▶ **1.** Se examinan 150 piezas recién fabricadas y se registra el número de imperfecciones por pieza (se supone que las piezas no deben tener imperfecciones) resultando los siguientes datos:

No de imperf. por pieza	0	1	2	3	4	5	6	7
Frecuencia observada	18	37	42	30	13	7	2	1

Sea X el número de imperfecciones en una pieza seleccionada al azar y suponga que X tiene distribución Poisson de parámetro λ.

a) Encuentre un estimador insesgado de λ y calcule la estimación para los datos anteriores.
b) ¿Cuál es la desviación estándar de este estimador? Calcule una estimación del error estándar del estimador.
c) Considere los siguientes estimadores de λ, basados en una muestra aleatoria $X_1, ..., X_n$:

$$\hat{\lambda_1} = \bar{X}, \qquad \hat{\lambda_2} = (X_1 + X_n)/2 \quad \text{y} \quad \hat{\lambda_3} = (X_1 + 2X_2 + X_n)/3, \qquad \text{con } n \geq 3$$

¿Cuáles son estimadores insesgados para λ?
d) Entre los estimadores insesgados para λ del ítem c), ¿cuál tiene menor varianza?

a) Podemos utilizar la media muestral como estimador insesgado de λ dado que es una distribución Poisson y $E(x) = \lambda$.

$$\bar{X} = \frac{\sum (X_i \cdot f_i)}{n} = \hat{\lambda}$$

$n = 150$

$\Rightarrow 0 \cdot 18 + 1 \cdot 37 + 2 \cdot 42 + 3 \cdot 30 + 4 \cdot 13 + 5 \cdot 7 + 6 \cdot 2 + 7 \cdot 1 = 317$

$\frac{317}{150} \approx 2{,}11$

b) la d.e. del estimador es $\sqrt{\lambda}$

$\therefore \sqrt{\hat{\lambda}} = \sqrt{2{,}11} \approx 1{,}45$

c) Considere los siguientes estimadores de λ, basados en una muestra aleatoria $X_1, ..., X_n$:

$$\hat{\lambda_1} = \bar{X}, \qquad \hat{\lambda_2} = (X_1 + X_n)/2 \quad \text{y} \quad \hat{\lambda_3} = (X_1 + 2X_2 + X_n)/3, \qquad \text{con } n \geq 3$$

¿Cuáles son estimadores insesgados para λ?

Para el **inciso c** del ejercicio 1, tenemos tres estimadores propuestos para λ y debemos determinar cuáles son insesgados. Los estimadores son:

1. $\hat{\lambda}_1 = \bar{X}$
2. $\hat{\lambda}_2 = \frac{X_1 + X_n}{2}$
3. $\hat{\lambda}_3 = \frac{X_1 + 2X_2 + X_n}{3}$

Donde:

- $X_1, X_2, \ldots, X_n$ son observaciones de la variable aleatoria X, que sigue una distribución de Poisson con parámetro λ.

Para determinar si un estimador es insesgado, verificamos si su valor esperado es igual a λ. Es decir, si $E(\hat{\lambda}) = \lambda$, entonces el estimador es insesgado.

$E(\hat{\lambda}_1) = E(\bar{X}) = \lambda$ ✓

$E(\hat{\lambda}_2) = E((X_1 + X_n)/2) = E(X_1 + X_n)/2 = \frac{\lambda + \lambda}{2} = \lambda$ ✓

$E(\hat{\lambda}_3) = E((X_1 + 2X_2 + X_n)/3) = \dots = \dfrac{E(X_1) + 2E(X_2) + E(X_n)}{3}$

$= \dfrac{\lambda + 2\lambda + \lambda}{3} = \dfrac{4\lambda}{3} \neq \lambda$

$\therefore$ Los insesgados son $\hat{\lambda}_1$ y $\hat{\lambda}_2$

d) $Var(\hat{\lambda}_1) = \dfrac{\lambda}{n}$ con $n = 150$

$\dfrac{Var(X_1)}{n} = \dfrac{\lambda}{n} = \dfrac{\lambda}{150}$

$Var(\hat{\lambda}_2) = V\left(\dfrac{X_1 + X_2}{2}\right) = \dfrac{V(X_1)}{2^2} + \dfrac{V(X_2)}{2^2} = \dfrac{\lambda + \lambda}{4} = \dfrac{\lambda}{2}$

$\min\left(\dfrac{\lambda}{150}, \dfrac{\lambda}{2}\right) = \dfrac{\lambda}{150} = \lambda_1$

