Интервали поверења

деф. $(X_1, X_2, \dots, X_n)$ узорак из расподеле која зависи од параметра θ, U_n и V_n статистике.

- Случајни интервал (U_n, V_n) је $(1-\alpha)\cdot 100-\%$-ни двострани интервал поверења за параметар θ ако:

$$P\{U_n \le \theta \le V_n\} = 1-\alpha$$

Једнострани интервали поверења:

доњи $(1-\alpha)\cdot 100-\%$-ни: ограничава θ са доње стране:

$$P\{U_n \le \theta\} = 1-\alpha$$

горњи $(1-\alpha)\cdot 100-\%$-ни: ограничава θ са горње стране

$$P\{\theta \le V_n\} = 1-\alpha$$

Тривијални примери:

- $X_i \sim \mathcal{E}(\lambda)$: $(0, +\infty)$ је 100%-ни интервал поверења за λ
- $X_i \sim Ber(p)$: $(0,1)$ је 100%-ни интервал поверења за p

Од ова два интервала немамо користи јер они обухватају цео домен нашег параметра, тј. не нуде нам нове информације о локацији параметра.
Зато не посматрамо 100%-не интервале, већ углавном 99%-не, 95%-не ...
Желимо да интервал поверења буде што ужи, али то повлачи мању вероватноћу да он обухвати параметар

Како тражимо U_n и V_n?

- обично пођемо од неке сл. величине чију расподелу знамо, а која зависи и од нашег узорка и од параметра.

1.

$(X_1, \dots, X_n)$ псу $\mathcal{N}(m, \sigma^2)$, $\sigma^2 = 1{,}79$, $n=25$, $\bar{x}_n = 2{,}6$

$1-\alpha = 0{,}95 \Rightarrow \alpha = 0{,}05$

тврђење $X_i \sim \mathcal{N}(m_i, \sigma_i^2)$, $i \in \{1, 2, \dots, n\}$, X_i независне

Тада $a_1 X_1 + \dots a_n X_n \sim \mathcal{N}(a_1 m_1 + \dots a_n m_n, a_1^2 \sigma_1^2 + \dots a_n^2 \sigma_n^2)$

Дакле, $\bar{X}_n = \frac{X_1}{n} + \dots + \frac{X_n}{n} \sim \mathcal{N}(\underbrace{\frac{1}{n} m + \dots \frac{1}{n} m}_{= m}, \underbrace{\frac{1}{n^2}\sigma^2 + \dots \frac{1}{n^2}\sigma^2}_{= \frac{\sigma^2}{n}})$

$$\Rightarrow \underbrace{\frac{\bar{X}_n - m}{\sigma}\sqrt{n}}_{\text{стожер}} \sim \mathcal{N}(0,1)$$

Поступак тражења U_n и V_n је следећи:
Тражимо c_1 и c_2 т.д.

$$P\left\{ c_1 \le \frac{\bar{X}_n - m}{\sigma}\sqrt{n} \le c_2 \right\} = 1-\alpha$$

$$\Leftrightarrow P\left\{ \frac{c_1 \sigma}{\sqrt{n}} \le \bar{X}_n - m \le \frac{c_2 \sigma}{\sqrt{n}} \right\} = 1-\alpha$$

$$\Leftrightarrow P\left\{ \underbrace{\bar{X}_n - c_2 \frac{\sigma}{\sqrt{n}}}_{U_n} \le m \le \underbrace{\bar{X}_n - c_1 \frac{\sigma}{\sqrt{n}}}_{V_n} \right\} = 1-\alpha$$

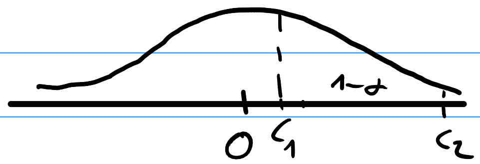
$\Rightarrow (U_n, V_n)$ је $(1-\alpha) \cdot 100\%$-ни инт. пов. за m

ширина интервала: $= V_n - U_n = (c_2 - c_1)\frac{\sigma}{\sqrt{n}}$

да би $c_2 - c_1$ било најмање, потребно је да важи: $\boxed{c_2 = -c_1}$

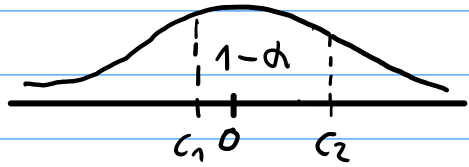
желимо да буде што мања

1°

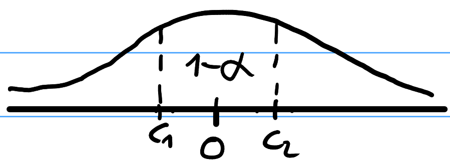


овде су c_1 и c_2 даљи него у 2°

2°



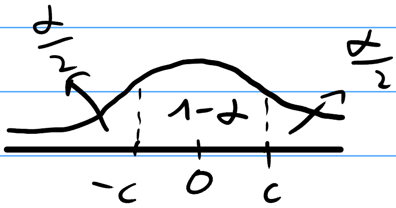
3°



$c_1=-c_2$

Овде су c_1 и c_2 на мањем растојању него у 1° и 2°, па бирамо c_1 и c_2 т.д. $c_1=-c_2$

треба нам c т.д. $P\left\{-c \le \frac{\bar{X}_n - m}{\sigma}\sqrt{n} \le c\right\} = 1-\alpha$



$$\phi(c) = 1-\frac{\alpha}{2} \Rightarrow c = \phi^{-1}\left(1-\frac{\alpha}{2}\right)$$

у R-у:

$$c = qnorm\left(1-\frac{\alpha}{2}\right)$$

$$\Rightarrow (U_n, V_n) = \left(\bar{X}_n - c\frac{\sigma}{\sqrt{n}}, \bar{X}_n + c\frac{\sigma}{\sqrt{n}}\right)$$

Када убацимо $\bar{x}_n$ и σ, добијамо:

$$I = (2{,}076, 3{,}124)$$

Веома често тумачење интервала поверења, АЛИ ПОГРЕШНО је

$$P\{2{,}076 \le m \le 3{,}124\} = 0{,}95$$

↳ фиксно

Правилно тумачење је следеће:

Вероватноћа да интервал $\left(\bar{X}_n - c\frac{\sigma}{\sqrt{n}}, \bar{X}_n + c\frac{\sigma}{\sqrt{n}}\right)$ обухвати m је 0,95. Дакле, ако бисмо симулирали 100 оваквих интервала, приближно 95 њих ће обухватити m. Видети код у R-у.