

Nombre: _____ Legajo: _____

Pregunta	1	2	3	4	5	Total
Puntos	2	2	2	2	2	10
Calificación						

No se contestan preguntas.
Justifique sus cálculos.
Redondee los resultados finales a 4 decimales.

Leitmotiv: hijos

1. Se toma una muestra de 100 recién nacidos y se los pesa. A partir de esas mediciones, se forma la siguiente tabla de datos agrupados:

Marca [kg]	2.7	2.9	3.1	3.3	3.5	3.7	3.9	4.1	4.3	4.5
Frecuencia	2	3	14	10	15	18	16	14	4	4

- (a) ($\frac{1}{2}$ pto.) Grafique el histograma.
 (b) ($\frac{1}{2}$ pto.) Grafique la poligonal de frecuencias acumuladas.
 (c) ($\frac{1}{2}$ pto.) Calcule la media y el desvío muestrales. Detalle los cálculos.
 (d) ($\frac{1}{2}$ pto.) Calcule la mediana muestral. Detalle los cálculos.

Respuesta:

Las marcas \bar{x}_i y las frecuencias f_i son las que están en la tabla. $n = \sum_i f_i$ es la cantidad de datos.

$$\bar{x} = \frac{1}{n} \sum \bar{x}_i f_i = 3.6460, \quad s = \sqrt{\frac{1}{n-1} \sum (\bar{x}_i - \bar{x})^2 f_i} = 0.4234.$$

Los datos no se acumulan “hasta” las marcas, sino “hasta” los extremos derechos de los intervalos de clase (*bines*). Por lo tanto, la mediana (que divide a la mitad los datos) tiene que estar en el intervalo $(3.6, 3.8]$. Usando interpolación lineal:

$$\text{mediana} = \frac{50 - 44}{62 - 44} \cdot (3.8 - 3.6) + 3.6 = 3.6667.$$

2. Una corriente pseudocientífica (al parecer, extendida en Japón) asevera que existe una relación entre el tipo sanguíneo y la personalidad de un individuo. En particular, el divulgador científico chino Xius Lu ha determinado que los niños recién nacidos pueden ser clasificados en pacíficos y no pacíficos (C y C^c), lo cual está relacionado con el factor Rh de su sangre (M si es Rh^+ y M^c si es Rh^-). Lu ha determinado, mediante un extenso estudio de bebés japoneses, que:

$$P(M|C) = 0.1, \quad P(M|C^c) = 0.9, \quad P(M^c) = 0.5. \quad (1)$$

- (a) (1 pto.) ¿Cuál es la probabilidad de que un niño tomado al azar sea pacífico?
 (b) (1 pto.) Si un bebé es Rh^- , ¿cuál es la probabilidad de que sea pacífico?

Respuesta:

Sabemos que

$$\begin{aligned} P(M) &= 1 - P(M^c) = P(M|C)P(C) + P(M|C^c)[1 - P(C)] \\ 1 - 0.5 &= 0.1P(C) + 0.9[1 - P(C)] \\ 0.5 &= 0.9 - 0.8P(C) \\ \Rightarrow P(C) &= 0.5. \end{aligned}$$

$$\begin{aligned} P(C|M^c) &= \frac{P(M^c|C)P(C)}{P(M^c|C)P(C) + P(M^c|C^c)P(C^c)} \\ &= \frac{[1 - P(M|C)]P(C)}{[1 - P(M|C)]P(C) + [1 - P(M|C^c)]P(C^c)} \\ &= \frac{[1 - 0.1]0.5}{[1 - 0.1]0.5 + [1 - 0.9]0.5} \\ &= 0.9. \end{aligned}$$

3. Considere una comunidad con n nacimientos. Se sabe que la probabilidad de que un bebé nazca con ojos claros es p_n .
- ($1/2$ pto.) ¿Cuál es la probabilidad de que al menos 3 de los bebés tengan ojos claros? Explique claramente qué asume y escriba la respuesta en función de n y p_n .
 - (1 pto.) Sea r_n la probabilidad de que exactamente 2 bebés tengan ojos claros. Determine qué valor de p_n maximiza dicha probabilidad para todo $n \geq 2$. Escriba su respuesta en función de n .
 - ($1/2$ pto.) Asuma que utiliza el valor p_n hallado en el punto anterior. Calcule $\lim_{n \rightarrow \infty} r_n$.

Respuesta:

Sean los eventos

C_i = el bebé del i -ésimo nacimiento tiene ojos claros con $i = 1, \dots, n$.

