ERRORES

Errores en el Proceso de Medición

En todo proceso de medición existen limitaciones dadas por

- ✓ los instrumentos usados
- ✓ el método de medición
- ✓ el observador

El mismo proceso de medición introduce errores o incertezas.

En muchas situaciones puede pensarse que al medir una magnitud μ lo que observamos es

$$X = \mu + \varepsilon$$

donde ε representa un error de medición aditivo, que eventualmente podria ser una suma de pequeños errores.

ERRORES

Errores en el Proceso de Medición

Los instrumentos que usamos para medir como las magnitudes mismas son fuente de incertezas al momento de medir.

Los instrumentos tienen una *precisión finita, por lo tanto* siempre existe una variación mínima de la magnitud que puede detectar.

Ejemplo: con una regla graduada en milímetros, no podemos detectar variaciones menores que una fracción del milímetro.

Las magnitudes a medir no están definidas con infinita precisión.

Ejemplo: Si queremos medir el largo de una mesa, si usamos instrumentos cada vez más precisos empezamos a notar las irregularidades, con lo cual la magnitud a medir puede resultar imprecisa.

ERRORES

Errores en el Proceso de Medición:

Tipos de Errores:

Errores sistemáticos: (sesgo) surgen por falla del equipo o del diseño. No se pueden evaluar realizando medidas repetidas.

Ejemplo: balanza que aún vacía pesa más que 0.

·Errores aleatorios: surgen por efectos de variables no controladas. Siempre están presentes, nunca se pueden eliminar. Podemos minimizar su efecto y realizando medidas repetidas independientes se puede evaluar la variabilidad que introducen.

ERRORES

Errores en el Proceso de Medición

Precisión: la precisión de un instrumento o un método de medición está asociada a la sensibilidad o menor variación de la magnitud que se pueda detectar con dicho instrumento o método.

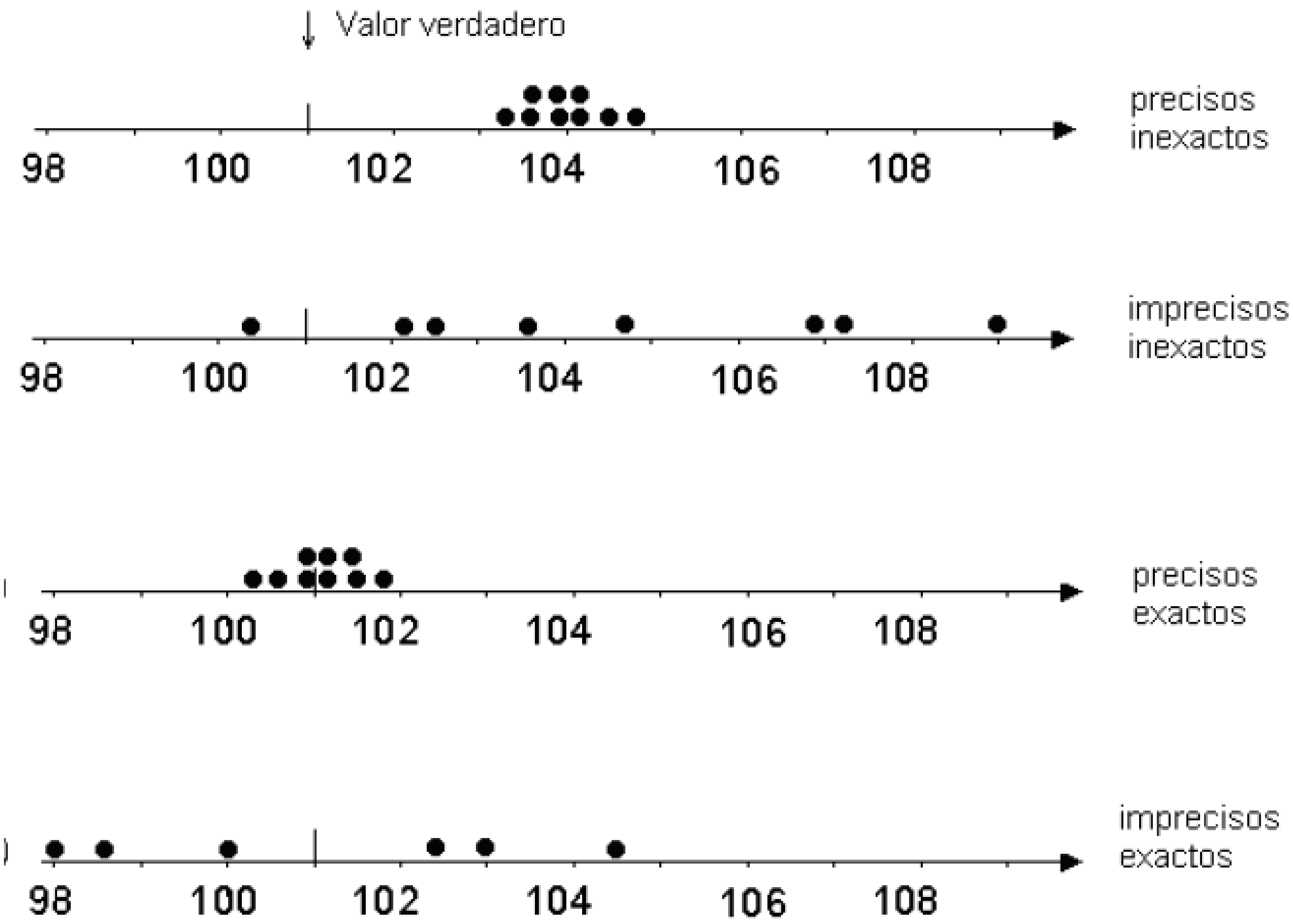
Ejemplo: un cronómetro es más preciso que un reloj común

