_n\}$.

(a) Encuentre la función de densidad de $X_{\text{máx}}$.

(b) Calcule $\mathbb{E}[X_{\text{máx}}]$ y $Var(X_{\text{máx}})$.

(c) Calcule el límite en probabilidad de $X_{\text{máx}}$.

(d) Demuestre que $n\,(\theta - X_{\text{máx}})$ converge en distribución a una exponencial.

10. Ejercicio 10

Sean $X_1, \ldots, X_n$ variables aleatorias con distribución de Poisson, cada una de ellas con media 1.

(a) Usando la desigualdad de Markov, obtener una cota para $P\left(\sum_{i=1}^{20} X_i > 15\right)$

(b) Usando el Teorema Central del Límite, apoximar $P\left(\sum_{i=1}^{60} X_i > 15\right)$.

11. Ejercicio 11

Las variables $Z_1, \ldots, Z_n$ son independientes con distribución $U(-20, 10)$. Hallar en forma aproximada

$$P\left(\sum_{i=1}^{900} Z_i > -4470\right).$$

12. Ejercicio 12

Sean $X_1, \ldots, X_n$ variables aleatorias independientes e idénticamente distribuidas con $\mathbb{E}[X] = \mu$ y $Var(X) = \sigma^2$. Demuestre que

$$\frac{\sqrt{n}\left(\overline{X}_n - \mu\right)}{S} \xrightarrow{d} \mathcal{N}(0,1)$$

13. Ejercicio 13

Una variable aleatoria X continua tiene distribución exponencial de parámetro $\lambda > 0$, $X \sim Exp(\lambda)$, si su función de densidad es

$$f_X(x) = \lambda e^{-\lambda x} I_{(0,+\infty)}(x).$$

Para $k > 1$, calcule

$$P\left(|X - \mathbb{E}(X)| \geq k\sqrt{Var(X)}\right)$$

y compare con la cota obtenida a través de la desigualdad de Tchebyshev.

14. Ejercicio 14

Sea $X \sim \mathcal{Poi}(\lambda)$, es decir una variable discreta con distribución Poisson de parámetro λ. Demuestre que $P(X \geq 2\lambda) \leq 1/\lambda$.

15. Ejercicio 15

En una línea de producción los productos pasan por 4 procesos sucesivos (preparación, armado, control y embalaje) hasta quedar listos para la venta. Sean X_1, X_2, X_3 y X_4 los tiempos que tarda en cumplirse cada uno de los procesos (en minutos). Se sabe que los cuatro procesos actúan en forma independiente y que $X_1 \sim U(3,5)$, $X_2 \sim \mathcal{N}(3, 1/4)$, $X_3 \sim \mathcal{N}(4, 1/4)$ y $X_4 \sim U(2,4)$. Sea Y la variable aleatoria que mide el tiempo total que tarda un producto en pasar por toda la línea de producción.

a) Halle $E(Y)$ y $Var(Y)$.

b) Acotar la probabilidad de que el tiempo total se aleje del tiempo total esperado en al menos dos minutos.

16. Ejercicio 16

El número de aviones que aterrizan en un aeropuerto por hora es una variable aleatoria que sigue una distribución de Poisson con parámetro $\lambda = 30$. Determine una cota inferior para la probabilidad de que el número de aviones que aterrizan en un período de una hora esté entre 20 y 40.

17. Ejercicio 17

Sea p la probabilidad de que una persona elegida al azar de la población adulta argentina apoye cierta política pública. Para estimar p se encuesta a n personas, se cuenta la cantidad de personas a favor y se calcula la media muestral $\overline{X}_n$.

a) Halle una cota superior para $P(|\overline{X}_n - p| \geq 0{,}1)$ que no dependa de p.

b) ¿A cuánta gente debería encuestarse para que $P(|\overline{X}_n - p| \geq 0{,}1) \leq 0{,}1$?

18. Ejercicio 18

Sean $X_i \overset{iid}{\sim} Exp(\lambda)$. Muestre que $V_n = \text{mín}\{X_i\} \xrightarrow{p} 0$ y que $V_n = \text{mín}\{X_i\} \xrightarrow{m.c.} 0$.

19. Ejercicio 19

En una línea de producción los productos pasan por 4 procesos sucesivos (preparación, armado, control y embalaje) hasta quedar listos para la venta. Sean X_1, X_2, X_3 y X_4 los tiempos que tarda en cumplirse cada uno de los procesos (en minutos). Se sabe que los cuatro procesos actúan en forma independiente y que $X_1 \sim U(3,5)$, $X_2 \sim \mathcal{N}(3,1/4)$, $X_3 \sim \mathcal{N}(4,1/4)$ y $X_4 \sim U(2,4)$. Sea Y la variable aleatoria que mide el tiempo total que tarda un producto en pasar por toda la línea de producción.

a) Halle $E(Y)$ y $Var(Y)$.

b) Acotar la probabilidad de que el tiempo total se aleje del tiempo total esperado en al menos dos minutos.

