Hay dos razones por las cuales el intervalo (6.63, 11.37) tiene mayor longitud que el obtenido anteriormente (7.69, 10.31).

- la varianza estimada 9.5, es mayor que el valor "conocido" 4 considerado previamente.
- aún cuando la varianza estimada hubiese sido 4, el intervalo estimado tendría mayor longitud debido a que el valor crítico de la t_8 es 2.306 mientras que el de la Normal es 1.96 para $1-\alpha = 0.95$. El intervalo hubiese sido (9-2.306*2/3, 9+2.306*2/3) = (7.46, 10.54)

OBSERVACIÓN

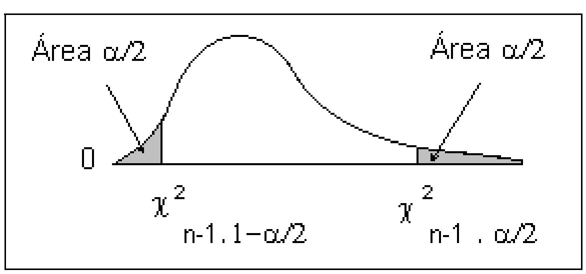
La longitud de un intervalo de confianza para μ no siempre es mayor cuando la varianza es desconocida ya que puede ocurrir que el desvío estándar s resulte mucho menor que σ. Sin embargo en promedio la longitud del intervalo es mayor cuando σ es desconocida: se puede demostrar que

$$t_{\alpha,n-1} E(S) \geq z_{\alpha} \ \sigma$$

2. INTERVALOS DE CONFIANZA PARA LA VARIANZA DE UNA DISTRIBUCIÓN NORMAL

Si X1, ... ,Xn es una muestra aleatoria de una distribución Normal de parámetros μ y σ^2, entonces podemos construir un intervalo de confianza para σ^2 utilizando el hecho que

$$(n-1)\frac{S^2}{\sigma^2} \sim \chi^2_{n-1}$$ (ji-cuadrado con n-1 grados de libertad)



Luego

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

ó equivalentemente

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

Por lo tanto, cuando $S^2 = s^2$, un intervalo del 100*(1-α)% de confianza para σ^2 está dado por

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

Ejemplo. Se espera que un procedimiento estandarizado produzca arandelas con muy pequeña desviación en su espesor. Suponga que se eligen al azar 10 de tales arandelas y se mide su espesor obteniéndose en pulgadas: 0.123, 0.124, 0.126, 0.120, 0.130, 0.133, 0.125, 0.128, 0.124, 0.126. Interesa calcular un intervalo del 90% de confianza para el desvío del grosor de las arandelas producidas por este procedimiento.

Solución.

s^2=1.366 x 10^{-6}

$$\chi^2_{9,0.05} = 16.917; \quad \chi^2_{9,0.95} = 3.334$$

$$\frac{9 \text{ x } 1.366 \text{ x } 10^{-5}}{16.917} = 7.267 \text{ x } 10^{-6}; \quad \frac{9 \text{ x } 1.366 \text{ x } 10^{-5}}{3.334} = 36.875 \text{ x } 10^{-6}$$

Luego, con una confianza del 90%

$$\sigma^2 \in \left(7.267 \text{ x } 10^{-6};\ 36.875 \text{ x } 10^{-6}\right)$$

Tomando raíz cuadrada, con una confianza del 90%,

$$\sigma \in \left(2.686 \text{ x } 10^{-3};\ 6.072 \text{ x } 10^{-3}\right)$$

3.1 INTERVALOS CON NIVEL DE CONFIANZA APROXIMADOS PARA LA MEDIA DE UNA VARIABLE BINOMIAL. Muestras Grandes.

Consideremos una población de artículos que pueden o cumplir con ciertas normas en proporciones p y 1-p, desconocidas. Si elegimos una muestra de n artículos al azar y registramos

$$X_i = \begin{cases} 1 & \text{si el artículo cumple con las normas} \\ 0 & \text{si el artículo no cumple con las normas} \end{cases}$$

Entonces $X = \sum_{i=1}^{n} X_i$ es la cantidad de artículos de la muestra que cumplen con las normas. Si podemos suponer que cada artículo cumple o no con las normas en forma independiente, resulta que $X \sim Bi(n,p)$ siendo p la proporción de artículos en la población que cumplen con las normas.