$\therefore$ λ_1 tiene menor varianza

► **2.** X_1 , ... , X_m es una muestra aleatoria de una distribución de resistencia de vigas de concreto con media μ_1 y desviación estándar σ_1 e $Y_1, ..., Y_n$ es una muestra aleatoria de una distribución de resistencia de cilindros de concreto con media μ_2 y desviación estándar σ_2. Suponga que $X_1, ..., X_m$ e $Y_1, ..., Y_n$ son muestras aleatorias independientes entre sí. Se obtuvieron las siguientes observaciones:
Resistencias de vigas de concreto:

5,9 7,2 7,3 6,3 8,1 6,8 7,0 7,6 6,8 6,5 7,0 6,3 7,9 9,0 8,2 8,7 7,8 9,7 7,4 7,7

Resistencias de cilindros de concreto:

6,1 5,8 7,8 7,1 7,2 9,2 6,6 8,3 7,0 8,3 7,8 8,1 7,4 8,5 8,9 9,8 9,7 14,1 12,6 11,2

a) Calcule una estimación para μ_1, μ_2, σ_1, σ_2
b) Dé un estimador insesgado de $\mu_1 - \mu_2$. Calcule una estimación de dicha diferencia.
c) Obtenga la varianza y la desviación estándar (error estándar) del estimador del inciso b)
d) Calcule una estimación puntual de la relación σ_1/σ_2.

a) $X_j = H_1(\mu_1, \sigma_1^2)$
$Y_j = H_2(\mu_2, \sigma_2^2)$

$\hat{\mu} = \overline{X}$

$\hat{\mu}_1$ es la media de la resistencia de las vigas $= 7,86$
$\hat{\mu}_2$ " " " " los cilindros $\approx 8,58$

$\hat{\sigma} = S = \sqrt{\dfrac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X})^2}$ c/ corrección de Bessel

$\hat{\sigma}_1 = 0,97$ $\hat{\sigma}_2 = 2,10$

Para el **inciso b** del ejercicio 2, necesitamos:

1. **Encontrar un estimador insesgado de** $\mu_1 - \mu_2$, es decir, la diferencia entre las medias de las resistencias de vigas y cilindros de concreto.
2. **Calcular una estimación** de esta diferencia usando los datos de las muestras.

Paso a Paso para Resolver el Inciso b

1. Estimador Insesgado de $\mu_1 - \mu_2$

Si $X_1, X_2, \ldots, X_m$ y $Y_1, Y_2, \ldots, Y_n$ son muestras independientes con medias μ_1 y μ_2 y varianzas σ_1^2 y σ_2^2, respectivamente, un estimador insesgado de la diferencia $\mu_1 - \mu_2$ es:

$$\hat{\mu}_1 - \hat{\mu}_2 = \bar{X} - \bar{Y}$$

donde:

- $\bar{X}$ es la media muestral de la resistencia de las vigas (X).
- $\bar{Y}$ es la media muestral de la resistencia de los cilindros (Y).

Esto se debe a que la media de la diferencia $\bar{X} - \bar{Y}$ es $E(\bar{X} - \bar{Y}) = \mu_1 - \mu_2$, por lo tanto, este es un estimador insesgado.

2. Cálculo de la Estimación de $\mu_1 - \mu_2$

Usando los valores calculados previamente:

- Media muestral de vigas, $\hat{\mu}_1 = 7.46$
- Media muestral de cilindros, $\hat{\mu}_2 = 8.58$

La estimación de $\mu_1 - \mu_2$ es:

$$\hat{\mu}_1 - \hat{\mu}_2 = 7.46 - 8.58 = -1.12$$

Paso a Paso para Resolver el Inciso c

1. Varianza del Estimador de la Diferencia $\hat{\mu}_1 - \hat{\mu}_2$

Dado que $\hat{\mu}_1$ y $\hat{\mu}_2$ son estimadores insesgados de μ_1 y μ_2, y que las muestras son independientes, la varianza de la diferencia $\hat{\mu}_1 - \hat{\mu}_2$ es la suma de las varianzas individuales:

$$\mathrm{Var}(\hat{\mu}_1 - \hat{\mu}_2) = \frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}$$

donde:

- σ_1 y σ_2 son las desviaciones estándar de las resistencias de vigas y cilindros, respectivamente.
- $m = 20$ es el tamaño de la muestra de las vigas.
- $n = 20$ es el tamaño de la muestra de los cilindros.