Exactitud: La exactitud de un instrumento o método de medición está asociada a la calidad de la calibración del mismo, a la proximidad al valor verdadero.

Ejemplo: Imaginemos que el cronómetro que usamos es capaz de determinar la centésima de segundo, pero adelanta dos minutos por hora, mientras que un reloj de pulsera común no lo hace. En este caso decimos que el cronómetro es más preciso que el reloj común, pero menos exacto.

ERRORES

Errores en el Proceso de Medición: Precisión y Exactitud



ERRORES

Errores en el Proceso de Medición

Tenemos errores por diversos orígenes:

Error de apreciación (mínima division de escala)

Error de definición (falta de definición del objeto)

Error de interacción (interacción en el método de medición):

Ejemplo: Si usamos un termómetro para medir una temperatura, existe un intercambio de calor entre el termómetro y aquello a lo que queremos medirle la temperatura, de modo que el resultado de la medición es un valor modificado del original debido a la interacción. Esta interacción podrá o no ser significativa, de acuerdo a si medimos la temperatura de un metro cúbico de agua si el volumen en cuestión es una fracción del mililitro.

ERRORES

Errores en el Proceso de Medición

$$\text{MEDICION}= \mu+\underbrace{\varepsilon_1+\varepsilon_2+\varepsilon_3+\ldots+\varepsilon_n}_{\text{errores}}$$

Cantidad a medir: (desconocido, pero no aleatorio)

$$VAR(\mu+\varepsilon_1+\varepsilon_2+\varepsilon_3+\ldots+\varepsilon_n)=\sigma^2_1+\sigma^2_2+\sigma^2_3+\ldots+\sigma^2_n=\sigma^2$$

Si llamamos $\varepsilon=\varepsilon_1+\varepsilon_2+\varepsilon_3+\ldots+\varepsilon_n$, bajo ciertas condiciones podría ocurrir que:

$X=\mu+\varepsilon$ donde $\varepsilon \sim N(0, \sigma^2)$ esto es equivalente a $X \sim N(\mu, \sigma^2)$

MODELO

PROPAGACION DE INCERTIDUMBRE

$$X = \mu + \xi$$

si $Z = f(X)$ el método de propagación de la incertidumbre dice que es una buena aproximación tomar

$$V(Z) = [f'(\mu)]^2 V(\xi)$$

Para justificar el método basta tomar el polinomio de Taylor de orden 1 centrado en X, evaluarlo en μ para obtener que

$$f(X) = f(\mu) + f'(\mu)\xi + R$$

Si se desprecia el termino de error R y se utiliza las propiedades de la varianza.

$$V(f(X)) = V(f(\mu) + f'(\mu)\xi + R) \cong V(f'(\mu)\xi) = [f'(\mu)]^2 V(\xi)$$

