

ESERCITAZIONE 8

- 1) In un processo di controllo qualità emerge che su 371 pezzi controllati, 18 sono difettosi. Trattando il campione come un campione casuale
- Calcola $L(\theta, \underline{x}_n)$ e il valore $L(p, \underline{x}_n)$ con p = proporzione dei pezzi difettosi nella popolazione (parametro ignoto).
 - Dettemo $IC_{90\%}$ approssimato per p
 - Allo luce dei dati osservati riteni più verosimile $p=0.07$ oppure $p=0.03$?

i) Possiamo vedere lo schema osservato come realizzazione di $X_i | p \sim \text{Bern}(p)$

$$X_i = \begin{cases} 1 = \text{difettoso} & p \\ 0 = \text{non difettoso} & 1-p \end{cases} \quad p = \frac{\text{pezzi difettosi}}{\text{pezzi totali}} = \text{probabilità di incontrare un pezzo difettoso.}$$

Quindi Dato $X_i | p \sim \text{Bern}(p) \quad \forall i=1, \dots, n$ col costrutto lo funzione di verosimiglianza: ricordando che $f_X(x, p) = p^x (1-p)^{1-x}$

$$L(p, \underline{x}_n) = \prod_{i=1}^n f_X(x_i, p) = \prod_{i=1}^n (p^{x_i} (1-p)^{1-x_i}) = p^{\sum x_i} (1-p)^{n - \sum x_i}$$

$$= p^{18} (1-p)^{371-18}$$

per $n=371$
 $\sum x_i = 18$
 il valore dei difetti

ii) Costruiamo $IC_{90\%}$ approx per p :

Ricordo che $\frac{\hat{\theta}_{MV} - \theta}{\sqrt{I_n^{-1}(\hat{\theta}_{MV})}} \sim N(0,1) \rightarrow$ questo caso quadratico puntuale

$$\text{Calcolo } \hat{\theta}_{MV} \Rightarrow \frac{d}{dp} L(p, \underline{x}_n) = \frac{d}{dp} l(p, \underline{x}_n) = \frac{d}{dp} \log(p^{\sum x_i} (1-p)^{n - \sum x_i})$$

$$= \frac{\sum x_i}{p} - \frac{(n - \sum x_i)}{1-p} = 0 \Rightarrow \sum x_i - p \sum x_i = n p - p \sum x_i \Rightarrow \hat{p}_{MV} = \hat{\theta}_{MV} = \bar{X}_n$$

Come calcolo $I_n^{-1}(\hat{\theta}_{MV})$? Osservo che $\hat{\theta}_{MV} = \bar{X}_n = \frac{\sum x_i}{n}$ è funzione dello statistica sufficiente minimale $(\sum x_i)$ (si vede con criterio di FATTORIZZAZIONE DI FISHER), il modello appartiene alla famiglia esponenziale (famiglia Bernoulli) e quindi la statistica $(\sum x_i)$ è anche completa ed è garantito il raggiungimento da parte della variata dello stimatore ottimo al limite inferiore di Cramer-Rao (ovvero $V(\hat{\theta}_{MV}) = \text{CR}(0) = I_n^{-1}(\theta)$).

Inoltre $E(\hat{\theta}_{MV}) = E(\bar{X}_n) = E(X) = p$ è non distorto per p e quindi

Rao-Blackwell e Lehman-Scheffe $\hat{\theta}_{MV} = \bar{X}_n$ è unico per p .