Usaremos las estimaciones de σ_1 y σ_2 obtenidas en el inciso a:

- $\hat{\sigma}_1 = 0.97$
- $\hat{\sigma}_2 = 2.10$

2. Cálculo del Error Estándar de la Diferencia

El **error estándar** es la raíz cuadrada de la varianza del estimador:

$$\text{Error estándar} = \sqrt{\mathrm{Var}(\hat{\mu}_1 - \hat{\mu}_2)}$$

Resultados del Inciso c

1. **Varianza del estimador de la diferencia** $\mathrm{Var}(\hat{\mu}_1 - \hat{\mu}_2)$:

$$\mathrm{Var}(\hat{\mu}_1 - \hat{\mu}_2) = 0.27$$

2. **Error estándar** del estimador de la diferencia (desviación estándar):

$$\text{Error estándar} = 0.52$$

Esto nos proporciona una medida de la precisión del estimador de la diferencia $\hat{\mu}_1 - \hat{\mu}_2$.

d) $\frac{\hat{\sigma}_1}{\hat{\sigma}_2} = \frac{0,97}{2,10} \approx 0,46$

3. Se seleccionan al azar n_1 fumadores (hombres) y n_2 fumadoras. Sean X_1 y X_2 variables aleatorias independientes que denotan, respectivamente, el número de fumadores y fumadoras que fuman cigarrillos con filtro. Denotamos con p_1 y p_2 las respectivas probabilidades de que un hombre y una mujer seleccionados al azar fumen cigarrillos con filtro.

a) Demuestre que $(X_1/n_1) - (X_2/n_2)$ es un estimador insesgado para $p_1 - p_2$.

b) ¿Cuál es el error estándar del estimador del inciso a)?

c) Si $n_1 = n_2 = 200$, $x_1 = 127$ y $x_2 = 176$, obtenga una estimación de $p_1 - p_2$.

d) Con los datos del inciso c) obtenga una estimación del error estándar del estimador de $p_1 - p_2$.

a) tengo que ver $\forall p_1 - p_2 \in \mathbb{R}$

Paso 1: Definir el estimador de p_1 y p_2

Dado que:

- X_1 es el número de hombres que fuman cigarrillos con filtro entre un total de n_1 hombres, entonces $\frac{X_1}{n_1}$ es un estimador de p_1.
- X_2 es el número de mujeres que fuman cigarrillos con filtro entre un total de n_2 mujeres, por lo tanto $\frac{X_2}{n_2}$ es un estimador de p_2.

Así, $\frac{X_1}{n_1} - \frac{X_2}{n_2}$ es un estimador para la diferencia $p_1 - p_2$.

Queremos demostrar que:

$$E\left(\frac{X_1}{n_1} - \frac{X_2}{n_2}\right) = p_1 - p_2$$

Dado que $X_1 \sim B(n_1, p_1)$
$X_2 \sim B(n_2, p_2)$

$$E(X_1) = n_1 p_1 \qquad E(X_2) = n_2 p_2$$

$$\Rightarrow E\left(\frac{X_1}{n_1} - \frac{X_2}{n_2}\right) = \frac{E(X_1)}{n_1} - \frac{E(X_2)}{n_2} = \frac{n_1 p_1}{n_1} - \frac{n_2 p_2}{n_2} = p_1 - p_2$$

b)

Paso 1: Varianza del Estimador

Recordemos que el **error estándar** es la raíz cuadrada de la **varianza** del estimador. Dado que X_1 y X_2 son variables aleatorias binomiales independientes, la varianza del estimador $\frac{X_1}{n_1} - \frac{X_2}{n_2}$ se calcula como:

$$\text{Var}\left(\frac{X_1}{n_1} - \frac{X_2}{n_2}\right) = \text{Var}\left(\frac{X_1}{n_1}\right) + \text{Var}\left(\frac{X_2}{n_2}\right)$$

Esto se debe a que X_1 y X_2 son independientes, por lo que la covarianza entre ellas es cero.

