2.3 INTERVALOS DE CONFIANZA PARA LA VARIANZA

Para utilizar a S^2 como estimador de σ^2 necesitamos conocer su distribución de probabilidad. De esta manera podremos establecer un cierto coeficiente de confianza de la estimación.

No existe una distribución conocida para S^2 pero sí para cierta transformación del mismo. Si la muestra proviene de una población en la cual la variable en estudio se distribuye normalmente, tenemos:

$$(n-1)\frac{S^2}{\sigma^2} \sim \chi^2_{n-1}$$

donde χ^2_{n-1} es la distribución Chi-cuadrado con *n - 1* grados de libertad.

Una vez que contamos con esta información, podemos establecer un intervalo de confianza para estimar σ^2, de la siguiente manera:

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

Despejando convenientemente, obtenemos el siguiente intervalo:

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

Trataremos de aclarar estos conceptos mediante un ejemplo.

El Departamento de Control de Calidad de una empresa decide estudiar la homogeneidad de los productos comprados a un proveedor de materiales plásticos. Para ello toma 6 piezas de la producción de un envío y las somete a pruebas de resistencia.

Las siguientes observaciones representan las cargas de rotura en unidades de 1.000 libras por pulgada cuadrada:

Cargas de rotura
15,3
18,7
22,3
17,6
19,1
14,8

El Departamento de Calidad ha determinado que interesa conocer la variabilidad de las cargas de rotura, pues la uniformidad de este insumo es fundamental en el proceso de producción de la fábrica.

Como primera medida y habiendo comprobado que la variable cargas de rotura se distribuye normalmente, decide estimar la variabilidad poblacional del material plástico y para ello construye un intervalo de confianza. De acuerdo a los requerimientos del problema, debe contar con la siguiente información:

$$S^2 = 7{,}575 \quad (n - 1) = 5$$
$$X^2_{5;0,05} = 1{,}15 \quad X^2_{5;0,95} = 11{,}07$$

El encargado de control de calidad decide trabajar con una confianza del 90% por lo que el intervalo de confianza para la varianza poblacional del material plástico queda conformado de la siguiente manera:

$$P\left[\frac{5 \times 7{,}575}{11{,}07} \leq \sigma^2 \leq \frac{5 \times 7{,}575}{1{,}15}\right] = 0{,}90$$

$$P(3{,}42 \leq \sigma^2 \leq 32{,}93) = 0{,}90$$

Este intervalo también puede expresarse en términos de la desviación estándar, solamente tomando raíz cuadrada en cada uno de los límites.

Así:

$$P(1{,}85 < \sigma < 5{,}74) = 0{,}90$$

Podemos observar que de acuerdo al valor que toma la varianza la precisión del intervalo deja mucho que desear. Evidentemente, se ha tomado una muestra muy chica y el investigador industrial debería ampliar considerablemente su tamaño si desea obtener una estimación más precisa de la variabilidad en la población de piezas.

2.4 INTERVALOS PARA TOTALES

A partir de la estimación por intervalos realizada para la media, por ejemplo, en el caso de una variable con distribución normal proveniente de una población con varianza conocida:

$$\bar{X} \pm z \frac{\sigma}{\sqrt{n}}$$

Se puede obtener la estimación por intervalos de los totales, multiplicando cada uno de los límites por el tamaño de la muestra

$$n\left(\bar{X} - z \frac{\sigma}{\sqrt{n}}\right) < T < n\left(\bar{X} + z \frac{\sigma}{\sqrt{n}}\right)$$

De esta forma se concluye que con un 1-α % de confianza, el total de n observaciones de la variable x se encuentra entre ambos valores determinados.