Duración de Baterías

```
#Intervalo de nivel 0.95

duracion<- c(237,242,232,242,248,230,244,243,254,
             262,234,220,225,246,232,218,228,240)

ene<- length(duracion)

mean(duracion)
[1] 237.6111

sd(duracion)
[1] 11.46934

t.alpha2<- qt(1-0.025, ene-1)

semi<- t.alpha2*sd(duracion)/sqrt(ene)

c(mean(duracion)- semi,mean(duracion)+ semi)
[1] 231.9075 243.3147
```

Volviendo al Teorema...

Notemos que...

si $X_1, \ldots, X_n$ i.i.d., $X_i \sim N(\mu, \sigma^2)$, entonces

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

sin embargo, al reemplazar μ por $\bar{X}_n$ tenemos que

$$\frac{\sum_{i=1}^{n} \left(X_i - \bar{X}_n \right)^2}{\sigma^2} \sim \chi_{n-1}^2 \ , \quad \text{es decir } \frac{(n-1)S_n^2}{\sigma^2} \sim \chi_{n-1}^2$$

$$\text{ya que} \quad S_n^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X}_n)^2$$

Intervalos de confianza para σ^2 con μ desconocido

Tenemos $X_1, \ldots, X_n$ i.i.d., $X_i \sim N(\mu, \sigma^2)$. Ahora buscamos intervalo de confianza para σ^2.

- Pivote: $\frac{(n-1)S_n^2}{\sigma^2} \sim \chi^2_{n-1}$
- Sea $\chi^2_{k,\beta}$ tal que $\mathbb{P}(U > \chi^2_{k,\beta}) = \beta$ cuando $U \sim \chi^2_k$, entonces

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

Luego,

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

es un intervalo de confianza de nivel $1 - \alpha$ para σ^2 bajo el modelo normal cuando μ desconocida.

Intervalos de confianza para σ^2 con μ conocida. (Ejercicio)

$\mu = \mu_o$

Buscamos intervalo de confianza para σ^2

- Pivote:

$$\frac{\sum_{i=1}^{n}(X_i - \mu_o)^2}{\sigma^2} \sim \chi^2_n$$

- Entonces:

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

y por lo tanto,

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

es un intervalo de confianza de nivel $1 - \alpha$ para σ^2 bajo el modelo normal cuando $\mu = \mu_o$ conocida.

Seguimos en el mundo normal...

- $X_i \sim \mathcal{N}(\mu, \sigma^2)$ i.i.d. Buscamos intervalo de confianza para μ.
- σ desconocido .....llegó la t...
- IC nivel $1 - \alpha$ para μ

$$\left(\bar{X}_n - t_{n-1,\alpha/2} \, \frac{S_n}{\sqrt{n}} \quad , \quad \bar{X}_n + t_{n-1,\alpha/2} \, \frac{S_n}{\sqrt{n}} \right)$$

- longitud $\rightarrow L = 2 \, t_{n-1,\alpha/2} \, \frac{S_n}{\sqrt{n}}$

Seguimos en el mundo normal...

- $X_i \sim \mathcal{N}(\mu, \sigma^2)$ i.i.d. Buscamos intervalo de confianza para μ.
- σ desconocido .....llegó la t...
- IC nivel $1-\alpha$ para μ

$$\left(\bar{X}_n - t_{n-1,\alpha/2}\, \frac{S_n}{\sqrt{n}} \quad , \quad \bar{X}_n + t_{n-1,\alpha/2}\, \frac{S_n}{\sqrt{n}} \right)$$

- longitud $\rightarrow L = 2\, t_{n-1,\alpha/2}\, \frac{S_n}{\sqrt{n}}$

- L es una v.a. Si queremos que $l \leq l_o$ ¿Cómo hacemos?

- Se debe trabajar con una muestra piloto, sugerencia: ver Notas de Estadística de Boente & Yohai.

Un ejemplo fuera del mundo normal...

Intervalos de confianza para λ de una $\mathcal{E}(\lambda)$

Sea $X_1, X_2, \ldots, X_n$ una m.a. de una distribución $E(\lambda)$. Hallar un intervalo de confianza de nivel $1-\alpha$ para λ.

Un ejemplo fuera del mundo normal...

Intervalos de confianza para λ de una $\mathcal{E}(\lambda)$

Sea $X_1, X_2, \ldots, X_n$ una m.a. de una distribución $E(\lambda)$. Hallar un intervalo de confianza de nivel $1-\alpha$ para λ.