Sabemos que $Var(X) = np(1-p)$

$$\therefore \quad Var\left(\frac{X_1}{n_1}\right) = \frac{Var(X_1)}{n_1^2} = \frac{n_1 p_1 (1-p_1)}{n_1^2} = \frac{p_1(1-p_1)}{n_1}$$

$$\text{Var}(X_1 - X_2) = \text{Var}(X_1) + \text{Var}(X_2)$$

$$Var\left(\frac{X_2}{n_2}\right) = \dots = \frac{p_2(1-p_2)}{n_2}$$

$$\Rightarrow var\left(\frac{X_1}{n_1} - \frac{X_2}{n_2}\right) = \frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}$$

$\therefore$

$$\text{Error estandar} = \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$$

c) $\hat{p}_1 = \frac{x_1}{n_1} = \frac{127}{200} \approx 0{,}635$ $\hat{p}_2 = \frac{x_2}{n_2} = \frac{176}{200} \approx 0{,}88$

Estimar $p_1 - p_2$

$\hat{p}_1 - \hat{p}_2 = 0{,}635 - 0{,}88 = -0{,}245$

Conclusión

La estimación puntual de la diferencia $p_1 - p_2$ es:

$$\hat{p}_1 - \hat{p}_2 = -0.245$$

Esto sugiere que, en promedio, la proporción de hombres que fuman cigarrillos con filtro es aproximadamente 24.5% menor que la de mujeres en esta muestra.

- Si $\hat{p}_1 - \hat{p}_2 > 0$, significa que $p_1 > p_2$ (mayor proporción en hombres).
- Si $\hat{p}_1 - \hat{p}_2 < 0$, significa que $p_1 < p_2$ (mayor proporción en mujeres).

d) Hay que sustituir en la fórmula:

Error estandar $= \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} = \sqrt{\frac{0{,}635(1-0{,}635)}{200} + \frac{0{,}88(1-0{,}88)}{200}}$

$\approx \boxed{0{,}0411}$

▶ **4.** Sea $X_1, ..., X_n$ una muestra aleatoria, donde cada X_i tiene media μ y varianza σ^2.
a) Demuestre que $\bar{X}^2$ no es un estimador insesgado para μ^2
b) ¿Para qué valor de k es el estimador $\hat{X}^2 - kS^2_{n-1}$ insesgado para μ^2?

a) Para demostrar que $\bar{x}^2$ no es un estimador insesgado para μ^2, necesitamos calcular su valor esperado y compararlo con μ^2

$E(\bar{X}) = \mu$

$Var(\bar{x}) = \frac{\sigma^2}{n}$ y sabemos $Var(\bar{x}) = E(\bar{x}^2) - (E(\bar{x}))^2$

$\therefore \frac{\sigma^2}{n} = E(x^2) - \mu^2$

$E(x^2) = \mu^2 + \frac{\sigma^2}{n} \neq \mu^2$

b)

$E\left(\bar{x}^2 - k\frac{S^2}{n-1}\right) = \mu^2$ $E(x^2) = \mu^2 - \frac{\sigma^2}{n}$

Esperanza de S^2: Recordemos que S^2 es un estimador insesgado de σ^2, por lo que:

$$\mathbb{E}\left(\frac{S^2}{n-1}\right) = \frac{\sigma^2}{n-1}$$

$\therefore$

$E\left(\bar{x}^2 - k\frac{S^2}{n-1}\right) = \mu^2 + \frac{\sigma^2}{n} - k\frac{\sigma^2}{n-1}$

Despejemos k

$$\cancel{\mu^2} + \frac{\sigma^2}{n} - k\,\frac{\sigma^2}{n-1} = \cancel{\mu^2}$$

$$\frac{\sigma^2}{n} = k\,\frac{\sigma^2}{n-1}$$

$$\frac{1}{n} = \frac{k}{n-1}$$

$$\boxed{k = \frac{n-1}{n}}$$

▶ **5.** Sea $X_1, ..., X_n$ una muestra aleatoria de la función de densidad

$$f(x) = \begin{cases} 0{,}5(1+\theta x) & -1 \leq x \leq 1 \\ 0 & \text{en cualquier otro caso} \end{cases}$$

donde $-1 \leq \theta \leq 1$ (esta distribución aparece en física de partículas). Demuestre que $\hat{\theta} = 3\bar{X}$ es un estimador insesgado de θ y calcule su varianza.

Para demostrar que $\hat{\theta} = 3\bar{X}$ es un estimador insesgado de θ, debemos verificar que el **valor esperado** de $\hat{\theta}$ sea igual a θ.