Si tenemos más de una medición

$$(X_1, \ldots, X_n) = (\mu_1, \ldots, \mu_n) + (\xi_1, \ldots, \xi_n).$$

Si $Z = f(X_1, \ldots, X_n)$

$$V(Z) = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial f}{\partial \mu_i} \frac{\partial f}{\partial \mu_j} Cov(\xi_i, \xi_j)$$

donde las derivadas parciales están evaluadas en $(X_1, \ldots, X_n)$. Para el caso particular donde los errores no están correlacionados la ecuación anterior queda

$$V(Z) = \sum_{i=1}^{n} \frac{\partial f}{\partial \mu_i}^2 V(\xi_i)$$

INTERVALOS DE CONFIANZA

Cuando se obtiene una estimación puntual de un parámetro, es conveniente acompañar dicha estimación por una **medida** de la precisión de la estimación.

Un modo de hacerlo es informar el estimador y su error standard.

Otro modo es reemplazar la estimación puntual por un intervalo de valores posibles para el parámetro.

Ejemplo: Supongamos que tenemos una m.a. $X_1, X_2, \ldots, X_n$

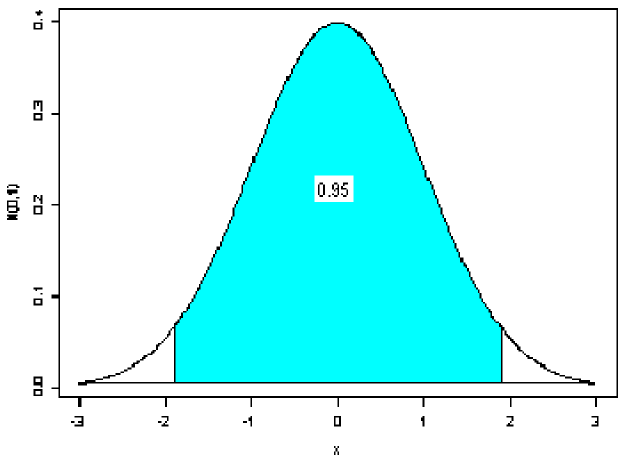
de una distribución $N(\mu, \sigma_o^2)$ con varianza σ_o^2 conocida.

Por ser los datos normales, sabemos que

$$\overline{X} \sim N\left(\mu, \frac{\sigma_o^2}{n}\right) \quad \Leftrightarrow \quad \frac{\overline{X} - \mu}{\frac{\sigma_o}{\sqrt{n}}} \sim N(0,1)$$

y, por lo tanto,

$$P\left(-1.96 \le \sqrt{n}\,\frac{\overline{X} - \mu}{\sigma_o} \le 1.96\right) = 0.95$$



A partir de esta expresión obtenemos

$$P\left(-1.96\frac{\sigma_o}{\sqrt{n}} \leq \overline{X} - \mu \leq 1.96\frac{\sigma_o}{\sqrt{n}}\right) = 0.95 \quad \Leftrightarrow \quad P\left(\overline{X} - 1.96\frac{\sigma_o}{\sqrt{n}} \leq \mu \leq \overline{X} + 1.96\frac{\sigma_o}{\sqrt{n}}\right) = 0.95$$

Es decir, que la probabilidad de que el intervalo

$$\left[\overline{X} - 1.96\frac{\sigma_o}{\sqrt{n}}, \overline{X} + 1.96\frac{\sigma_o}{\sqrt{n}}\right]$$

contenga al verdadero valor del parámetro μ es 0.95. Este intervalo se denomina **intervalo de confianza para μ de nivel de confianza 0.95**.

A partir de esta expresión obtenemos

$$P\left(-1.96\frac{\sigma_o}{\sqrt{n}} \leq \overline{X} - \mu \leq 1.96\frac{\sigma_o}{\sqrt{n}}\right) = 0.95 \quad \Leftrightarrow \quad P\left(\overline{X} - 1.96\frac{\sigma_o}{\sqrt{n}} \leq \mu \leq \overline{X} + 1.96\frac{\sigma_o}{\sqrt{n}}\right) = 0.95$$

Es decir, que la probabilidad de que el intervalo

$$\left[\overline{X} - 1.96\frac{\sigma_o}{\sqrt{n}}, \overline{X} + 1.96\frac{\sigma_o}{\sqrt{n}}\right]$$

contenga al verdadero valor del parámetro μ es 0.95. Este intervalo se denomina **intervalo de confianza para μ de nivel de confianza 0.95**.