Asumimos que los eventos C_i son **independientes** y que para todos ellos tenemos **la misma probabilidad** $P(C_i) = p_n$. Sea X_n el número de bebés con ojos claros. Sabemos que $X_n \sim \text{Bino}(n, p_n)$. Luego

$$\begin{aligned} P(X_n > 3) &= 1 - P(X_n \leq 2) = 1 - \sum_{k=0}^2 \binom{n}{k} p_n^k (1 - p_n)^{n-k} \\ &= 1 - \left[(1 - p_n)^n + np_n(1 - p_n)^{n-1} + \frac{n(n-1)}{2} p_n^2 (1 - p_n)^{n-2} \right] \\ &= 1 - (1 - p_n)^{n-2} \left[(1 - p_n)^2 + np_n(1 - p_n) + \frac{n(n-1)}{2} p_n^2 \right]. \end{aligned}$$

Para el punto b:

$$r_n(p_n) = P(X_n = 2) = \frac{n(n-1)}{2} p_n^2 (1 - p_n)^{n-2}.$$

Llamemos p_n^* al valor que maximiza r_n . Para $n = 2$, es claro que $p_n^* = 1$.

Para $n > 2$,

$$\frac{dr_n}{dp_n} = \frac{n(n-1)}{2} p_n (1-p_n)^{n-3} (2-np_n).$$

El valor crítico es $p_n^* = 2/n$. Para ver que es un máximo, hacemos la segunda derivada:

$$\left. \frac{d^2 r_n}{dp_n^2} \right|_{p_n^*} = \frac{n(n-1)}{2} [\text{Algo} \times (2-np_n^*) - (1-p_n^*)^{n-3} np_n^*] = -n(n-1) < 0.$$

Para el punto c: por la aproximación Poisson a la binomial, la respuesta es

$$\lim_{n \rightarrow \infty} r_n(p_n^*) = \frac{2^2}{2!} e^{-2} = 0.2707.$$

4. El humor, positivo o negativo, de un bebé al despertarse puede ser modelado como una variable aleatoria X con densidad de probabilidad

$$f_X(x) = \begin{cases} \alpha |x|^3 & x \in (-1, +1) \\ 0 & x \notin (-1, +1) \end{cases} \quad (2)$$

con $\alpha \in \mathbb{R}^+$.

- (1 pto.) Determine la media y la varianza de X .
- ($1/2$ pto.) Sea Y una variable aleatoria normal con el mismo valor medio y el mismo desvío que X . Determine la probabilidad de que Y pertenezca al rango de valores posibles de X .
- ($1/2$ pto.) Sea W una variable aleatoria uniforme con el mismo valor medio y el mismo desvío que X . Determine la probabilidad de que W pertenezca al rango de valores posibles de X .

Respuesta:

Dado que f_X es par,

$$1 = \int_{\mathbb{R}} f_X(x) dx = 2 \int_0^1 \alpha x^3 dx = \frac{\alpha}{2} \Rightarrow \alpha = 2.$$

Por la simetría, $\mu_X = E[X] = 0$. Luego,

$$\sigma_X^2 = E[X^2] = 4 \int_0^1 x^5 dx = \frac{2}{3}.$$

Sea $Y \sim \mathcal{N}(0, \sqrt{2/3})$. Luego

$$\begin{aligned} P(Y \in (-1, 1)) &= \Phi\left(+\sqrt{\frac{3}{2}}\right) - \Phi\left(-\sqrt{\frac{3}{2}}\right) \\ &= 2\Phi\left(+\sqrt{\frac{3}{2}}\right) - 1 \\ &= 0.7793 \quad \text{por computadora} \\ &\approx 2 \cdot 0.88970 - 1 = 0.7794 \quad \text{interpolación lineal en tabla} \\ &\approx 2 \cdot 0.88975 - 1 = 0.7795 \quad \text{valor medio en tabla} \\ &\approx 2\Phi(1.22) - 1 = 0.7776 \quad \text{valor más cercano en tabla.} \end{aligned}$$

Dado que $E[W] = 0$, entonces $W \sim \text{Uniforme}(-a, a)$ para $a \in \mathbb{R}^+$. Luego:

$$\sigma_W^2 = \frac{(2a)^2}{12} = \frac{a^2}{3} = \frac{2}{3} \Rightarrow a = \sqrt{2}.$$

