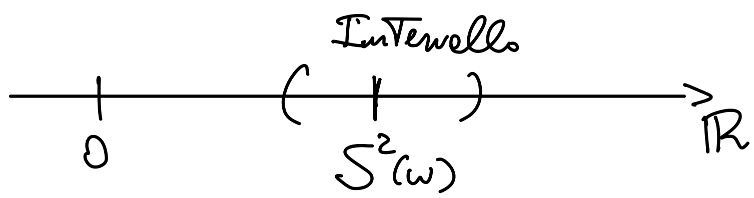
- Intervalli di fiducia per la varianza di un campione gaussiano

$X_1,\dots,X_n \quad , \quad X_k \sim N(\mu,\sigma^2) \quad , \text{ stimiamo } \sigma^2$

Statistica campionaria

$$S^2 := \frac{1}{n-1}\sum_{k=1}^{n}(X_k-\bar{X})^2 \quad \text{corretta}$$

$$\frac{n-1}{n}S^2 \quad \text{stimatore di massima verosimiglianza}$$



Cerchiamo $d>0$ t.c.

$$1-\alpha = \mathbb{P}_{\sigma^2}\left(|S^2-\sigma^2|<d\right) = \mathbb{P}_{\sigma^2}\left(\cancel{\Big|\underbrace{\frac{(n-1)S^2}{\sigma^2}}_{=C_{n-1}}-(n-1)\Big|<\frac{(n-1)d}{\sigma^2}}\right)$$

σ^2 apart. a $[S^2(\omega)-d, S^2(\omega)+d]$

$$1-\alpha = \mathbb{P}_{\sigma^2}\left(d_1 < \overbrace{\frac{(n-1)S^2}{\sigma^2}}^{C_{n-1}} < d_2\right) = \mathbb{P}_{\sigma^2}\left(\frac{(n-1)S^2}{d_2} < \sigma^2 < \frac{(n-1)S^2}{d_1}\right)$$

$$\shortparallel$$

$$F_{\chi^2(n-1)}(d_2) - F_{\chi^2(n-1)}(d_1)$$

$d_2 = \chi^2_{(1-\alpha/2,\, n-1)}$

$d_1 = \chi^2_{(\alpha/2,\, n-1)}$

$\Rightarrow$ Int. bilatero $\left[\dfrac{(n-1)\bar{\sigma}^2}{\chi^2_{(1-\alpha/2,\, n-1)}} , \dfrac{(n-1)\bar{\sigma}^2}{\chi^2_{(\alpha/2,\, n-1)}}\right]$

Intervalli unilateri

- destro

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

$$\shortparallel$$

$$F_{\chi^2(n-1)}(d) = 1-\alpha \quad \Rightarrow \quad d = \chi^2_{(1-\alpha,\, n-1)}$$

Intervallo unilatero destro al livello di fiducia $1-\alpha$

è $\left[\frac{(n-1)\bar{\sigma}^2}{\chi^2_{(1-\alpha,\, n-1)}}, +\infty\right)$

- sinistro

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

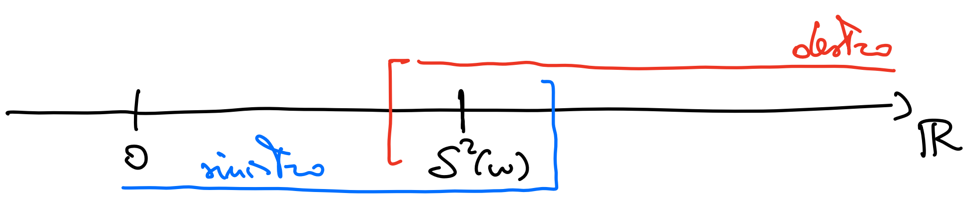
$$\shortparallel$$

$$1 - F_{\chi^2(n-1)}(d) = 1-\alpha$$

$$\Rightarrow \quad d = \chi^2_{(\alpha,\, n-1)}$$

Intervallo unilatero sinistro al livello di fiducia $1-\alpha$ è

$$\left(0, \frac{(n-1)\bar{\sigma}^2}{\chi^2_{(\alpha,\, n-1)}}\right]$$



$S^2(\omega) = \bar{\sigma}^2$

Esempi

- Vogliamo stimare la deviazione standard di una bilancia. Pesiamo un oggetto per 10 volte ottenendo i valori in g

12.3, 12.4, 13.3, 12.5, 12.6, 12.8, 12.0
12.4, 13.0, 12.6

Qual è un intervallo di fiducia sinistro per σ al livello di fiducia del 90 %?

$X_1 \dots X_{10}$, $X_k \sim N(\mu, \sigma^2)$, stimare σ^2.