En general, tendremos

$$P\left(-z_{\alpha/2} \leq \sqrt{n}\,\frac{\overline{X} - \mu}{\sigma_o} \leq z_{\alpha/2}\right) = 1 - \alpha$$

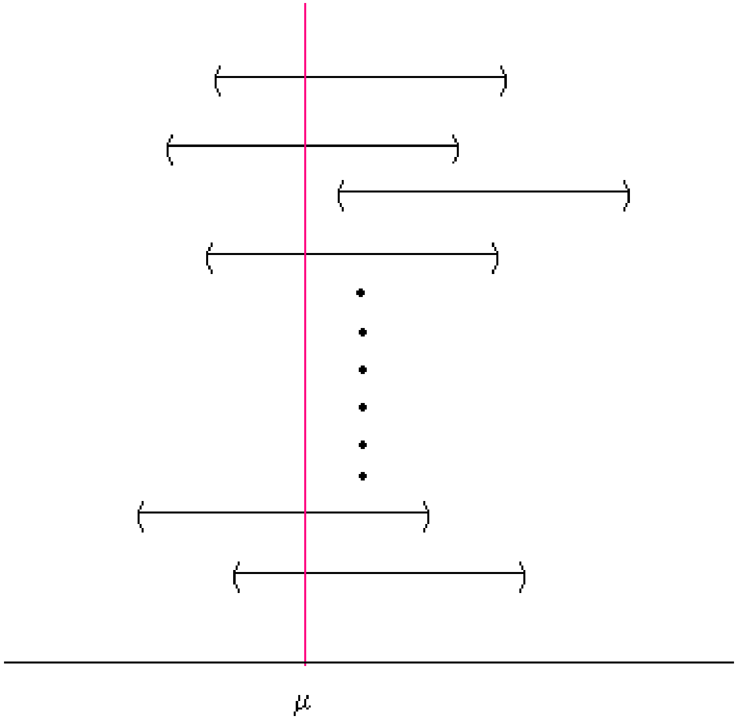
luego el siguiente **intervalo de confianza es de nivel 1 - α para μ**

$$\left[\overline{X} - z_{\alpha/2}\frac{\sigma_o}{\sqrt{n}}, \overline{X} + z_{\alpha/2}\frac{\sigma_o}{\sqrt{n}}\right]$$

Interpretación: Veamos una simulación en R

Supongamos que, en base a diferentes muestras calculamos los correspondientes intervalos de confianza para μ.

Entonces el $(1 - \alpha)$ 100% de ellos contendrán al verdadero valor μ.



Ejemplo:

Supongamos que tenemos una muestra normal con n=49 con verdadero valor del desvío standard es σ_o = 35 y que se observa $\bar{x} = 160$ y construimos un intervalo de confianza para la media de nivel 0.95.

De acuerdo al enunciado tenemos que $X_1, X_2, \ldots, X_{49}$ son v. a. i.i.d. N(μ, 35^2). Como las v.a. son normales independienes y la varianza es conocida, el intervalo para μ será de la forma

$$\left(\bar{X} - z_{a/2} \frac{\sigma_o}{\sqrt{n}}, \bar{X} + z_{\alpha/2} \frac{\sigma_o}{\sqrt{n}} \right)$$

Como, $z_{\alpha/2} = z_{0.025} = 1.96$ y $\sigma_o = 35, n = 49$ obtenemos

$$\left(160 - 1.96 \frac{35}{\sqrt{49}}, 160 + 1.96 \frac{35}{\sqrt{49}} \right) = (160 - 9.8, 160 + 9.8) = (150.2, 169.8)$$

INTERVALOS DE CONFIANZA PARA LOS PARAMETROS DE LA DISTRIBUCION NORMAL

Distribución t:

Sean dos v.a. $Z \sim N(0,1)$ y $U \sim \chi^2_n = \Gamma\left(\frac{n}{2}, \frac{1}{2}\right)$ independientes, entonces

$$T = \frac{Z}{\sqrt{U/n}} \sim t_n$$

Se dice que *T* tiene distribución ***t* de Student con *n* grados de libertad**.

Esta distribución está tabulada para diferentes valores de *n*. Su densidad es simétrica respecto al 0 y tiene forma de campana, pero tiene colas más pesadas que la distribución normal standard.

Cuando n tiende a infinito, la distribución de Student tiende a la distribución normal standard.