Paso 1: Valor Esperado de $\bar{X}$

Primero, necesitamos encontrar $\mathbb{E}(\bar{X})$ en términos de θ. Dado que $\bar{X}$ es la media muestral de $X_1, X_2, \ldots, X_n$, y todos los X_i son idénticamente distribuidos, tenemos:

$$\mathbb{E}(\bar{X}) = \mathbb{E}(X)$$

Entonces, debemos encontrar $\mathbb{E}(X)$ para la distribución dada por $f(x)$.

$$E(X) = \int_{-1}^{1} x f(x)\,dx = \int_{-1}^{1} x\, 0{,}5(1+\theta x)\,dx = \int_{-1}^{1} 0{,}5x + 0{,}5\theta x^2\,dx$$

$$= \int_{-1}^{1} 0{,}5x\,dx + \int_{-1}^{1} 0{,}5\theta x^2\,dx = \ldots = \frac{1}{3}\theta$$

Comprobar $\hat{\theta} = 3\bar{X}$ es insesgado

$$E(\bar{X}) = E(X) = \frac{\theta}{3}$$

$$\Rightarrow E(\hat{\theta}) = E(3\bar{X}) = 3E(\bar{X}) = 3\,\frac{\theta}{3} = \theta$$

∴ sí es estimador insesgado.

$$Var(\hat{\theta}) = Var(3\bar{X}) = 9 \cdot Var(\bar{X})$$

$$Var(\bar{X}) = \frac{Var(X)}{n} \qquad Var(X) = \frac{1}{3} - \frac{\theta^2}{9}$$

$$\therefore Var(\hat{\theta}) = \ldots = \frac{9 \cdot \left(\frac{1}{3} - \frac{\theta^2}{9}\right)}{n} = \boxed{\frac{3 - \theta^2}{n}}$$

► **6.** Se denota con X la proporción de tiempo que un estudiante, seleccionado al azar, emplea trabajando en cierta prueba de aptitud. Se supone que X tiene función de densidad dada por:

$$f(x) = \begin{cases} (\theta+1)x^\theta & 0 \leq x \leq 1 \\ 0 & \text{en cualquier otro caso} \end{cases}$$

donde $-1 < \theta$.
a) Obtenga por el método de los momentos un estimador de θ.
b) Se toma una muestra aleatoria de 10 estudiantes obteniéndose las siguientes observaciones:

0,91   0,79   0,90   0,65   0,86   0,47   0,73   0,97   0,94   0,77

Calcule con esta información una estimación de θ, usando el estimador obtenido en el inciso a).

a)

Inciso a): Obtener un estimador de θ mediante el método de los momentos.

- Para aplicar el método de los momentos, calculamos el primer momento de X, $E(X)$.
- Usamos la fórmula $E(X) = \int_0^1 x \cdot f(x)\,dx$, reemplazando $f(x)$ y resolviendo la integral en función de θ.
- Luego, igualamos el momento teórico $E(X)$ con el momento muestral (la media muestral $\bar{X}$) para despejar θ.

$$E(x) = \int_0^1 x \cdot (\theta+1) x^\theta \, dx = (\theta+1) \int_0^1 x^{\theta+1} \, dx = \ldots$$

$$= (\theta+1) \int_0^1 x^a \, dx = (\theta+1) \frac{x^{a+1}}{a+1} \qquad a \neq -1$$

$$= (\theta+1) \left[\frac{x^{\theta+2}}{\theta+2} \right]_0^1 = (\theta+1) \left(\frac{1}{\theta+2} - \frac{0}{\theta+2} \right) = \boxed{\frac{\theta+1}{\theta+2}}$$

Igualamos momento muestral:

$$\bar{X} = \frac{\theta+1}{\theta+2}$$

Ahora despejamos θ

$$\bar{X} = \frac{\theta+1}{\theta+2}$$

$$\bar{x}(\theta+2) = \theta+1$$

$$\bar{X}\theta + 2\bar{X} = \theta + 1$$

$$\bar{X}\theta - \theta = 1 - 2\bar{X}$$

$$\theta(\bar{x}-1) = 1 - 2\bar{x}$$

$$\theta = \frac{1-2\bar{X}}{\bar{x}-1}$$

$$\theta = \frac{1-2\bar{X}}{\bar{X}-1}$$

$\bar{X} - 1$

Este es el estimador de θ por el método de los momentos. En el inciso b), simplemente reemplazaremos $\bar{X}$ con la media de los datos observados para obtener una estimación puntual de θ.

b)

$$\bar{X} = \frac{0,91 + 0,79 + 0,90 + 0,65 + 0,86 + 0,47 + 0,73 + 0,97 + 0,94 + 0,77}{10} = \frac{7,99}{10} = 0,799$$

Reemplazamos:

$$\theta = \frac{1 - 2 \cdot 0,799}{0,799 - 1} = \frac{-0,598}{-0,201} \approx \boxed{2,975}$$

