Estadística. Grupo m3

Hoja 4. Intervalos de confianza

Mayte Rodríguez

Método de la cantidad pivotal

- Si μ es un parámetro de localización (la distribución de $X - \mu$ no depende de μ), entonces

$$T(\mathbf{X}, \mu) = T((X_1, \ldots, X_n), \mu) = \bar{X} - \mu$$

es una cantidad pivotal.

- Si σ es un parámetro de escala (la distribución de $\frac{X}{\sigma}$ no depende de σ) , entonces

$$T(\mathbf{X}, \sigma) = \frac{\bar{X}}{\sigma}$$

es una cantidad pivotal.

- Si μ es un parámetro de localización y σ es un parámetro de escala (la distribución de $\frac{\bar{X}-\mu}{\sigma}$ no depende de (μ, σ)), entonces

$$\begin{aligned} T_1(\mathbf{X}, (\mu, \sigma)) &= \frac{\bar{X} - \mu}{\sigma} \\ T_2(\mathbf{X}, \mu) &= \frac{\bar{X} - \mu}{S} \end{aligned}$$

Método de la cantidad pivotal

- Estadístico suficiente.
- La cantidad

$$T(\mathbf{X},\theta) = -2\sum_{j=1}^{n} \ln F_\theta(X_j) \sim \chi^2_{2n}$$

es una cantidad pivotal siempre que F_θ sea continua y monótona en θ.

Ejercicio 1

Sea $(X_1, X_2, ..., X_n)$ una muestra aleatoria simple de $X \sim N(\theta, \sigma^2 = \theta^2)$ para $\theta > 0$.
Construir una cantidad pivotal y utilizarla para hallar un intervalo de cofianza para θ al nivel de confianza $1 - \alpha$.

Como $\bar{X} \sim N(\theta, \theta^2/n)$ entonces

$$T(\mathbf{X}, \theta) = \frac{\bar{X} - \theta}{\theta/\sqrt{n}} \sim N(0, 1)$$

es una cantidad pivotal.
Al nivel de confianza $1 - \alpha$, hay que encontrar a y b de forma que

$$1 - \alpha = P(a < T(\mathbf{X}, \theta) < b).$$

Ejercicio 1

Como la distribución de $T(\mathbf{X}, \theta)$ es simétrica, el intervalo de longitud mínima es el que tiene probabilidad de colas iguales. Entonces, la probabilidad de cada cola será $\alpha/2$ de forma que

$$\begin{aligned} P(T(\mathbf{X}, \theta) &\geq b) = \alpha/2 \\ P(T(\mathbf{X}, \theta) &\leq a) = \alpha/2 \end{aligned}$$

y entonces $a = -z_{\alpha/2}$ y $b = z_{\alpha/2}$. El intervalo de confianza para θ se obtiene al escribir la desigualdad en términos de θ

$$\begin{aligned} a &< T(\mathbf{X}, \theta) < b \\ -z_{\alpha/2} &< T(\mathbf{X}, \theta) < z_{\alpha/2} \\ -z_{\alpha/2} &< \frac{\bar{X} - \theta}{\theta/\sqrt{n}} < z_{\alpha/2} \\ -z_{\alpha/2}\theta/\sqrt{n} &< \bar{X} - \theta < z_{\alpha/2}\theta/\sqrt{n} \end{aligned}$$

Ejercicio 1

Entonces

$$\begin{aligned} -z_{\alpha/2}\theta/\sqrt{n} &< \bar{X}-\theta \\ -z_{\alpha/2}\theta/\sqrt{n}+\theta &< \bar{X} \\ \theta(1-z_{\alpha/2}/\sqrt{n}) &< \bar{X} \end{aligned}$$

