Introducción

- En el tema de estimación puntual vimos como p.ej. el estimador de máxima verosimilitud utiliza los datos para proporcionar una estimación $\hat{\theta}$ del valor de θ.
- Por sí sola, una estimación puntual como $\hat{\theta} = 5.13$ no contiene información sobre su precisión, es solamente un número, sea que se base en 10 datos o en 10000.
- Por este motivo las estimaciones puntuales se complementan con **intervalos de confianza**.

Introducción

- Por ejemplo para estimar un valor esperado desconocido μ podríamos decir que nuestra estimación de μ es $\overline{X} = 5.13$ con un intervalo de confianza del 90% de $[4.80, 5.46]$.
- Otra forma de describir dicho intervalo es $\overline{X} \pm 0.33 = 5.13 \pm 0.33$.
- Cuando queremos estimar un parámetro desconocido θ, la estimación por un intervalo de confianza (IC) consiste en dar un rango de valores creíbles de θ en el cual confiemos al nivel $\alpha = (1-\alpha)\%$. La precisión dada por el intervalo es el ancho del intervalo.
- La estimación mediante IC no tiene porqué contener al verdadero valor de θ, la probabilidad de que sí lo contenga es $1-\alpha$.

Definiciones y comentarios

- Supongamos una variable aleatoria X con distribución que depende de un parámetro θ (θ podría ser un vector). Sea $M = \{x_1, \ldots, x_n\}$ una muestra de n realizaciones u observaciones de X.
- Dado $\alpha \in (0, 1)$ un **intervalo de confianza** al nivel $1 - \alpha$ para θ es un intervalo $[a, b]$ donde a y b dependen de la muestra y tal que $P(\theta \in [a, b]) = 1 - \alpha$
- Los extremos a y b del intervalo son variables aleatorias que dependen de M y son tales que la probabilidad de que el intervalo capture en su interior al valor verdadero de θ es $1 - \alpha$.

Comentarios

- Un IC no es más que un par de estimadores puntuales a, b que proporcionan los límites inferior y superior del intervalo.
- Nuevos datos producirán un nuevo intervalo.
- El valor real (y que no conocemos) de θ es un número dado, está fijo aunque no sepamos su valor.
- Pese a la notación, el nivel de confianza $1 - \alpha$ no es la probabilidad de que θ pertenezca a $[a, b]$, es **la probabilidad de que** $[a, b]$ **capture a** θ **en su interior**.
- θ está fijo en una posición desconocida, $[a, b]$ intenta capturarlo, como lanzar una red.

Datos normales: IC para μ, (σ conocido)

- Sea $M = \{x_1, \ldots, x_n\}$ i.i.d donde $X \sim N(\mu, \sigma^2)$ donde **supondremos σ^2 conocido**.
- El intervalo $\left[\overline{X}_n - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \overline{X}_n + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right]$ es un IC al nivel de confianza $(1-\alpha)$ para μ. El centro es $\overline{X}_n$ y el radio es $z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$.
- Una forma de verlo es que si realizamos muchas veces el experimento de obtener n lecturas de X y para cada vez hallamos el IC, entonces el $(1-\alpha)\%$ de esos IC tendrán el valor real de μ en su interior (y el otro $\alpha\%$ no).

Observaciones

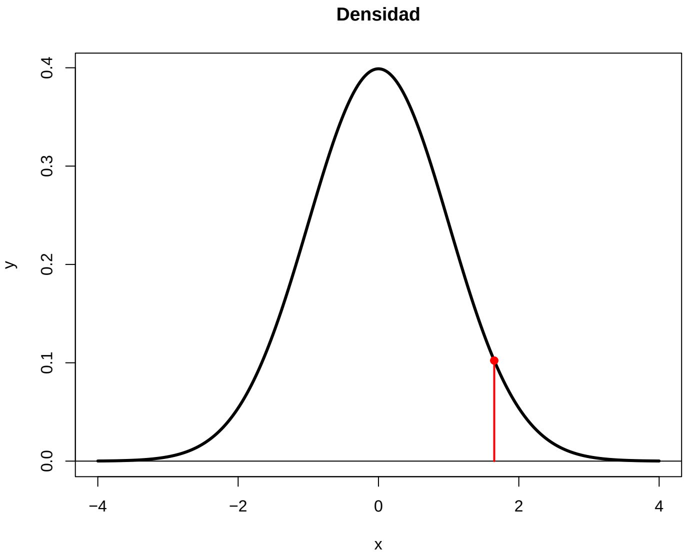
- Mientras mayor sea n, más angosto será el intervalo.
- Mientras mayor sea σ^2, más ancho será el intervalo.
- Mientras más cercano a cero sea α, más cerca de uno estará $1-\alpha$ por lo que IC será más ancho (estamos pidiendo que casi con certeza el IC capture al parámetro en su interior).
- Cuando veamos el test de hipótesis para la media de poblaciones normales con varianza conocida donde $H_0 : \mu = \mu_0$, veremos que el IC es precisamente la región de aceptación $\left[\mu_0 - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \mu_0 + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right]$,si el estadístico del test ($\overline{X}_n$ en este caso) cae en el mismo, no rechazamos H_0.
- La región crítica es el complemento del IC.

