Pontificia Universidad Católica de Chile
Facultad de Matemáticas
Departamento de Estadística

Primer Semestre 2024

Curso : **Probabilidad y Estadística**
Sigla : **EYP1113**
Profesores : **Ana María Araneda L., Ricardo Aravena C., Ricardo Olea O., Felipe Ossa M. y Inés Varas C.**

PAUTA INTERROGACIÓN 2

Problema 1

La velocidad de sedimentación globular (VSG) mide la distancia que recorren los eritrocitos en un tubo de ensayo en una hora. Esta medición permite seguir el progreso de enfermedades inflamatorias crónicas o determinar la efectividad de un tratamiento.

Estudios muestran que la VSG, en mujeres, presenta un comportamiento Log-Normal con un valor esperado de 15.25 $\texttt{mm/hr}$ y un coeficiente de variación del 33 %; en cambio en hombres, el comportamiento de la VSG es de tipo Normal, donde el 10.75 % tiene una VSG mayor a 20.36 $\texttt{mm/hr}$, mientras que un 6.68 % presenta una VSG menor a 16.25 $\texttt{mm/hr}$.

(a) ¿Qué porcentajes, en mujeres y en hombres, se encuentran en preocupación médica, es decir, presentan una VSG sobre los 18.67 $\texttt{mm/hr}$?

(b) Asuma que la VSG en las muestras de sangre que llegan a un laboratorio son independientes y siguen las distribuciones antes descritas. Adicionalmente, se sabe que $1/3$ de las muestras son de hombres. ¿Cuál es la probabilidad de que en las próximas 10 mediciones al menos dos presenten valores que impliquen una preocupación médica?

(c) Bajo los supuestos del apartado anterior, ¿cuál es valor esperado del número de muestras para análisis VSG que deben procesar en el laboratorio para encontrar la primera muestra que implique preocupación médica?

Solución

Sea Y la VSG de mujeres y X la VSG de hombres.

Del enunciado se tiene que

$$Y \sim \text{Log-Normal}(\lambda,\, \zeta) \quad \text{y} \quad X \sim \text{Normal}(\mu,\, \sigma).$$

Además

$$\mu_Y = 15.25, \quad \delta_Y = 0.33,$$

$$x_{89.25\,\%} = \mu + \sigma \cdot \Phi^{-1}(0.8925) = 20.36 \tag{1}$$

$$x_{6.68\,\%} = \mu - \sigma \cdot \Phi^{-1}(0.9332) = 16.25, \tag{2}$$

donde $\Phi^{-1}(0.8925) = 1.24$ y $\Phi^{-1}(0.9332) = 1.50$.

Despejando

$$\zeta = \sqrt{\ln(1 + \delta_Y^2)} = 0.3215098, \quad \lambda = \ln(\mu_Y) - \frac{1}{2} \cdot \zeta^2 = 2.672895, \tag{3}$$

$$\sigma = 1.5 \quad \text{y} \quad \mu = 18.5.$$

(a) Se pide, para hombres,

$$p = 1 - \Phi\left(\frac{18.67 - 18.5}{1.5}\right) \approx 1 - \Phi(0.11) = 1 - 0.5438 = 0.4562$$

y para mujeres

$$q = 1 - \Phi\left(\frac{\ln(18.67) - 2.672895}{0.3215098}\right) \approx 1 - \Phi(0.79) = 1 - 0.7852 = 0.2148$$

(b) Por probabilidades totales

$$\theta = \frac{1}{3} \cdot 0.4562 + \frac{2}{3} \cdot 0.2148 \approx 0.2953$$

Sea Z el numero de mediciones que implican preocupación médica

$$Z \sim \text{Binomial}(n = 10,\, \theta)$$

Se pide

$$P(Z \geq 2) = 1 - P(Z < 2) = 1 - P(Z \leq 1) = 1 - 0.03020249 - 0.12656158 = 1 - 0.1567641 \approx 0.8432$$

(c) Sea W el número de muestras a procesar para encontrar la primera muestra que implique preocupación médica.

$$W \sim \text{Geométrica}(\theta).$$

Se pide

$$\text{E}(W) = \frac{1}{\theta} = \frac{1}{0.2953} = 3.386387.$$

Asignación de Puntaje:

Logro 1: Obtener la ecuación (1) **[0.1 Ptos]** y la ecuación (2) **[0.1 Ptos]**. Obtener $\mu = 18.5$ **[0.2 Ptos]**, $\sigma = 1.5$ **[0.2 Ptos]** y $p = 0.4562$ **[0.4 Ptos]**.

Logro 2: Obtener las ecuaciones en (3) **[0.2 Ptos]**, obtener $\lambda = 2.672895$ **[0.2 Ptos]**, $\zeta = 0.321509$ **[0.2 Ptos]** y $q = 0.2148$ **[0.4 Ptos]**.