20. Ejercicio 20

Dadas observaciones independientes $X_1, \ldots, X_n$ tales que $E(X_i) = 0$ y $a_i \leq X_i \leq b_i$, la desigualdad de Hoeffding dice que dado $\varepsilon > 0$ fijo se cumple la siguiente desigualdad

$$P\left(\sum_{i=1}^{n} X_i \geq \varepsilon\right) \leq \exp\left(-2\,\frac{\varepsilon^2}{\sum_{i=1}^{n}(b_i - a_i)^2}\right). \tag{1}$$

a) Si $X_1, \ldots, X_n \sim Be(p)$ pruebe que para todo $\varepsilon > 0$ se cumple que

$$P(|\overline{X}_n - p| > \varepsilon) \leq 2e^{-2n\varepsilon^2} \tag{2}$$

donde $\overline{X}_n = n^{-1}\sum_{i=1}^{n} X_i$.

b) Muestre que para valores grandes de n la cota de Hoeffding es mejor que la de Tchebyshev.

21. Ejercicio 21

Sean $X_1, X_2, \ldots, X_{50}$ variables aleatorias independientes cada una con función de densidad

$$f_X(x) = \begin{cases} 2x & \text{si } 0 \le x \le 1, \\ 0 & \text{en caso contrario.} \end{cases}$$

Sean $S_{50} = X_1 + X_2 + \cdots + X_{50}$ y $\overline{X}_{50} = S_{50}/50$.

a) Usar el Teorema Central del Límite para aproximar $P(S_{50} \le 30)$ y $P(\overline{X}_{50} \ge 0{,}7)$.

b) Hallar el valor de a tal que $P(S_{50} \ge a) \approx 0{,}86$.

22. Ejercicio 22

Supongamos que una medición tiene media (o esperanza) igual a μ y su varianza es $\sigma^2 = 25$. Sea $\overline{X}_n$ el promedio de n de tales mediciones independientes entre sí.

a) Aproximar $P\left(\left|\overline{X}_{64} - \mu\right| < 1\right)$.

b) Usando Tchebychev, determinar cuán grande debe ser n para que

$$P\left(\left|\overline{X}_n - \mu\right| < 1\right) \ge 0{,}95.$$

c) Usando el Teorema Central del Límite, hallar n tal que $P\left(\left|\overline{X}_n - \mu\right| < 1\right) \approx 0{,}95$. Comparar con el resultado obtenido en (b).

23. Ejercicio 23

Sean $X_1, \ldots, X_n$ i.i.d. cuya función de densidad es

$$f(x) = \frac{3}{x^4} I_{[1,+\infty)}(x)$$

Calcule aproximadamente

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

24. Ejercicio 24

Un código fuente de un software tiene $n = 100$ páginas. Sea X_i la cantidad de errores en la i-ésima página. Supongamos que $X_i \sim \mathscr{P}(1)$ y que son independientes. Sea $Y = \sum_{i=1}^{n} X_i$ el número total de errores. Usando el Teorema Central del Límite aproxime $P(Y < 90)$.

25. Ejercicio 25

Sean $X_1, \ldots, X_n$ una muestra aleatoria tal que X_i tiene función de densidad dada por $f_X(x) = 2xI_{(0,1)}(x)$. Hallar el límite en probabilidad de

$$Y_n = (X_1 X_2 \ldots X_n)^{1/n} = \left(\prod_{i=1}^{n} X_i\right)^{1/n}$$

26. Ejercicio 26

Sea $X_1, \ldots, X_n$ una muestra aleatoria tal que $E(X_1) = 0$, $E(X_1^2) = 2$ y $E(X_1^4) < \infty$. Hallar el límite en distribución de

1. $$V_n = \frac{\sum_{i=1}^{n} X_i}{\sum_{i=1}^{n} X_i^2}.$$

2. $$W_n = \frac{\sqrt{n}\sum_{i=1}^{n} X_i}{\sum_{i=1}^{n} X_i^2}.$$

3. $$Y_n = \frac{\sum_{i=1}^{n} X_i}{\sqrt{\sum_{i=1}^{n} X_i^2}}.$$

4. $$Z_n = \exp\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} X_i\right)$$

5. $$R_n = \exp\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} X_i\right) + \frac{1}{n}\sum_{i=1}^{n} X_i$$

6. $$S_n = \frac{1}{n^{3/2}}\left(\sum_{i=1}^{n} X_i\right)\left(\sum_{i=1}^{n} X_i^2\right)$$

27. Ejercicio 27

Decimos que una sucesión de variables aleatorias $X_1, X_2, \ldots$ converge en **media cuadrática** a una variable aleatoria X si

$$\lim_{n\to\infty} E\left(|X_n - X|^2\right) = 0\,.$$

a) Demuestre que la convergencia en media cuadrática implica convergencia en probabilidad.

b) (*) Mediante un contraejemplo, pruebe que la afirmación recíproca no es cierta.