Recordemos:
- Una $\mathcal{E}(\lambda)$ es una $\Gamma(1, \lambda)$.
- La distribución de la suma de exponenciales de igual parámetro es Gamma, es decir:

$$\sum_{i=1}^{n} X_i \sim \Gamma(n, \lambda)$$

- Una constante positva por una Gamma es Gamma:

$$U \sim \Gamma(\alpha, \lambda) \text{ y } a > 0 \Rightarrow aU \sim \Gamma\left(\alpha, \frac{\lambda}{a}\right)$$

Intervalos de confianza para λ de una $\mathcal{E}(\lambda)$

Entonces,

$$\sum_{i=1}^{n} X_i \sim \Gamma(n, \lambda)$$

$$\lambda \sum_{i=1}^{n} X_i \sim \Gamma(n, 1)$$

$$2\lambda \sum_{i=1}^{n} X_i \sim \Gamma\left(n, \frac{1}{2}\right) = \Gamma\left(\frac{2n}{2}, \frac{1}{2}\right) = \chi^2_{2n}$$

Luego, el pivote para la exponencial:

$$G\left(X_1, X_2, \ldots, X_n, \lambda\right) = 2\lambda \sum_{i=1}^{n} X_i$$

Intervalos de confianza para λ de una $\mathcal{E}(\lambda)$

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

$$P\left(\frac{\chi_{2n,1-\alpha/2}}{2\sum_{i=1}^{n} X_i} \leq \lambda \leq \frac{\chi_{2n,a/2}}{2\sum_{i=1}^{n} X_i}\right) = 1-\alpha$$

Por lo tanto, el intervalo de confianza para λ es

$$\left[\frac{\chi_{2n,1-\alpha/2}}{2\sum_{i=1}^{n} X_i}, \frac{\chi_{2n,a/2}}{2\sum_{i=1}^{n} X_i}\right]$$

¿Y si no conocemos la distribución?

¿Y si no conocemos la distribución?
¿Cuántos datos tenemos?

Regiones de confianza con nivel asintótico $(1-\alpha)$

Sea $X_1, X_2, \ldots, X_n$ m.a. $X_i \sim F(x, \theta)$, $\theta \in \Theta$. Se dice que $S_n(X_1, \ldots, X_n)$ es una sucesión de regiones de confianza con nivel asintótico $1-\alpha$ si:

$$\lim_{n\to\infty} \mathbb{P}_\theta(\theta \in S_n(X_1, \ldots, X_n)) = 1-\alpha \quad \forall\, \theta \in \Theta\ .$$

Procedimiento para obtener RC con nivel asintótico

Teorema Sea $X_1, \ldots, X_n$ una muestra aleatoria de una distribución perteneciente a la familia $F(x, \theta)$, $\theta \in \Theta$.
Supongamos que

- $\forall n$, $\exists$ v.a. $U_n = G_n(X_1, \ldots, X_n, \theta)$ tal que $U_n \xrightarrow{\mathcal{D}} U$, donde U es una variable aleatoria con distribución independiente de θ
- A y B puntos de continuidad de F_U tales que $\mathbb{P}(A \leq U \leq B) = 1 - \alpha$.

Luego, si

$$\mathcal{S}_n(X_1, \ldots, X_n) = \{\theta : A \leq G_n(X_1, \ldots, X_n, \theta) \leq B\}$$

$\mathcal{S}_n(\mathbf{X})$ es una sucesión de RC con nivel asintótico $(1 - \alpha)$.

Ejemplos

Intervalo de confianza asintótico para la media de una distribución

- Consideremos el caso en que $X_1, \ldots, X_n$ son una muestra aleatoria $\mathbb{E}(X_i) = \mu$ y $\mathbb{V}ar(X_i) = \sigma^2$, ambas desconocidas. Veremos que

$$\left[\bar{X}_n - z_{\frac{\alpha}{2}} \frac{S_n}{\sqrt{n}}, \bar{X}_n + z_{\frac{\alpha}{2}} \frac{S_n}{\sqrt{n}}\right]$$

  es un intervalo para $\mu = \mathbb{E}(X_i)$ tiene nivel asintótico $1 - \alpha$.

Vayamos al pizarrón.