Si el alumno aproxima ζ por c.o.v., es decir, $\zeta \approx 0.33$ no descontar puntaje y q sería igual a 0.2182.

Logro 3: Obtener $\theta = \frac{1}{3} \cdot 0.4562 + \frac{2}{3} \cdot 0.2148 \approx 0.2953$ **[1.0 Ptos]**.

Si aproximo ζ por c.o.v., el valor de θ sería igual a 0.2975. No descontar puntaje en este caso.

Logro 4: Definir $Z \sim$ Binomial$(n = 10,\, \theta)$ **[0.4 Ptos]** y reconocer que se pide calcular $P(Z \geq 2)$ **[0.1 Ptos]** Resultado final $P(Z \geq 2) \approx 0.843$ **[0.5 Ptos]**.

Si aproximo ζ por c.o.v., el valor de $P(Z \geq 2)$ sería igual a 0.8467. No descontar puntaje en este caso.

Logro 5: Definir $W \sim$ Geométrica(θ) **[1.0 Ptos]**.

Logro 6: Obtener E$(W) = 3.386387$ **[1.0 Ptos]**.

Si aproximo ζ por c.o.v., el valor de E(W) sería igual a 3.361345. No descontar puntaje en este caso.

+ 1 Punto Base

Problema 2

La OMS establece que la contaminación del aire es un factor de riesgo crítico para las enfermedades no transmisibles, causando aproximadamente una cuarta parte de todas las muertes de adultos por enfermedades cardíacas y accidentes cerebrovasculares. Particularmente peligrosa es la contaminación del aire por material particulado respirable fino PM2.5, que corresponde al material formado por partículas cuyo diámetro aerodinámico es menor o igual a $2.5\ \mu\mathtt{m}$.

Suponga que el diámetro aerodinámico de una partícula PM2.5 se comporta como una variable aleatoria con distribución Uniforme entre 0 y $2.5\ \mu\mathtt{m}$. y que la concentración de estas partículas en el aire en un día dado sigue una distribución Exponencial con valor esperado igual a 25 $\mu\mathtt{g}/\mathtt{m}^3$.

(a) Encuentre el valor esperado y varianza del volumen de una partícula PM2.5 escogida al azar. Puede asumir que estas partículas son esféricas y que el volumen de una esfera está dado por $(4/3)\,\pi\,r^3$, donde r corresponde al radio de la partícula.

(b) La norma chilena establece estado de alerta ambiental cuando la concentración de material particulado respirable fino PM2.5 está entre 80 y 109 $\mu\mathtt{g}/\mathtt{m}^3$. Encuentre la probabilidad de que en un día dado se produzca un estado de alerta ambiental.

Solución

(a) El volumen de una partícula está dado por:

$$V = \frac{4}{3}\pi\left(\frac{X}{2}\right)^3 = \frac{\pi}{6}X^3,$$

donde X corresponde al diámetro de la partícula, que distribuye Uniforme$(0,\,2.5)$.

Se pide:

$$\mu_V = \mathsf{E}(V) = \frac{\pi}{6}\cdot\mathsf{E}(X^3) = \frac{\pi}{6}\int_0^{2.5} x^3\,\left(\frac{1}{2.5-0}\right)\,dx = \frac{\pi}{6}\cdot\frac{1}{2.5}\cdot\left.\frac{x^4}{4}\right|_0^{2.5} = \frac{\pi}{6}\cdot\frac{1}{2.5}\cdot\frac{2.5^4}{4} = 2.04.$$

Para la varianza se requiere:

$$\mathsf{E}(V^2) = \frac{\pi^2}{6^2}\cdot\mathsf{E}(X^6) = \frac{\pi^2}{6^2}\int_0^{2.5} x^6\,\left(\frac{1}{2.5-0}\right)\,dx = \frac{\pi^2}{90}\cdot\left.\frac{x^7}{7}\right|_0^{2.5} = \frac{\pi^2}{6^2}\cdot\frac{1}{2.5}\cdot\frac{2.5^7}{7} = 9.56.$$

Luego, la varianza pedida corresponde a:

$$\sigma_V^2 = \mathsf{E}(V^2) - \mu_V^2 = 9.56 - 2.04^2 = 5.40.$$

b) Sea Y la concentración de partículas PM2.5 en un día cualquiera.

$$Y \sim \text{Exponencial}(\lambda).$$

Del formulario se tiene que

$$\mu_Y = \frac{1}{\lambda} = 25 \to \lambda = 0.04.$$

Luego, la probabilidad de alerta ambiental en un día dado está dada por:

$$P(80 < Y \leq 109) = F_Y(109) - F_Y(80) = \left(1 - e^{-109\,\lambda}\right) - \left(1 - e^{-80\,\lambda}\right) = e^{-80\,\lambda} - e^{-109\,\lambda} = 0.028.$$

Asignación de Puntaje:

Logro 1: Obtener $\text{E}(X^3) = \dfrac{2.5^3}{4}$ **[0.5 Ptos]** y $\text{E}(V) = 2.04$ **[0.5 Ptos]**.