Proposición: Sea $X_1, X_2, \ldots, X_n$ una m.a. de una distribución N(μ, σ²), entonces

a) $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \quad \Leftrightarrow \quad \sqrt{n}\,\frac{\bar{X}-\mu}{\sigma} \sim N(0,1)$

b) $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$ con $S^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}$

c) $\bar{X}$ y S^2 son independientes

d) $\sqrt{n}\,\frac{\bar{X}-\mu}{S} \sim t_{n-1}$

Intervalo de confianza para la media de la distribución normal con varianza desconocida:

Sea $X_1, X_2, \ldots, X_n$ una m.a. de una distribución N(μ, σ^2), entonces

$$\sqrt{n}\,\frac{\overline{X}-\mu}{S} \sim t_{n-1}$$

$$P\left(-t_{n-1,\alpha/2} \leq \sqrt{n}\,\frac{\overline{X}-\mu}{S} \leq t_{n-1,\alpha/2}\right) = 1-\alpha$$

de donde se deduce el siguiente **intervalo de confianza de nivel 1 - α para μ,**

$$\left[\overline{X} - t_{n-1,\alpha/2}\frac{S}{\sqrt{n}}, \overline{X} + t_{n-1,\alpha/2}\frac{S}{\sqrt{n}}\right]$$

Intervalo de confianza para la varianza de la distribución normal con media conocida

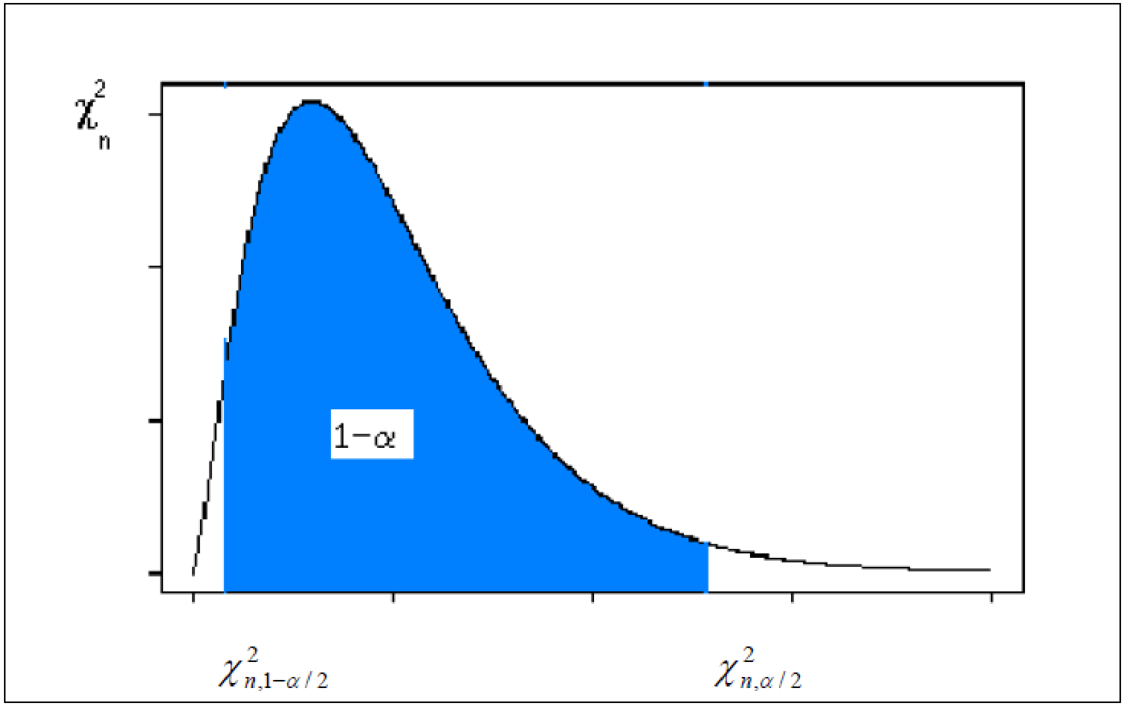
Sea $X_1, X_2, \ldots, X_n$ una m.a. de una distribución $N(\mu_o, \sigma^2)$, con media μ_o conocida, entonces

$$\frac{X_i - \mu_o}{\sigma} \sim N(0,1) \quad \forall 1 \leq i \leq n \quad \Rightarrow \quad \left(\frac{X_i - \mu_o}{\sigma}\right)^2 \sim \chi_1^2 = \Gamma\left(\frac{1}{2}, \frac{1}{2}\right) \quad \forall 1 \leq i \leq n$$

Como además las v.a. son independientes

$$\sum_{i=1}^{n} \left(\frac{X_i - \mu_o}{\sigma}\right)^2 \sim \chi_n^2 = \Gamma\left(\frac{n}{2}, \frac{1}{2}\right)$$

¿Cómo elegimos los percentiles de la distribución χ^2 que encierran un área igual a $1 - \alpha$?