Ejemplos

Intervalo de confianza asintótico para el parámetro p de la distribución binomial

Sea $X \sim \mathrm{Bi}(n, p)$. Podemos pensar a X como el número de éxitos en las n repeticiones en un experimento binomial con probabilidad de éxito p. Consideremos para cada $i = 1, \ldots, n$,

$$X_i = \begin{cases} 1 & \text{si se obtuvo éxito en la } i \text{ ésima repetición} \\ 0 & \text{si se obtuvo fracaso en la } i \text{ ésima repetición} \end{cases}.$$

Estas v.a. son independientes y para todo i tenemos que $X_i \sim \mathrm{Bi}(1, p)$ y

$$X = \sum_{i=1}^{n} X_i$$

Además: $\mathbb{E}(X_i) = p$ y $\mathbb{V}ar(X_i) = p(1-p)$ Luego, por T.C.L. para n suficientemente grande

$$\frac{X - np}{\sqrt{np(1-p)}} \overset{a}{\sim} N(0,1) \quad \text{y} \quad \frac{\overline{X} - p}{\sqrt{p(1-p)/n}} \overset{a}{\sim} N(0,1)$$

Ejemplos

Intervalo de confianza asintótico para el parámetro p de la distribución binomial

Sea $X_1, X_2, \ldots, X_n$ una m.a. de una distribución $B(1,p)$. En este contexto, consideremos $\hat{p} = \overline{X}$. Por lo visto,

$$\frac{\overline{X} - p}{\sqrt{p(1-p)/n}} \xrightarrow{\mathcal{D}} N(0,1)$$

implica que

$$P\left(-z_{a/2} \leq \frac{\overline{X} - p}{\sqrt{\frac{p(1-p)}{n}}} \leq z_{a/2}\right) \overset{n\to\infty}{\longrightarrow} 1 - \alpha$$

¿Cómo despejamos p?

Alternativa 1: Explorar

$[\widehat{p}_{1,n}, \widehat{p}_{2,n}]$ raíces del polinomio en p

$$n\bar{X}_n^2 - p(2n\bar{X}_n + z_{\frac{\alpha}{2}}^2) + p^2(z_{\frac{\alpha}{2}}^2 + n)$$

Ejemplos

Intervalo de confianza asintótico para el parámetro p de la distribución binomial

Alternativa 2

Por la Ley de los Grandes Números,

$$\hat{p} = \overline{X} = \frac{\sum_{i=1}^{n} X_i}{n} \xrightarrow{p} p.$$

Reemplazamos en el denominador del pivote a p por su estimador, aplicando el Teorema de Slutzky, resulta que la distribución asintótica es la misma, o sea:

$$\frac{\overline{X} - p}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}} \xrightarrow{\mathcal{D}} N(0,1)$$

y por lo tanto

$$P\left(-z_{a/2} \leq \frac{\overline{X} - p}{\sqrt{\frac{\overline{X}(1-\overline{X})}{n}}} \leq z_{a/2}\right) \cong 1 - \alpha$$

Ejemplos

Intervalo de confianza asintótico para el parámetro p de la distribución binomial

Por lo tanto, un intervalo para p de nivel asintótico $1-\alpha$ es

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

o, lo que es lo mismo,

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

Ejemplo

¿Cómo resultaría en nuestro ejemplo de la estimación de la proporción de apoyo a las vacunas el intervalo de nivel asintótico 0.95?

Usando la distribución asintótica de los EMV

$X_1, \ldots, X_n$ i.i.d. donde X_i tienen función de densidad o de probabilidad puntual $f(x, \theta)$.
Bajo condiciones de regularidad hemos visto que

- si $\widehat{\theta}_n = \widehat{\theta}_n^{EMV}$, entonces $\sqrt{n}(\widehat{\theta}_n - \theta) \xrightarrow{\mathcal{D}} N(0, \frac{1}{I(\theta)})$
- $$\sqrt{n}\sqrt{I(\theta)}(\widehat{\theta}_n - \theta) \xrightarrow{\mathcal{D}} N(0, 1)$$
- La región

$$S(\mathbf{X}) = \{\theta : -z_{\frac{\alpha}{2}} \leq \sqrt{n}\sqrt{I(\theta)}(\widehat{\theta}_n - \theta) \leq z_{\frac{\alpha}{2}}\}$$

no tiene porqué ser un intervalo y puede ser difícil de calcular.

Usando la distribución asintótica de los EMV

- Si $I(\theta)$ es una función continua de θ, como $\widehat{\theta}_n \xrightarrow{P} \theta$, bajo condiciones de regularidad obtendremos que $I(\widehat{\theta}_n) \xrightarrow{P} I(\theta)$, luego

$$\sqrt{n}\sqrt{I(\widehat{\theta}_n)}(\widehat{\theta}_n - \theta) \xrightarrow{\mathcal{D}} N(0,1)$$

- Entonces, un intervalo de nivel asintótico $1-\alpha$ será

$$\left[\widehat{\theta}_n - z_{\frac{\alpha}{2}}\sqrt{\frac{1}{n\ I(\widehat{\theta}_n)}}, \widehat{\theta}_n + z_{\frac{\alpha}{2}}\sqrt{\frac{1}{n\ I(\widehat{\theta}_n)}}\right]$$