7. Se supone que el espesor de pintura de baja viscosidad (X) tiene distribución normal. Se observaron las siguientes observaciones de espesores de pintura de baja viscosidad:

0,83 0,88 0,88 1,04 1,09 1,12 1,29 1,31 1,48 1,49 1,59 1,62 1,65 1,71 1,76 1,83

a) Calcule una estimación puntual de la media de la distribución del espesor de pintura por el método de los momentos.
b) Calcule una estimación puntual de la mediana de la distribución del espesor de pintura por el método de máxima verosimilitud (MV).
c) Calcule una estimación del percentil 90 de la distribución del espesor de pintura por el método de MV.
d) Estime $P(X < 1{,}5)$ por el método de MV.
e) ¿Cuál es el error estándar del estimador usado en el inciso a)?

a) $E(x_i) = \mu$

Estimador: $\bar{X} = \frac{\sum_{j=1}^{n} X_j}{n}$

(0.83+0.88+0.88+1.04+1.09+1.12+1.29+1.31+1.48+1.49+1.59+1.62+1.65+1.71+1.76+1.83)/16 = 1.3481

b)

Estimador:

$$\hat{\mu} = \frac{\sum_{j=1}^{n} X_j}{n}$$

Estimación para los datos:
1.3481

Inciso b: Estimación de la mediana por el método de máxima verosimilitud (MV)

Para una distribución normal, la media y la mediana son iguales, ya que es simétrica. Por lo tanto, la mediana es igual a la media muestral calculada en el inciso a.

c)

Para una distribución normal $N(\mu, \sigma^2)$, el percentil p se calcula como:

$$P_p = \mu + z_p \cdot \sigma$$

donde z_p es el valor z correspondiente al percentil 90, que es aproximadamente $z_{0.9} \approx 1.28$.

d)

Inciso d: Estimación de $P(X < 1.5)$ usando MV

Para calcular esta probabilidad, usamos la función de distribución acumulativa (CDF) de una normal. En términos de la CDF de la normal estándar Φ:

$$P(X < 1.5) = \Phi\left(\frac{1.5 - \mu}{\sigma}\right)$$

$$F_Z\left(\frac{1.5 - 1.3481}{\sqrt{0.1074}}\right)$$

$$\simeq F_Z(0.46)$$

$$\simeq 0.6772$$

e)

$$\sigma_{\hat{\mu}} = \sqrt{V(\hat{\mu})} = \sqrt{V(\bar{X})} = \sqrt{\frac{\sigma^2}{n}} = \frac{\sigma}{\sqrt{n}}$$

8. El tiempo de respuesta X (en segundos) de cierta terminal de computadoras tiene distribución exponencial de parámetro $\lambda > 0$.

a) Obtenga el estimador de máxima verosimilitud de λ.

b) Se hacen 10 observaciones resultando los valores:

3,11 0,64 2,55 2,20 5,44 3,42 10,39 8,93 17,82 1,30

Calcule, con estos datos, una estimación para λ usando el estimador obtenido en el inciso a).

$X \sim \mathcal{E}(\lambda)$ con $\lambda > 0$ $\quad f(x) = \lambda e^{-\lambda x},\ x \geq 0 \wedge \lambda \geq 0$

a) Función de verosimilitud

$$L(\lambda) = \prod_{i=1}^{n} \lambda e^{-\lambda x} = \lambda^n e^{-\lambda \sum_{i=1}^{n} x_i}$$

Para simplificar se toma el log (log-verosimilitud)

$$\ln(L(\lambda)) = n\ln(\lambda) - \lambda \sum_{i=1}^{n} x_i$$

Para encontrar el valor que maximiza la verosimilitud, derivamos la log-v respecto λ e igualamos a 0.

$$\frac{d}{d\lambda} \ln(L(\lambda)) = \frac{d}{d\lambda}\left(n\ln(\lambda) - \lambda \sum_{i=1}^{n} x_i\right)$$

$$\frac{d}{d\lambda} \ln(L(\lambda)) = \frac{n}{\lambda} - \sum_{i=1}^{n} x_i = 0$$

$$\Rightarrow \frac{n}{\lambda} - \sum_{i=1}^{n} x_i = 0$$

$$\frac{n}{\lambda} = \sum_{i=1}^{n} x_i \quad \ldots\ldots \quad \lambda = \frac{n}{\sum_{i=1}^{n} x_i}$$

$$\Rightarrow \quad \hat{\lambda} = \frac{n}{n \cdot \overline{X}} = \boxed{\frac{1}{\overline{X}}}$$

b) $$\frac{3{,}11+0{,}64+2{,}55+2{,}20+5{,}44+3{,}42+10{,}39+8{,}93+17{,}82+1{,}30}{10} = \frac{55{,}8}{10} = 5{,}58 = \overline{X}$$