$n = 10$, $1-\alpha = 0.9 \Rightarrow \alpha = 0.1$

$$I_s = \left(0, \frac{(n-1)\,\overline{\sigma}^2}{\chi^2_{(\alpha, n-1)}}\right] = \left(0, \frac{9 \cdot \overline{\sigma}^2}{\chi^2_{(0.1, 9)}}\right]$$

$$\overline{\sigma}^2 := \frac{1}{n-1} \sum_{k=1}^{n} (x_k - \overline{x})^2 = \frac{1}{n-1} \left[\sum_{k=1}^{n} x_k^2 - n\,\overline{x}^2 \right] \sim 0.1366$$

$$\chi^2_{(0.1, 9)} \sim 4.1682$$

$$I_s = (0, 0.2949] \quad , \quad \sigma < 0.543$$

Stime sul peso dell'oggetto? $X_k \sim N(\mu, \sigma^2)$, stimare μ.
σ^2 ignota

Con livello di fiducia del 99 %

$$I = \left[\bar{x} - \frac{\bar{\sigma}}{\sqrt{n}} \, t_{(1-\alpha/2,\, n-1)} \,,\ \bar{x} + \frac{\bar{\sigma}}{\sqrt{n}} \, t_{(1-\alpha/2,\, n-1)} \right]$$

precisione della stima $\frac{\bar{\sigma}}{\sqrt{n}} \, t_{(1-\alpha/2,\, n-1)} \sim 0.3798$

precisione relativa $= \frac{\text{prec.}}{|\bar{x}|} \sim 3 \times 10^{-2}$

$1-\alpha = 0.99$, $\alpha = 0.01$, $1-\alpha/2 = 0.995$

$t_{(0.995,\, 9)} \sim 3.2498$

$$I = [\, 12.21 \,,\ 12.97 \,]$$

Esempio Una ditta produce dadi di cui vogliamo stimare la variabilità dello spessore.
Misuriamo lo spessore di un campione di 100 dadi e troviamo $\bar{x} = 0.53$ cm e $\bar{\sigma}^2 = 0.012$ cm.
Costruiamo un int. di fiducia destro al livello del 95 %.

$X_1 \dots X_{100}$, $X_k \sim N(\mu, \sigma^2)$ μ, σ^2 ignote

Stimare σ^2, $I_d = \left[\frac{(n-1)\,\bar{\sigma}^2}{\chi^2_{(1-\alpha,\, n-1)}} \,,\ +\infty \right)$

$n = 100$, $1-\alpha = 0.95$, $\bar{\sigma}^2 = 0.012$

$\chi^2_{(0.95,\, 99)} \sim 124.3421$ ($R \to 123.2252$)

$$I_d = [0.0096, +\infty) \quad [R \to [0.0096, +\infty)]$$

$$\sigma > 0.0977$$

- Intervallo di fiducia per la media di un campione statistico di v.a. di Bernoulli

$X_1, \dots, X_n \quad , \quad X_k \sim B(1, p) \quad , \quad E[X_k] = p$

$\mathrm{Var}(X_k) = p(1-p)$

Stimare p.

Statistica campionaria $\overline{X}$, stima corretta.

$$\overline{X} = \frac{X_1 + \dots + X_n}{n} \sim \frac{1}{n} \cdot B(n, p)$$

$$\frac{X_1 + \dots + X_n - np}{\sqrt{n\,p(1-p)}} \approx N(0,1) \quad \text{per } n \text{ grande} \quad \left(\begin{array}{c}\text{Teo del} \\ \text{Limite} \\ \text{Centrale}\end{array}\right)$$

$$\parallel$$

$$\frac{\sqrt{n}}{\sqrt{p(1-p)}} (\overline{X} - p) \approx N(0,1)$$

Cerchiamo $d > 0$ t.c.

$$1 - \alpha = \mathbb{P}_p\left(\Big| \underbrace{\frac{\sqrt{n}}{\sqrt{p(1-p)}} (\overline{X} - p)}_{\approx N(0,1)} \Big| < d \right) =$$

$$= \cancel{\mathbb{P}_p\left(-\frac{\sqrt{p(1-p)}}{\sqrt{n}} d + \overline{X} < p < \overline{X} + \frac{\sqrt{p(1-p)}}{\sqrt{n}} d \right)}$$

p viene approssimato da $\overline{X}$

$$1-\alpha = \mathbb{P}_p\left(\underbrace{\left| \frac{\sqrt{n}}{\sqrt{\overline{X}(1-\overline{X})}} (\overline{X}-p) \right|}_{\approx N(0,1)} < d \right) =$$

$$= \mathbb{P}_p\left(\overline{X} - \frac{\sqrt{\overline{X}(1-\overline{X})}}{\sqrt{n}} d < p < \overline{X} + \frac{\sqrt{\overline{X}(1-\overline{X})}}{\sqrt{n}} d \right)$$

$$= \Phi(d) - \Phi(-d) = 2\Phi(d) - 1 = 1-\alpha$$

$$\Longleftrightarrow \Phi(d) = 1 - \alpha/2 \Longleftrightarrow d = q_{1-\alpha/2}.$$