Los elegimos de manera tal que quede un área igual a $\alpha/2$ en cada extremo.

Entonces,

$$P\left(\chi^2_{n,1-\alpha/2} \leq \frac{\sum_{i=1}^{n}(X_i-\mu_o)^2}{\sigma^2} \leq \chi^2_{n,\alpha/2}\right) = 1-\alpha$$

Se obtiene el siguiente intervalo

$$\left[\frac{\sum_{i=1}^{n}(X_i-\mu_o)^2}{\chi^2_{n,\alpha/2}}, \frac{\sum_{i=1}^{n}(X_i-\mu_o)^2}{\chi^2_{n,1-\alpha/2}}\right]$$

Ahora un intervalo para σ

Suponiendo como antes que observamos $\bar{x} = 160$ y $s = 35$, hallemos un intervalo de confianza para σ de nivel 0.95.

Por tratarse de una muestra normal con media desconocida, el intervalo para σ^2 será de la forma

$$\left[\frac{(n-1)S^2}{\chi^2_{n-1,\alpha/2}}, \frac{(n-1)S^2}{\chi^2_{n-1,1-\alpha/2}}\right]$$

con $\chi^2_{n-1,\alpha/2} = \chi^2_{48,0.025} = 69.02$ y $\chi_{n-1,1-\alpha/2} = \chi_{48,0.975} = 30.75$ Obtenemos

$$\left(\frac{48 \cdot 35^2}{69.02}, \frac{48 \cdot 35^2}{30.75}\right) = (851.93, 1912.20)$$

y tomando raíz cuadrada, un intervalo de confianza para σ de nivel 0.95 será

$$\left(\sqrt{\frac{48 \cdot 35^2}{69.02}}, \sqrt{\frac{48 \cdot 35^2}{30.75}}\right) = \left(\sqrt{851.93}, \sqrt{1912.20}\right) = (29.19, 43.73)$$

Volviendo al ejemplo:

Supongamos ahora que la varianza es desconocida, pero que el valor observado de S es s=35.

El correspondiente intervalo de confianza para μ será de la forma

$$\left(\overline{X} - t_{n-1,\alpha/2}\frac{S}{\sqrt{n}}, \overline{X} + t_{n-1,\alpha/2}\frac{S}{\sqrt{n}}\right)$$

con $t_{n-1,\alpha/2} = t_{48,0.025} = 2.01$. Obtenemos

$$\left(160 - 2.01\frac{35}{\sqrt{49}}, 160 + 2.01\frac{35}{\sqrt{49}}\right) = (160 - 10.05, 160 + 10.05) = (149.95, 170.05)$$

y como es lógico, resulta un intervalo más largo.

Ahora un intervalo para σ

Suponiendo como antes que observamos $\bar{x} = 160$ y $s = 35$, hallemos un intervalo de confianza para σ de nivel 0.95.

Por tratarse de una muestra normal con media desconocida, el intervalo para σ^2 será de la forma

$$\left(\frac{(n-1)S^2}{\chi^2_{n-1,\varepsilon/2}}, \frac{(n-1)S^2}{\chi^2_{n-1,1-\varepsilon/2}} \right)$$

con $\chi^2_{n-1,\alpha/2} = \chi^2_{48,0.025} = 69.02$ y $\chi_{n-1,1-\alpha/2} = \chi_{48,0.975} = 30.75$ Obtenemos

$$\left(\frac{48 \cdot 35^2}{69.02}, \frac{48 \cdot 35^2}{30.75} \right) = (851.93, 1912.20)$$

y tomando raíz cuadrada, un intervalo de confianza para σ de nivel 0.95 será

$$\left(\sqrt{\frac{48 \cdot 35^2}{69.02}}, \sqrt{\frac{48 \cdot 35^2}{30.75}} \right) = \left(\sqrt{851.93}, \sqrt{1912.20} \right) = (29.19, 43.73)$$