Ahora estimamos

$$\lambda(\hat{\lambda}) = \frac{1}{5{,}58} \approx \boxed{0{,}179}$$

9. Se supone que la resistencia al corte de soldaduras eléctricas (lb/plg^2) tiene distribución normal. Se determina la resistencia al corte de 10 soldaduras eléctricas obteniéndose:

392 376 401 367 389 362 409 415 358 375

a) Estime el verdadero promedio y la desviación estándar de la resistencia al corte, usando los estimadores de máxima verosimilitud.
b) Estime el percentil 95 de la distribución (use la propiedad de invarianza del estimador por MV).
c) Estime $P(X \leq 400)$ usando la propiedad de invarianza del estimador por MV.

$X \sim N(?, ?)$

a) Estimador para μ

$\hat{\mu} = \overline{X}$

Estimación:

(392+376+401+367+389+362+409+415+358+375)/10 = 384.4

Estimador para σ

$$\hat{\sigma} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (x_i - \overline{X})^2} \approx \boxed{18{,}86}$$

Por ser MV.

b) Para $N(\mu,\sigma^2)$ el percentil 95 se calcula como:

$$P_{95} = \mu + Z_{0,95} \cdot \sigma$$

donde $Z_{0,95}$ es el valor crítico para el percentil 95 en una distribución normal estandar (aprox $Z_{0,95} = 1,645$ para una cola unitaria o $Z_{0,95} = 1,96$ para una cola bilateral.

Usaremos $Z = 1,645$ para una cola unitara

$$\hat{P}_{95} = \hat{\mu} + Z_{0,95}\,\hat{\sigma} = 384,4 + 1,645 \cdot 18,86 = \boxed{415,4247}$$

c) $P(X \leq 400) = P\left(Z \leq \frac{400-\hat{\mu}}{\hat{\sigma}}\right) = P\left(Z \leq \frac{400-384,4}{18,86}\right) \approx P(Z \leq 0,826)$

$\simeq 0,796 = 79,6\%$

10. **a)** Sea $X_1, ..., X_n$ una muestra aleatoria de una distribución uniforme en el intervalo $[0, \theta]$. El estimador de máxima verosimilitud de θ es $\hat{\theta} = Y = \text{máx}(X_i)$. Obtenga la distribución acumulada de Y utilizando el hecho de que

$$Y \leq y \text{ si y sólo si cada } X_i \leq y$$

y demuestre que la función de densidad de Y es:

$$f(x) = \begin{cases} \dfrac{ny^{n-1}}{\theta^n} & 0 \leq y \leq \theta \\ 0 & \text{en cualquier otro caso} \end{cases}$$

b) Utilice el resultado de la parte a) para demostrar que el estimador de máxima verosimilitud de θ es sesgado pero que $(n+1)\text{máx}(X_i)/n$ es insesgado.
c) Probar que $2\bar{X}$ es un estimador insesgado para θ.
d) Entre los estimadores insesgados dados en los ítems b) y c), ¿cuál elegiría?

a) Dado que $Y = \max(X_1, ..., X_n)$ y cada X_i tiene una distribución uniforme en $[0, \theta]$ tenemos que:

i) Función de distribución acumulada $F_Y(y) = P(Y \leq y)$:

- Para que $Y \leq y$, es necesario que todas las observaciones $X_i \leq y$ (dado que Y es el máximo)
- La probabilidad de que una observación $X_i \leq y$ es $\frac{y}{\theta}$, dado que $X_i \sim U(0, \theta)$
- Dado que los X_i son independientes, la prob. de que todas las obs. estén por debajo de y es:

$$P(Y \leq y) = P(X_1 \leq y) \cdot P(X_2 \leq y) \cdots P(X_n \leq y) = \left(\frac{y}{\theta}\right)^n$$

$\therefore F_Y(y) = \left(\frac{y}{\theta}\right)^n \quad 0 \leq y \leq \theta$

ii) Función densidad $f_Y(y)$:

- La densidad $f_Y(y)$ se obtiene derivando FDA con respecto a y

$$f_Y(y) = \frac{d}{dy} F_Y(y) = \frac{d}{dy}\left(\frac{y}{\theta}\right)^n = \frac{n y^{n-1}}{\theta^n}$$

