Práctico 5: Esperanza y varianza poblacionales y muestrales

1. En una Clínica Médica se realizan análisis de sangre a 1.500 personas a diario, para detectar cierta afección. La probabilidad de que el test dé positivo es 0,002. Para reducir costos, la clínica decide implementar un nuevo método, con el objetivo de minimizar el número de ensayos de laboratorio. El método consiste en hacer grupos de 30 personas y mezclar las muestras de sangre en cada grupo. Se realizan entonces 50 análisis, uno para cada grupo. Si el análisis de un grupo da negativo, ninguna de las personas de ese grupo está afectada. En caso contrario (si el análisis del grupo da positivo) se analiza separadamente a las 30 personas del grupo (o sea, se hacen 30 análisis, uno para cada uno).
   (a) Consideremos uno cualquiera de esos 50 grupos, por ejemplo el i-ésimo. Sea X_i la variable aleatoria que indica el número de análisis que hay que hacer en total dentro de ese grupo. ¿Qué valores puede tomar la variable X_i y con qué probabilidades? ¿Cuál es la esperanza de X_i?¿Cuál es la esperanza del número de análisis necesarios para las 1.500 personas?
   (b) Si en vez de considerar 50 grupos de 30 personas se consideran r grupos de k personas ($rk = 1500$) ¿cuál es la esperanza (en función de k) del número de análisis necesarios para esas 1.500 personas?
   (c) ¿Conviene hacer grupos de 20 o de 30 personas?

2. (a) Sea $X_1, X_2, ..., X_n$ variables aleatorias de distribución Poisson de parámetro λ. Calcular la esperanza de la variable aleatoria $Y = X_1 + X_2 + ... + X_n$.
   (b) Se observan los siguientes resultados para 15 experimentos independientes igualmente distribuidos con distribución de Poisson:

i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
x_i	1	3	4	2	8	4	5	4	3	6	6	5	4	5	2

   Usando la *Ley de los Grandes Números*, estimar el valor del parámetro λ.

3. La bacteria *Helicobacter Pylori* genera deficiencias en la absorción de nutrientes debido a su interferencia en la secreción de ácidos por parte del estómago. Sea X la variable aleatoria que representa el tiempo de vida de dicha bacteria donde la unidad de medida es en días. Se asume que la distribución de X es exponencial de parámetro $\lambda = 0.1$.
   (a) Calcular la probabilidad de que el tiempo de vida de la bacteria sea mayor a una semana.
   (b) Indicar $\mathrm{E}(X)$ y $\mathrm{Var}(X)$.

4. Sea X una variable aleatoria con densidad uniforme en el intervalo $[a, b]$.
   (a) Escribir su función densidad.
   (b) Calcular la esperanza de X en los siguientes casos:
      i. Si $a = 0$ y $b = 1$.
      ii. Si $a = 2$ y $b = 4$.
      iii. Concluir cuánto vale $\mathrm{E}(X)$ para cualquier par de parámetros $a < b$.

5. Sean X e Y variables aleatorias independientes con distribución $N(1, 4)$ y $N(3, 9)$ respectivamente.
   (a) Calcular $\mathrm{E}(X + Y)$ y $\mathrm{E}(\frac{4X-Y}{3})$.
   (b) Calcular $\mathrm{Var}(X + Y)$ y $\mathrm{Var}(\frac{4X-Y}{3})$.

6. Sea $X_1, ..., X_n$ variables aleatorias independientes con la misma distribución (**i.i.d.**) que una variable aleatoria X tal que $\mathrm{E}(X) = \mu$ y $\mathrm{var}(X) = \sigma^2$.

   Definimos una nueva variable aleatoria denominada *media muestral* como:

$$\overline{X_n} = \frac{1}{n}(X_1 + ... + X_n)$$

   Demostrar que se cumple:

(a) $\mathrm{E}(\overline{X_n}) = \mu$
(b) $\mathrm{Var}(\overline{X_n}) = \frac{\sigma^2}{n}$

7. Un atleta de salto largo se prepara para las olimpíadas. Podemos suponer que la distribución de un salto de este atleta en metros es normal. Es decir que si $X =$ *"longitud de un salto del atleta"*, entonces X (aleatoria) es $X \sim N(\mu, \sigma^2)$.

   Se sabe que el 7,65% de las veces realiza un salto de menos de 7 metros, y que el 12,65% de las veces realiza saltos de más de 8.10 metros.

   (a) Con la información proporcionada, hallar el salto promedio que realiza este atleta y su desviación estándar.
   (b) La competencia consta de realizar 3 saltos y contabilizar la suma (en metros) de los saltos realizados.

      SUGERENCIA: Suma de variables aleatorias normales independientes tiene distribución normal de parámetros a calcular.

      i. ¿Cuál es el puntaje que se espera que saque el atleta?
      ii. ¿Cuál es la desviación estándar del puntaje obtenido?
      iii. Si queda descalificado todo aquel atleta que sume menos de 21 metros, ¿cuál es la probabilidad de que quede descalificado?
   (c) Ahora la competencia consta en comparar el mejor de tres saltos. Calcular la probabilidad de que el mejor entre tres saltos de nuestro atleta sea de más de 8 metros. (Utilizar probabilidad del complemento e independencia entre saltos).

Ejercicios complementarios

1. Dadas dos variables aleatorias geométricas X_1 y X_2 de parámetros $p_1 < p_2$.

   (a) ¿Cuál de las dos distribuciones tiene menor esperanza?
   (b) ¿Cuál de las dos distribuciones tiene menor varianza?
      i. Sabiendo que el científico ya realizó 10 experimentos, la probabilidad de que el próximo experimento sea positivo ¿es mayor, menor o igual a la probabilidad de ser positivo en el primer intento? ¿Y si ya realizó 100 experimentos?
      ii. Indicar cuál es el número esperado de experimentos que deberá realiza el científico.

2. Sea X una variable aleatoria de distribución exponencial de parámetro λ. Demostrar la relación $(\mathrm{E}(X))^2 = \mathrm{Var}(X)$.

3. El modelo de Bolzman para el movimiento de las partículas de un gas, asume que el vector velocidad de cada partícula es $V = (V_x, V_y, V_z)$ donde las componentes son variables aleatorias normales, independientes e identicamente distribuidas. Se considera la velocidad: $v = |V| = \sqrt{V_x^2 + V_y^2 + V_z^2}$. Puede demostrarse a partir de la distribución de V que la densidad de v es

$$p(x) = \frac{4x^2}{\alpha^3\sqrt{\pi}} e^{-x^2/\alpha^2}$$

   si $x > 0$, y $p(x) = 0$ si $x \leq 0$. Hallar la velocidad media de una molécula y su varianza.

4. Una variable X tiene distribución χ_k^2 si su densidad es

$$f(x) = \begin{cases} C(k)x^{k/2-1}e^{-x/2} & \text{si } x > 0 \\ 0 & \text{si } x \leq 0 \end{cases} \tag{1}$$

   donde $C(k)$ es una constante (que depende de k) que hace que el área bajo la gráfica de f sea uno. Probar que si $X_1, \ldots, X_k$ son variables aleatorias iid con distribución normal con media 0 y varianza 1 entonces si denotamos $X = (X_1 \ldots, X_k)$ verifica que

$$\|X\|^2 \sim \chi_k^2$$