Datos normales: IC para μ, (σ desconocido)

1. Supongamos que nuestra muestra M proviene de una distribución $N(\mu, \sigma^2)$ **con ambos parámetros desconocidos**.
2. Si $\overline{X}_n \sim N(\mu, \sigma^2/n)$ entonces $\sqrt{n}\left(\frac{\overline{X}_n - \mu}{\sigma}\right) \sim N(0,1)$.
3. Esto es, si a una variable aleatoria normal se le resta su media y se la divide entre su desvío, obtenemos una variable normal estándar.
4. Si en lugar de usar σ usamos la varianza muestral σ_n (el promedio de los cuadrados menos el cuadrado del promedio), entonces $\sqrt{n}\left(\frac{\overline{X}_n - \mu}{\sigma_n}\right)$ tiene **distribución de Student con n-1 grados de libertad**.
5. El IC al nivel $1-\alpha$ para μ es $\left[\overline{X}_n - t_{n-1}(\alpha/2)\frac{\sigma_n}{\sqrt{n}}, \overline{X}_n + t_{n-1}(\alpha/2)\frac{\sigma_n}{\sqrt{n}}\right]$. El centro es $\overline{X}_n$, el radio es $t_{n-1}(\alpha/2)\frac{\sigma_n}{\sqrt{n}}$.

Observaciones

- El punto $z_{\alpha/2}$ es el punto en la campana de la normal estándar tal que a su derecha el área bajo la campana es $\alpha/2$ y el área a la izquierda de $-z_{\alpha/2}$ es $\alpha/2$.
- Esto es, el área bajo la campana en el intervalo $[-z_{\alpha/2}, z_{\alpha/2}]$ es igual a $1-\alpha$.
- De modo similar, $t_{n-1}(\alpha/2)$ es el punto bajo la campana de la Student con n-1 grados de libertad tal que el área bajo dicha campana a su derecha es $\alpha/2$ y el área a la izquierda de $-t_{n-1}(\alpha/2)$ es $\alpha/2$.
- O sea que el área bajo la campana de la Student con n-1 grados de libertad en el intervalo $[t_{n-1}(\alpha/2), -t_{n-1}(\alpha/2)]$ es $1-\alpha$.


Densidad
y
x
0.0
0.1
0.2
0.3
0.4
−4
−2
0
2
4

Datos no normales

- Supongamos que M es i.i.d pero la distribución de X no es normal. Sabemos que $E(X) = \mu$, el valor esperado o esperanza de X es μ.
- Se cumple que si llamamos $Z_n = \sqrt{n}\left(\frac{\overline{X}_n - \mu}{\sigma_n}\right)$, Z_n tiene una distribución que **tiende a la distribución normal estándar cuando n crece**.

  El IC al nivel $1 - \alpha$ para μ para este caso es $\left[\overline{X}_n \pm z_{\alpha/2}\frac{\sigma_n}{\sqrt{n}}\right]$
- Es parecido al caso de muestras normales con σ conocido, pero con σ_n en lugar de σ.

Ejemplo: Población normal, σ conocido

- Generamos 10 datos $x_1, \ldots, x_{10}$ i.i.d $\sim N(\mu, \sigma^2)$ con $\mu = 30$ y $\sigma = 5$.
- datos=normrnd(30,5,1,10) con esta sentencia una fila con 10 valores $\sim N(30, 25)$.
- Fijemos $\alpha = 0.05$ de modo que $1 - \alpha = 0.95$. Buscaremos el valor $z_{\alpha/2}$, el punto que deja a su derecha el 2.5% del área bajo la campana de la normal estándar y el 97.5% del área a su izquierda.
- Lo hacemos con el comando zalfa2=norminv(0.975,0,1) obtenemos ans=1.96.
- El promedio es $\overline{X}_{10} = 32.734$, se obtiene haciendo prom=mean(datos). Como n=10, $\sqrt{10} = 3.162$

Ejemplo: Población normal, σ conocido

- Como $\sigma = 5$, el IC al nivel 0.95 para μ obtenido de estos datos es $[32.734 \pm 1.96\frac{5}{3.162}] = [32.734 \pm 3.099]$, o sea $[29.635, 35.833]$.
- Si generamos otros 10 datos con la misma distribución y repetimos el cálculo obtendremos otro IC ligeramente distinto $(\overline{X}_{10} = 30.145)$.
- El ancho del intervalo solo depende de σ y de n, queda igual. El nuevo IC es $[30.145 \pm 3.099] = [27.046, 33.244]$.
- Si repetimos muchas veces esto tendremos muchos IC de los cuales uno espera que el 95 % logre capturar el verdadero valor $\mu = 30$ en su interior, mientras que el restante 5 % fallará.
- Como habitualmente uno tiene una muestra y hace esto una sola vez, confiamos en que nuestro IC sea uno de ese 95% de los que capturan a μ.

Ejemplo: Población normal, σ desconocido

- Para la misma muestra (la de promedio 30.145) supongamos que no conocemos σ. Fijado $\alpha = 0.05$ debemos buscar el valor $t_{n-1}(\alpha/2)$ en la distribución Student con 9 grados de libertad (recordemos que n=10).
- Esto lo hacemos con t=tinv(0.975,9). Obtenemos t=2.262. En lugar de σ usamos $\sigma_n = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{X}_n)^2}$
- Para ello hacemos sigma=std(datos), obtenemos sigma=5.413. Cuando hacemos $\frac{\sigma_n}{\sqrt{10}}$ obtenemos 1.711.
- Por tanto el IC queda $[30.145 \pm 2.262 \times 1.711] = [26.275, 34.015]$.
- Queda más ancho que cuando se conocía σ, lo que tiene sentido.