Para construir un intervalo de confianza para p nos basaremos en la aproximación de la distribución Binomial por la distribución Normal cuando n es suficientemente grande.

$$\frac{X - np}{\sqrt{np(1-p)}} \quad \underset{aprox}{\sim} \quad N(0,1)$$

Por lo tanto, para cualquier α en el intervalo (0,1)

$$P\left\{ -z_{\alpha/2} < \frac{X - np}{\sqrt{np(1-p)}} < z_{\alpha/2} \right\} \approx 1 - \alpha$$

ó equivalentemente

$$P\left\{ -z_{\alpha/2} < \frac{\hat{p} - p}{\sqrt{p(1-p)/n}} < z_{\alpha/2} \right\} \approx 1 - \alpha \qquad (4)$$

donde $\hat{p} = \frac{X}{n}$.

Para despejar un intervalo de confianza para p de la expresión (4) se reemplazan los menores (<) por iguales (=). Se obtiene así una ecuación cuadrática en p cuyas soluciones son los extremos del intervalo de confianza buscado

$$p = \frac{\hat{p} + \frac{z_{\alpha/2}^2}{2n} \pm z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n} + \frac{z_{\alpha/2}^2}{4n^2}}}{1 + (z_{\alpha/2}^2)/n}$$

Como n es grande los términos en z^2 son despreciables y el intervalo resultante es el mismo que se obtiene al reemplazar p por $\hat{p}$ en el denominador de la expresión (4).

Luego un intervalo de confianza con nivel de confianza aproximado 1-α para p está dado por

$$\left\{\hat{p} - z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}},\ \hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right\} \qquad (5)$$

Este intervalo puede utilizarse siempre que $n\hat{p} \geq 5$ y $n(1-\hat{p}) \geq 5$.

Ejemplo. Se elige al azar, de un lote grande, una muestra de 100 transistores. Mediante una prueba, se determina que 80 de ellos satisfacen las normas vigentes. Un intervalo de confianza del 95% para p, la verdadera proporción de transistores que cumplen con los requerimientos, está dado por

$$\left(0.8 - 1.96\sqrt{0.8(0.2)/100};\ 0.8 + 1.96\sqrt{0.8(0.2)/100}\right) = (0.7216;\ 0.8784)$$

Esto es, con un aprox. 95% de confianza, entre 72.16% y 87.84% de los transistores cumplen con los requerimientos.

3.2 PROCEDIMIENTO EN DOS PASOS: TAMAÑO DE MUESTRA NECESARIO PARA LA OBTENCIÓN DE UN INTERVALO DE CONFIANZA CON LONGITUD PREFIJADA PARA UNA PROPORCIÓN

La longitud del intervalo de confianza dado por la ecuación (5)

$$2\, z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

depende del parámetro que nos interesa estimar. Para hallar un tamaño de muestra de manera que la longitud del intervalo resultante sea aproximadamente L se procede en dos pasos:

- Paso 1: se toma una muestra inicial de tamaño n1 y se obtiene un estimador inicial $\tilde{p}$
- Paso 2: se utiliza la proporción estimada en el paso 1 para determinar el tamaño total n resolviendo la siguiente ecuación:

$$2\, z_{\alpha/2} \sqrt{\frac{\tilde{p}(1-\tilde{p})}{n}} = L$$

elevando ambos miembros de la igualdad al cuadrado

$$\left(2\, z_{\alpha/2}\right)^2 \tilde{p}(1-\tilde{p})/n = L^2$$

ó

$$n = \left(2\, z_{\alpha/2}\right)^2 \tilde{p}(1-\tilde{p})/L^2$$

Ejemplo. Un laboratorio introduce un procedimiento, que le resulta más económico, para la obtención de un reactivo que luego envasa en frascos. La probabilidad que el reactivo del un frasco elegido al azar cumpla las normas de calidad es desconocida (p). Interesa obtener un intervalo de 99% de confianza cuya longitud sea aproximadamente 0.05.