Es decir

$$\theta < \bar{X}/(1-z_{\alpha/2}/\sqrt{n})$$

Ejercicio 1

Por otra parte

$$\begin{aligned} \bar{X} - \theta &< z_{\alpha/2}\theta/\sqrt{n} \\ \bar{X} &< z_{\alpha/2}\theta/\sqrt{n} + \theta \\ \bar{X} &< \theta(1 + z_{\alpha/2}/\sqrt{n}) \end{aligned}$$

y entonces

$$\bar{X}/(1 + z_{\alpha/2}/\sqrt{n}) < \theta$$

Ejercicio 1

Por tanto

$$\bar{X}/(1+z_{\alpha/2}/\sqrt{n}) < \theta < \bar{X}/(1-z_{\alpha/2}/\sqrt{n})$$

y el intervalo de confianza al nivel $1-\alpha$ para θ es

$$IC_{1-\alpha}(\theta) = \left(\frac{\bar{X}}{1+z_{\alpha/2}/\sqrt{n}}, \frac{\bar{X}}{1-z_{\alpha/2}/\sqrt{n}}\right)$$

Ejercicio 2

Sean (X_1, X_2) una muestra aleatoria simple de $X \sim N(0, \sigma^2 = 1/\theta)$ y $T = \frac{X_1^2 + X_2^2}{2}$ un estadístico. Demostrar que $2T\theta$ es una cantidad pivotal y utilizarla para construir un intervalo de confianza para θ al nivel de confianza $1 - \alpha$.

Se sabe que $\frac{X_1}{1/\sqrt{\theta}} = \sqrt{\theta} X_1 \sim N(0, 1)$ y $\sqrt{\theta} X_2 \sim N(0, 1)$, de forma que
$2T\theta = \theta X_1^2 + \theta X_2^2 \sim \chi_2^2$.
Al nivel de confianza $1 - \alpha$, hay que encontrar a y b de forma que

$$1 - \alpha = P(a < 2T\theta < b).$$

Ejercicio 2

Si calculamos el intervalo que tiene probabilidad de colas iguales, la probabilidad de cada cola será $\alpha/2$ de forma que

$$\begin{aligned} P(2T\theta &\geq b) = \alpha/2 \\ P(2T\theta &\leq a) = \alpha/2 \end{aligned}$$

y entonces $a = \chi^2_{2;1-\alpha/2}$ y $b = \chi^2_{2;\alpha/2}$.

Ejercicio 2

El intervalo de confianza para θ se obtiene al escribir la desigualdad en términos de θ

$$\begin{aligned}
\chi^2_{2;1-\alpha/2} &< 2T\theta < \chi^2_{2;\alpha/2} \\
\chi^2_{2;1-\alpha/2}/(2T) &< \theta < \chi^2_{2;\alpha/2}/(2T) \\
\chi^2_{2;1-\alpha/2}/(X_1^2 + X_2^2) &< \theta < \chi^2_{2;\alpha/2}/(X_1^2 + X_2^2)
\end{aligned}$$

El intervalo de confianza al nivel $1-\alpha$ para θ es

$$IC_{1-\alpha}(\theta) = \left(\frac{\chi^2_{2;1-\alpha/2}}{X_1^2 + X_2^2}, \frac{\chi^2_{2;\alpha/2}}{X_1^2 + X_2^2} \right)$$

Ejercicio 3a

Se han medido los siguientes valores de una determinada magnitud: 521, 742, 593, 635, 788, 717, 606, 639, 666, 624. Suponiendo que esta magnitud se distribuye según una normal y para un nivel de confianza del 95%, se pide:

a) Hallar un intervalo de confianza para la media.
Si $X \sim N(\mu, \sigma^2)$, una cantidad pivotal para μ es $T_1(\mathbf{X}, \mu) = \frac{\bar{X}-\mu}{S}$. Para calcular su distribución utilizamos el lema de Fisher, de forma que

$$\begin{aligned} \bar{X} &\sim N(\mu, \sigma^2/n) \\ \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} &\sim N(0,1) \end{aligned}$$

y esta última cantidad depende de σ, que es desconocida.

Ejercicio 3a

Por otra parte

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

y entonces

$$\frac{\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}}{\sqrt{(n-1)\frac{S^2}{\sigma^2}/(n-1)}} = \frac{\bar{X}-\mu}{S/\sqrt{n}} = T(\mathbf{X},\mu) \sim t_{n-1}.$$

Al nivel de confianza $1-\alpha$, hay que encontrar a y b de forma que

$$1-\alpha = P(a < T(\mathbf{X},\mu) < b).$$

Ejercicio 3a

Como la distribución de $T(\mathbf{X}, \mu)$ es simétrica, el intervalo de longitud mínima es el que tiene probabilidad de colas iguales. Entonces, la probabilidad de cada cola será $\alpha/2$ de forma que

$$
\begin{aligned}
P(T(\mathbf{X}, \mu) &\geq b) = \alpha/2 \\
P(T(\mathbf{X}, \mu) &\leq a) = \alpha/2
\end{aligned}
$$

y entonces $a = -t_{n-1;\alpha/2}$ y $b = t_{n-1;\alpha/2}$.