Por lo tanto,

$$\begin{aligned} P(W \in (-1, 1)) &= F_W(+1) - F_W(-1) = 2F_W(+1) - 1 = 2\frac{1 + \sqrt{2}}{2\sqrt{2}} - 1 = \frac{1}{\sqrt{2}} \\ &= 0.7071. \end{aligned}$$

5. En nuestra especie, el ADN mitocondrial es heredado de nuestra madre biológica (nuestro padre biológico no aporta nada). Supongamos un modelo muy simple de una comunidad endogámica:

- Sólo son madres mujeres de la comunidad.
 - Cada mujer en la comunidad llega a adulta y tiene exactamente 2 bebés.
 - La probabilidad de que un bebé sea una biológicamente mujer es 0.5.
 - La primera generación está formada por dos mujeres, Ana y Bárbara, con diferente ADN mitocondrial (hijas de distintas madres).
- (a) (1 pto.) Determine la probabilidad de que todas las mujeres en la tercera generación hayan heredado el ADN mitocondrial de Ana. ¿Qué debe asumir para hacer sus cálculos?
- (b) (1 pto.) Explique por qué la probabilidad de que todas las mujeres en la tercera generación hayan heredado el ADN mitocondrial de una sola mujer de la primera generación **no** es igual al doble de la respuesta al ítem anterior.

Respuesta:

Definimos:

- $A(B)$ = número de hijas de Ana (Bárbara).
- $A_1(A_2)$ = número de hijas de la primera (segunda) hija de Ana.
- $B_1(B_2)$ = número de hijas de la primera (segunda) hija de Bárbara.
- E_A = todas las mujeres de la tercera generación heredan el ADN mitocondrial de Ana.

- $E_A (E_B)$ = todas las mujeres de la tercera generación heredan el ADN mitocondrial de Ana (Bárbara).

Ojo: Cero (0) mujeres en la tercera generación $\subset E_A \cap E_B$.

Si asumimos independencia en los nacimientos de hijos de una misma madre: $A, A_1, A_2, B, B_1, B_2 \sim \text{Bino}(2, 0.5)$. También asumiremos “independencia” de todas estas variables aleatorias. Las comillas son porque, claramente, es falso. En concreto:

- $A, B \sim \text{Bino}(2, 0.5)$ son independientes.
- $A_1 \sim \text{Bino}(2, 0.5)$ si $A > 0$ y $A_1 = 0$ caso contrario.
- $A_2 \sim \text{Bino}(2, 0.5)$ si $A > 1$ y $A_2 = 0$ caso contrario.
- $B_1 \sim \text{Bino}(2, 0.5)$ si $B > 0$ y $B_1 = 0$ caso contrario.
- $B_2 \sim \text{Bino}(2, 0.5)$ si $B > 1$ y $B_2 = 0$ caso contrario.
- A_1 y A_2 son independientes si $A = 2$.
- B_1 y B_2 son independientes si $B = 2$.

La idea:

$$\begin{aligned} E_A &= \text{No hay hijas descendientes de Bárbara en la tercera generación} \\ &= \{B = 0\} \cup [\{B = 1\} \cap \{B_1 = 0\}] \cup [\{B = 2\} \cap \{B_1 = 0\} \cap \{B_2 = 0\}] \end{aligned}$$

Los tres eventos en la unión son m.e.

$$\begin{aligned} P(E_A) &= P(B = 0) + P(\{B = 1\} \cap \{B_1 = 0\}) + P(\{B = 2\} \cap \{B_1 = 0\} \cap \{B_2 = 0\}) \\ &= P(B = 0) + P(B = 1)P(B_1 = 0|B = 1) + P(B = 2)P(B_1 = 0|B = 2)P(B_2 = 0|B = 2) \\ &= (0.5^2 + 2 \cdot 0.5^2 \times 0.5^2 + 0.5^2 \times 0.5^2 \times 0.5^2) = 0.390625 \\ &= 0.3906. \end{aligned}$$

Consideremos el evento E_U = todas las descendientes de la tercera generación tienen ADN mitocondrial. Luego $E_U = E_A \cup E_B$ y

$$P(E_U) = P(E_A) + P(E_B) - P(E_A \cap E_B) < P(E_A) + P(E_B) = 2P(E_A).$$

Hemos utilizado dos hechos. El primero es que $E_A \cap E_B$ es el evento a la no existencia de descendientes femeninos en la tercera generación: es fácil mostrar que es un evento de probabilidad positiva. El segundo es que $P(E_A) = P(E_B)$.