- Paso 1. En una muestra inicial de 30 frascos, 26 de ellos resultaron aceptables por lo que el estimador inicial es $\tilde{p}$ =26/30 = 0.87.
- Paso 2. El tamaño muestral requerido es

$$n = \frac{\left(2\, z_{0.005}\right)^2}{(0.05)^2}(26/30)(1-26/30) = \frac{4(2.58)^2}{(0.05)^2}\frac{26}{30}\frac{4}{30} = 1231$$

Deberíamos tomar una muestra adicional de 1201 frascos. Si, por ejemplo, 1040 de ellos resultan aceptables (manteniendo aprox. la proporción inicial) el intervalo de 99% de confianza para la verdadera proporción de componentes aceptables es:

$$\left(\frac{1066}{1231} - \sqrt{1066\left(1-\frac{1066}{1231}\right)}\frac{z_{0.005}}{1231};\ \frac{1066}{1231} + \sqrt{1066\left(1-\frac{1066}{1231}\right)}\frac{z_{0.005}}{1231}\right)$$

(0.84091; 0.89101)

OBSERVACIÓN

Hemos visto que la longitud del intervalo de confianza para p es L si n está dado por

$$n = \left(2\, z_{\alpha/2}\right)^2 \tilde{p}(1-\tilde{p})/L^2$$

Es fácil ver que la función g(p) = p(1-p) definida en el intervalo [0,1] toma su valor máximo 1/4 cuando p = 1/2. Luego una cota superior para n es:

$$n \le \left(z_{\alpha/2}\right)^2 / L^2$$

De esta manera, si se elige una muestra de tamaño mayor o igual a $(z_{\alpha/2})^2/L^2$, garantizamos la obtención de un intervalo de confianza para *p* de longitud no mayor a L sin tener que realizar un muestreo adicional.

4. INTERVALOS DE CONFIANZA PARA LA MEDIA DE UNA DISTRIBUCIÓN EXPONENCIAL

La distribución exponencial es utilizada en estudios sobre confiabilidad como modelo del tiempo hasta la falla de un dispositivo. Por ejemplo el tiempo de vida de un componente semiconductor podría estar modelado como una variable aleatoria con media 40 000 horas. Recíprocamente, suponiendo que un modelo exponencial es el adecuado para modelar el tiempo de vida de un componente, podríamos estar interesados en estimar su vida media mediante un intervalo de confianza.

Supongamos que X1, ... ,Xn es una muestra aleatoria de una distribución exponencial de parámetro λ , Xi ~ $\varepsilon(\lambda)$. Sabemos que E(Xi)=1/λ, luego la media muestral $\sum$Xi / n es un estimador insesgado y consistente de 1/λ. Para obtener un intervalo de confianza para 1/λ es necesario recordar que

$$2\lambda \sum_{i=1}^{n} X_i \sim \chi^2_{2n}$$

Luego, para cualquier $\alpha \in (0,1)$

$$P\left\{ \chi^2_{2n,1-\alpha/2} \le 2\lambda \sum_{i=1}^{n} X_i \le \chi^2_{2n,\alpha/2} \right\} = 1-\alpha$$

ó equivalentemente,

$$P\left\{ \frac{2\sum_{i=1}^{n} X_i}{\chi^2_{2n,\alpha/2}} \le \frac{1}{\lambda} \le \frac{2\sum_{i=1}^{n} X_i}{\chi^2_{2n,1-\alpha/2}} \right\} = 1-\alpha$$

Luego, un intervalo de 100(1-α)% de confianza para 1/λ es

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

Ejemplo. Una fábrica produce artículos cuyos tiempos de vida (en horas) se suponen independientes con función de densidad exponencial común a todos:

$$f(x) = \lambda\, e^{-\lambda x}, \quad 0 < x < \infty$$

Si la suma los tiempos de vida de los primeros 10 artículos es 1 740 ¿cuál es un intervalo de confianza del 95% para 1/λ?

Como

$$\chi^2_{20,0.025=34.169} \qquad \chi^2_{20,0.975=9.661},$$

luego el intervalo es

$$\left(\frac{2 \text{ x } 1740}{34.169}, \frac{2 \text{ x } 1740}{9.661} \right)$$

Es decir que con un nivel de confianza del 95% el tiempo de vida medio se encuentra en el intervalo (101.847, 360.211).