Ejercicio 3a

El intervalo de confianza para μ se obtiene al escribir la desigualdad en términos de μ

$$
\begin{aligned}
a &< T(\mathbf{X},\mu) < b \\
-t_{n-1;\alpha/2} &< T(\mathbf{X},\mu) < t_{n-1;\alpha/2} \\
-t_{n-1;\alpha/2} &< \frac{\bar{X}-\mu}{S/\sqrt{n}} < t_{n-1;\alpha/2} \\
-t_{n-1;\alpha/2}S/\sqrt{n} &< \bar{X}-\mu < t_{n-1;\alpha/2}S/\sqrt{n} \\
\bar{X}-t_{n-1;\alpha/2}S/\sqrt{n} &< \mu < \bar{X}+t_{n-1;\alpha/2}S/\sqrt{n}
\end{aligned}
$$

Ejercicio 3a

El intervalo de confianza al nivel $1-\alpha$ para μ es

$$IC_{1-\alpha}(\mu) = \left(\bar{X} - t_{n-1;\alpha/2} S/\sqrt{n}, \bar{X} + t_{n-1;\alpha/2} S/\sqrt{n}\right)$$

Para los datos observados, $\bar{x} = 653.1$, $s = 78.18$ y el intervalo de confianza al nivel $1-\alpha = 0.95$ es

$$IC_{1-\alpha}(\mu) = \left(653.1 - 2.262\frac{78.18}{\sqrt{10}}, 653.1 + 2.262\frac{78.18}{\sqrt{10}}\right) = (597, 709)$$

donde $t_{n-1;\alpha/2} = t_{9;0.025} = 2.262.$

b) Hallar un intervalo de confianza para la varianza.

Utilizaremos como cantidad pivotal $T(\mathbf{X}, \sigma^2) = (n-1)\frac{S^2}{\sigma^2} \sim \chi^2_{n-1}$. De esta forma, al nivel de confianza $1-\alpha$, hay que encontrar a y b de forma que

$$1 - \alpha = P(a < T(\mathbf{X}, \sigma^2) < b).$$

Tomando probabilidad de colas iguales $a = \chi^2_{n-1;1-\alpha/2}$ y $b = \chi^2_{n-1;\alpha/2}$.

Ejercicio 3b

El intervalo de confianza para σ^2 se obtiene al escribir la desigualdad en términos de σ^2

$$\begin{aligned}
\chi^2_{n-1;1-\alpha/2} &< T(\mathbf{X}, \sigma^2) < \chi^2_{n-1;\alpha/2} \\
\chi^2_{n-1;1-\alpha/2} &< (n-1)\frac{S^2}{\sigma^2} < \chi^2_{n-1;\alpha/2} \\
(n-1)S^2/\chi^2_{n-1;\alpha/2} &< \sigma^2 < (n-1)S^2/\chi^2_{n-1;1-\alpha/2}
\end{aligned}$$

Ejercicio 3b

El intervalo de confianza al nivel $1-\alpha$ para σ^2 es

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

Para los datos observados, $s^2 = 78.18^2$ y el intervalo de confianza es

$$IC_{1-\alpha}(\sigma^2) = \left(\frac{9 * 78.18^2}{19.02}, \frac{9 * 78.18^2}{2.7} \right) = (2892, 20374).$$

donde $\chi^2_{n-1;\alpha/2} = \chi^2_{9;0.025} = 19.02$ y $\chi^2_{n-1;1-\alpha/2} = \chi^2_{9;0.975} = 2.7$