Logro 2: Obtener $\text{E}(X^6) = \dfrac{2.5^6}{7}$ **[0.5 Ptos]** y $\text{E}(V^2) = 9.56$ **[0.5 Ptos]**.

Logro 3: Obtener $\sigma_V^2 = 5.40$ **[1.0 Ptos]**.

Logro 4: Obtener $\lambda = 0.04$ **[1.0 Ptos]**.

Logro 5: Obtener $P(80 < Y \leq 109) = F_Y(109) - F_Y(80)$ **[1.0 Ptos]**.

Logro 6: Obtener $P(80 < Y \leq 109) = 0.028$ **[1.0 Ptos]**.

+ 1 Punto Base

Problema 3

Suponga que usted está estudiando el comportamiento de la velocidad del viento en una región determinada, la cual puede variar considerablemente en función de varios factores, como la ubicación geográfica, la hora del día, las condiciones climáticas, entre otros.

Para modelar la distribución de la velocidad del viento ($\mathtt{m/s}$), le recomiendan utilizar la distribución Rayleigh(α), cuya función de densidad está dada por

$$f_X(x) = \frac{2\,x\,e^{-x^2/\alpha}}{\alpha}, \quad x > 0, \alpha > 0.$$

Esta distribución es especialmente útil en situaciones donde la magnitud de la variable de interés (en este caso, la velocidad del viento) es una combinación de múltiples factores aleatorios que contribuyen de manera independiente al resultado final.

(a) Si a partir de mediciones observa que la velocidad mediana en el sector es de 5 $\mathtt{m/s}$, ¿qué valor sería adecuado asignar al parámetro α?

(b) Determine el coeficiente de variación de este modelo.. En el caso que dependa de α, utilice el valor obtenido en (a).

Hint: Puede ser de utilidad la función $\Gamma(\cdot)$ y sus propiedades:

$$(1) \quad \Gamma(k) = \int_0^\infty u^{k-1}\,e^{-u}\,du; \quad (2) \quad \Gamma(a+1) = a\,\Gamma(a); \quad (3) \quad \Gamma(n+1) = n!, \quad \textit{si}\; n \in \mathbb{N}_0; \quad (4) \quad \Gamma(1/2) = \sqrt{\pi}.$$

Solución

(a) Tenemos que

$$F_X(x_p) = \int_0^{x_p} \frac{2\,x\,e^{-x^2/\alpha}}{\alpha}\,dx = 1 - e^{-x^2/\alpha} \to x_p = \sqrt{-\alpha\,\ln(1-p)}.$$

Reemplazando

$$x_{50\,\%} = \sqrt{\alpha\,\ln(2)} = 5 \to \alpha = 36.06738.$$

(b) Tenemos que

$$\mu_X = \mathsf{E}(X) = \int_0^\infty x\cdot\frac{2\,x\,e^{-x^2/\alpha}}{\alpha}\,dx = \alpha^{1/2}\int_0^\infty y^{3/2-1}\,e^{-y}\,dy = \alpha^{1/2}\cdot\Gamma\left(\frac{3}{2}\right) = \alpha^{1/2}\cdot\frac{1}{2}\cdot\Gamma\left(\frac{1}{2}\right) = \frac{\sqrt{\alpha\,\pi}}{2}.$$

y

$$\mathsf{E}(X^2) = \int_0^\infty x^2\cdot\frac{2\,x\,e^{-x^2/\alpha}}{\alpha}\,dx = \alpha\int_0^\infty y^{2-1}\,e^{-y}\,dy = \alpha\,\cdot\,\Gamma(1) = \alpha.$$

Por lo tanto

$$\sigma_X^2 = \mathsf{E}(X^2) - \mu_X^2 = \frac{\alpha\,(4-\pi)}{4} \to \delta_X = \frac{\sigma_X}{\mu_X} = \sqrt{\frac{4-\pi}{\pi}} = 0.5227232.$$

Asignación de Puntaje:

Logro 1: Obtener $F_X(x_p) = 1 - e^{-x^2/\alpha}$. **[1.0 Ptos]**

Logro 2: Obtener $x_{50\,\%} = \sqrt{-\alpha\,\ln(1-0.5)}$. **[1.0 Ptos]**

Logro 3: Indicar que $\alpha = 36.06738$. **[1.0 Ptos]**

Logro 4: Obtener que $\mathrm{E}(X) = \dfrac{\sqrt{\alpha\,\pi}}{2}$. **[1.0 Ptos]**

Logro 5: Obtener que $\mathrm{E}(X^2) = \alpha$. **[1.0 Ptos]**

Logro 6: Obtener $\delta_X = \sqrt{\dfrac{4-\pi}{\pi}}$ o que $\delta_X = 0.5227232$ **[1.0 Ptos]**

+ 1 Punto Base