L'intervallo di fiducia bilatero per il parametro p al livello di fiducia $1-\alpha$ è

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

$\overline{X}(\omega) = \hat{p}$, precisione della stima $= \frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}} q_{1-\alpha/2}$

<u>Es</u> Costruire gli intervalli unilateri.

<u>Esempi</u> Una ditta produce transistor, e vuole stimare la bontà del prodotto. Su un campione di 1000 transistor, ne troviamo 900 funzionanti. Stimare la probabilità che un transistor

fenzioni al livello del 95 %.

$X_1 \ldots X_{1000}$, $n = 1000$, $X_k \sim B(1, p)$, stimare p.

$X_1 + \ldots + X_{1000} = 900$, $\hat{p} = \overline{X}(\omega) = \frac{900}{1000} = 0.9$

$1 - \alpha = 0.95$, $1 - \alpha/2 = 0.975$

$q_{1-\alpha/2} \sim 1.96$ [$R \to 1.959964$]

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

$$= [0.8814 ,\ 0.9186]$$

prec. della stima ~ 0.0186

OSS $X_k \sim N(\mu, \sigma^2)$ σ^2 nota

prec. della stima $\frac{\sigma}{\sqrt{n}} q_{1-\alpha/2}$

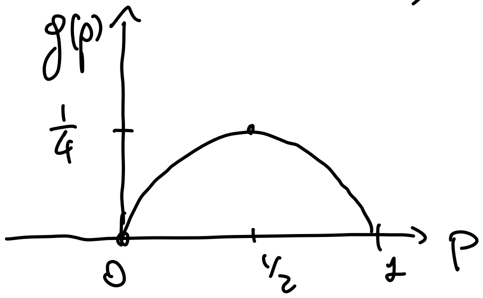
- n
- α

$X_k \sim N(\mu, \sigma^2)$ σ^2 ignota — prec. della stima $\tau_{(1-\alpha/2, n-1)} \frac{\overline{\sigma}}{\sqrt{n}}$

$X_k \sim B(1, p)$ — " — $q_{1-\alpha/2} \frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}}$

I idea Utilizzare un piccolo campione iniziale per avere $\bar{\sigma}$ o $\hat{p}$ da inserire nella formula.

II idea Nel caso $B(1,p)$, usiamo $g(p) = p(1-p)$



Possiamo dire che $g(p) \le \frac{1}{4} \ \forall p \in [0,1]$, e quindi se

$$q_{1-\alpha/2} \frac{\sqrt{1/4}}{\sqrt{n}} < c$$

allora

$$q_{1-\alpha/2} \frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}} < c$$

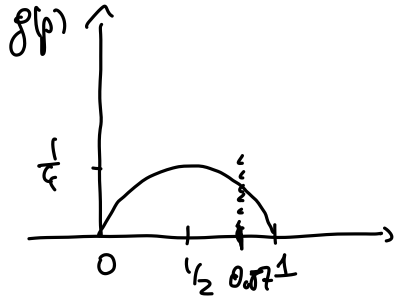
Esempio Un'azienda produce circuiti integrati che sono funzionanti con prob. p. Quanto deve essere grande il campione per ottenere una precisione della stima per $p \le 0.05$ al livello di fiducia del 95%?

- $X_1, \dots, X_n$, $X_k \sim B(1,p)$

prec. stima $\quad q_{1-\alpha/2} \dfrac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}} \underset{\forall \hat{p}}{\le} \boxed{q_{1-\alpha/2} \dfrac{\sqrt{1/4}}{\sqrt{n}} \le 0.05}$

$$\sqrt{n} \ge q_{0.975} \sqrt{\frac{1}{4}} \frac{1}{0.05} \sim \sqrt{384.15} \iff n \ge 385$$

• Supponiamo di sapere che $\hat{p} \ge 0.87$



$$g(p)\Big|_{[0.87,1]} \le g(0.87)$$

prec. di me $\quad q_{1-\alpha/2}\dfrac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}} \le \boxed{q_{1-\alpha/2}\dfrac{\sqrt{0.87(1-0.87)}}{\sqrt{n}} \le 0.05}$

$\forall \hat{p} \in [0.87, 1]$

$\Rightarrow \quad n \ge 174$