Ejercicio 4b

El número diario de piezas fabricadas por una máquina A en 5 días ha sido: 50, 48, 53, 60, 37; mientras que, en esos mismos días, una máquina B ha fabricado: 40, 51, 62, 55, 64.
Suponiendo independencia, se pide:
b) Construir un intervalo de confianza al nivel de confianza $1-\alpha = 0.90$ para el cociente de las varianzas.
Sabemos que $X \sim N(\mu_1, \sigma_1^2)$ y $Y \sim N(\mu_2, \sigma_2^2)$.
Para construir el intervalo de confianza para el cociente de varianzas $\frac{\sigma_1^2}{\sigma_2^2}$, observamos que

$$\begin{aligned} W_1 &= \frac{(n-1)S_1^2}{\sigma_1^2} \sim \chi^2_{n-1} \\ W_2 &= \frac{(n-1)S_2^2}{\sigma_2^2} \sim \chi^2_{n-1} \end{aligned}$$

y entonces

$$T\left((\mathbf{X},\mathbf{Y}),\frac{\sigma_1^2}{\sigma_2^2}\right)=\frac{W_1/(n-1)}{W_2/(n-1)}=\frac{S_1^2}{S_2^2}\frac{\sigma_2^2}{\sigma_1^2}\sim F_{n-1,n-1}$$

es una cantidad pivotal para $\frac{\sigma_1^2}{\sigma_2^2}$. De esta forma, al nivel de confianza $1-\alpha$, hay que encontrar a y b de forma que

$$1-\alpha=P\left(a<T\left((\mathbf{X},\mathbf{Y}),\frac{\sigma_1^2}{\sigma_2^2}\right)<b\right).$$

Ejercicio 4b

Tomando probabilidad de colas iguales $a = F_{n-1,n-1;1-\alpha/2}$ y $b = F_{n-1,n-1;\alpha/2}$.

El intervalo de confianza para $\frac{\sigma_1^2}{\sigma_2^2}$ se obtiene al escribir la desigualdad en términos de $\frac{\sigma_1^2}{\sigma_2^2}$

$$
\begin{aligned}
F_{n-1,n-1;1-\alpha/2} &< T\left((\mathbf{X},\mathbf{Y}),\frac{\sigma_1^2}{\sigma_2^2}\right) < F_{n-1,n-1;\alpha/2} \\
F_{n-1,n-1;1-\alpha/2} &< \frac{S_1^2}{S_2^2}\frac{\sigma_2^2}{\sigma_1^2} < F_{n-1,n-1;\alpha/2} \\
\frac{S_1^2}{S_2^2}/F_{n-1,n-1;\alpha/2} &< \frac{\sigma_1^2}{\sigma_2^2} < \frac{S_1^2}{S_2^2}/F_{n-1,n-1;1-\alpha/2}
\end{aligned}
$$

Ejercicio 4b

El intervalo de confianza al nivel $1-\alpha$ para $\frac{\sigma_1^2}{\sigma_2^2}$ es

$$IC_{1-\alpha}\left(\frac{\sigma_1^2}{\sigma_2^2}\right) = \left(\frac{\frac{S_1^2}{S_2^2}}{F_{n-1,n-1;\alpha/2}}, \frac{\frac{S_1^2}{S_2^2}}{F_{n-1,n-1;1-\alpha/2}}\right).$$

Las cuasivarianzas muestrales son $s_1^2 = 70.3$ y $s_2^2 = 92.3$. El intervalo de confianza al nivel $1-\alpha = 0.9$ es

$$IC_{1-\alpha}\left(\frac{\sigma_1^2}{\sigma_2^2}\right) = \left(\frac{70.3/92.3}{6.39}, \frac{70.3/92.3}{0.156}\right) = (0.119, 4.882)$$

donde $F_{n-1,n-1;\alpha/2} = F_{4,4;0.05} = 6.39$ y
$F_{n-1,n-1;1-\alpha/2} = F_{4,4;0.95} = \frac{1}{F_{4,4;0.05}} = \frac{1}{6.39} = 0.156.$

Ejercicio 4a

a) Construir un intervalo de confianza al nivel de confianza $1 - \alpha = 0.95$ para la diferencia de medias.
Podemos asumir que $X \sim N(\mu_1, \sigma^2)$ y $Y \sim N(\mu_2, \sigma^2)$ y entonces

$$\begin{aligned} \bar{X} &\sim N(\mu_1, \sigma^2/n) \\ \bar{Y} &\sim N(\mu_2, \sigma^2/n) \\ \bar{X} - \bar{Y} &\sim N(\mu_1 - \mu_2, 2\sigma^2/n) \end{aligned}$$

Ejercicio 4a

Por tanto,

$$\frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sigma\sqrt{2/n}} \sim N(0,1)$$

.