$\therefore$ $f_Y(y) = \frac{n y^{n-1}}{\theta^n}$, $0 \leq y \leq \theta$

b) i) Calcular esperanza de Y:

- La $E(Y)$ nos ayudará a ver si $\hat{\theta} = Y$ es insesgado
- La $E(Y)$ para una var con densidad $f_Y(y)$ se calcula:

$$E(Y) = \int_0^\theta y f_Y(y) dy = \int_0^\theta y \frac{n y^{n-1}}{\theta^n} dy$$

Esto da:

$$E(Y) = \frac{n}{\theta^n} \int_0^\theta y^n dy = \frac{n}{\theta^n} \cdot \frac{\theta^{n+1}}{n+1} = \frac{n\theta}{n+1}$$

Entonces

$$E(Y) = \frac{n\theta}{n+1}$$

ii) Corregir el sesgo:

- Dado $E(Y) = \frac{n\theta}{n+1}$, el estimador Y subestima el valor de θ
- Para obtener un estimador insesgado, podemos escalar Y por el factor $\frac{n+1}{n}$:

$$\hat{\theta}_{insesgado} = \frac{n+1}{n} Y = \frac{n+1}{n} \max(X_i)$$

$\therefore$ Este es un estimador insesgado de θ

c) i) $E(\bar{X})$

- Dado que $X_i \sim U(0,\theta)$, sabemos que $E(X_i) = \frac{\theta}{2}$
- como $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$, tenemos que:

$$E(\bar{X}) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n} \cdot n \cdot \frac{\theta}{2} = \frac{\theta}{2}$$

ii) $E(2\bar{X})$

- $E(2\bar{X}) = 2 \cdot E(\bar{X}) = 2 \cdot \frac{\theta}{2} = \theta$ ✓

∴ Esto demuestra que $2\bar{X}$ es un estimador insesgado de θ

d) Para elegir debemos considerar la varianza.

i) Varianza de $\hat{\theta}_1 = \frac{n+1}{n} \max(X_i)$

$E(\max(X_i)) = \frac{n\theta}{n+1}$

La varianza de $\max(X_i)$ en una distribución $U(0, \theta)$ es:

$$Var(\max(X_i)) = \frac{n\theta^2}{(n+1)^2(n+2)}$$

Dado que $\hat{\theta}_1 = \frac{n+1}{n} \max(X_i)$, la var es:

$$Var(\hat{\theta}_1) = \left(\frac{n+1}{n}\right)^2 Var(\max(X_i)) = \left(\frac{n+1}{n}\right)^2 \frac{n\theta^2}{(n+1)^2(n+2)}$$

simplificando:

$$Var(\hat{\theta}_1) = \frac{\theta^2}{n(n+2)}$$

ii) Varianza dado $\hat{\theta}_2 = 2\bar{X}$

- $E(X_i) = \frac{\theta}{2} \Rightarrow E(\bar{X}) = \frac{\theta}{2}$ ∴ $E(2\bar{X}) = \theta$

- La $var(X_i)$ para una va uniforme $= \frac{\theta^2}{12}$

- Como $\bar{X}$ es la media de n variables indep. $X_i \Rightarrow$

$$Var(\bar{X}) = \frac{Var(X_i)}{n} = \frac{\theta^2}{12n}$$

- Aplicación de la constante 2:

Dado que $\hat{\theta}_2 = 2\bar{X}$, la $var(\hat{\theta}_2)$ es

$$Var(\hat{\theta}_2) = (2)^2 \, Var(\bar{X}) = 4 \cdot \frac{\theta^2}{12n} = \frac{\theta^2}{3n}$$

iii) Comparación de varianzas

$$Var(\hat{\theta}_1) = \frac{\theta^2}{n(n+2)}$$

$$var(\hat{\theta}_2) = \frac{\theta^2}{3n}$$

Para comparar, observamos el comportamiento de cada varianza en función de n:

- A medida que n aumenta, ambas varianzas disminuyen, pero la varianza de $\hat{\theta}_1 = \frac{n+1}{n} \max(X_i)$ tiende a ser menor que la de $2\overline{X}$

Dado que $\frac{1}{n(n+2)} < \frac{1}{3n}$ para $n \geq 1$

$\therefore$ $\hat{\theta}_1$ es el estimador preferido porque tiene una menor varianza, lo que lo hace más eficiente para estimar θ en una distribución uniforme $[0, \theta]$