Pero esta cantidad depende de σ, que es desconocida.

Por el lema de Fisher

$$\begin{aligned} \frac{(n-1)S_1^2}{\sigma^2} &\sim \chi^2_{n-1} \\ \frac{(n-1)S_2^2}{\sigma^2} &\sim \chi^2_{n-1} \end{aligned}$$

Ejercicio 4a

Entonces

$$\frac{(n-1)S_1^2}{\sigma^2}+\frac{(n-1)S_2^2}{\sigma^2}=\frac{(n-1)S_1^2+(n-1)S_2^2}{\sigma^2}\sim\chi^2_{2n-2}$$

Por tanto

$$T((\mathbf{X},\mathbf{Y}),\mu_1-\mu_2)=\frac{\frac{\bar{X}_1-\bar{X}_2-(\mu_1-\mu_2)}{\sigma\sqrt{2/n}}}{\sqrt{\frac{(n-1)S_1^2+(n-1)S_2^2}{\sigma^2}/(2n-2)}}=\frac{(\bar{X}_1-\bar{X}_2-(\mu_1-\mu_2))}{S_p\sqrt{2/n}}\sim t_{2n-2}$$

donde

$$S_p^2=\frac{(n-1)S_1^2+(n-1)S_2^2}{2n-2}$$

Ejercicio 4a

Entonces

$$\begin{aligned}
a &< T((\mathbf{X},\mathbf{Y}),\mu_1-\mu_2) < b \\
-t_{2n-2;\alpha/2} &< T((\mathbf{X},\mathbf{Y}),\mu_1-\mu_2) < t_{2n-2;\alpha/2} \\
-t_{2n-2;\alpha/2} &< \frac{(\bar{X}-\bar{Y}-(\mu_1-\mu_2))}{S_p\sqrt{2/n}} < t_{2n-2;\alpha/2} \\
-t_{2n-2;\alpha/2}S_p\sqrt{2/n} &< (\bar{X}-\bar{Y}-(\mu_1-\mu_2)) < t_{2n-2;\alpha/2}S_p\sqrt{2/n} \\
\bar{X}-\bar{Y}-t_{2n-2;\alpha/2}S_p\sqrt{2/n} &< \mu_1-\mu_2 < \bar{X}-\bar{Y}+t_{2n-2;\alpha/2}S_p\sqrt{2/n}
\end{aligned}$$

Ejercicio 4a

y el intervalo de confianza al nivel $1-\alpha$ para $\mu_1 - \mu_2$ es

$$IC_{1-\alpha}(\mu_1-\mu_2) = (\bar{X}-\bar{Y}-t_{2n-2;\alpha/2}S_p\sqrt{2/n}, \bar{X}-\bar{Y}+t_{2n-2;\alpha/2}S_p\sqrt{2/n}).$$

Las medias y las cuasivarianzas muestrales son $\bar{x} = 49.6$, $\bar{y} = 54.4$, $s_1^2 = 70.3$ y $s_2^2 = 92.3$.
El intervalo de confianza al nivel $1-\alpha = 0.95$ es

$$IC_{1-\alpha}(\mu_1-\mu_2) = (49.6-54.4-2.306*9.0166\sqrt{2/5}, 49.6-54.4+2.306*9.0166\sqrt{2/5})$$

Ejercicio 4a

$$IC_{1-\alpha}(\mu_1 - \mu_2) = (-17.9502, 8.3502)$$

donde

$$s_p = \sqrt{\frac{(n-1)S_1^2 + (n-1)S_2^2}{2n-2}} = \sqrt{\frac{4*70.3 + 4*92.3}{8}} = 9.0166$$

y $t_{2n-2;\alpha/2} = t_{8;0.025} = 2.306$