E per Cramer-Rao $V(\hat{\theta}_{MV}) = V(\bar{X}_n) = c(\theta)$

$$V(\bar{X}_n) = \frac{V(X)}{n} = \frac{I_n^{-1}(\theta)}{n} = \frac{p(1-p)}{n}$$

$$\text{quindi } I_n^{-1}(\hat{\theta}_{MV}) = \frac{\hat{p}_{MV}(1-\hat{p}_{MV})}{n} = \frac{\bar{X}_n(1-\bar{X}_n)}{n}$$

Utilizzando $\frac{\hat{p}_{MV} - p}{\sqrt{I_n^{-1}(\hat{p}_{MV})}}$ come quantità pivotale si ottiene l'intervallo

$$\tilde{IC}_{1-\alpha} = \left[\hat{p}_{MV} \pm Z_{1-\frac{\alpha}{2}} \sqrt{I_n^{-1}(\hat{p}_{MV})} \right] \quad \text{con } 1-\alpha = 0.9 \rightarrow \alpha = 0.1$$

$$= \left[\bar{X}_n \pm 2.035 \sqrt{\frac{\bar{X}_n(1-\bar{X}_n)}{n}} \right] = \left[\frac{18}{371} \pm 1.64 \times \sqrt{\frac{\frac{18}{371}(1-\frac{18}{371})}{371}} \right] = [0.028, 0.066]$$

$$\text{iii) } \frac{L(p=0.07, \bar{X}_n)}{L(p=0.03, \bar{X}_n)} = \frac{0.07^{18} (1-0.07)^{371-18}}{0.03^{18} (1-0.03)^{371-18}} = 1.47 > 1 \quad \text{quindi } p=0.07 \text{ più verosimile di } p=0.03$$

2) $(X_1, \dots, X_n)$ con $f_X(x, \theta) = (\theta+1)x^\theta$ $x \in [0,1]$ $\theta > -1$ (Esercizio 2023)

i) Utilizzando la distribuzione congiunta di $\hat{\theta}_{MV}$ determino $\tilde{Q}(X_n, \theta)$ e $\tilde{IC}_{1-\alpha}(\bar{X}_n)$

Può darsi che $\hat{\theta}_{MV} \sim N(\theta, I_n^{-1}(\theta))$

$$l(\theta, \bar{X}_n) = \prod_{i=1}^n ((\theta+1)x_i^\theta) = (\theta+1)^n \left(\prod_{i=1}^n x_i \right)^\theta \rightarrow \text{CRITERIO FATT. FISHER } T(\bar{X}_n) = \prod_{i=1}^n x_i \text{ stat. suff. naturale}$$

$$l(\theta, \bar{X}_n) = n \log(\theta+1) + \theta \log \prod_{i=1}^n x_i$$

$$\frac{d}{d\theta} l(\theta) = \frac{n}{\theta+1} + \log \prod_{i=1}^n x_i = 0 \Rightarrow \frac{n}{\theta+1} = -\log T \Rightarrow \frac{\theta+1}{n} = -\frac{1}{\log T}$$

$$\hat{\theta}_{MV} = -\frac{n}{\log T} - 1 \quad \text{dovrebbe anche essere } \frac{d^2}{d\theta^2} l(\theta) \Big|_{\theta=\hat{\theta}_{MV}} < 0$$

$$\frac{d^2}{d\theta^2} l(\theta) = -\frac{n}{(\theta+1)^2} < 0 \quad \forall \theta \in \mathbb{R} \quad (\text{quello vale per } \theta = \hat{\theta}_{MV})$$

Calcolo $I_n(\theta) = \mathbb{E}_\theta \left[I_n^{-1}(\theta, \bar{X}_n) \right]$ (nel caso di problemi regolari di stima)

$$= \mathbb{E} \left[\frac{n}{(\theta+1)^2} \right] = \frac{n}{(\theta+1)^2} \Rightarrow c(\theta) = I_n^{-1}(\theta) = \frac{(\theta+1)^2}{n}$$

$$\text{Quindi } \hat{\theta}_{MV} \sim N\left(\theta, \frac{(\theta+1)^2}{n}\right) \quad \text{e con } \tilde{Q}(X_n, \theta) = \frac{\hat{\theta}_{MV} - \theta}{\sqrt{\frac{(\hat{\theta}_{MV}+1)^2}{n}}} = \frac{\hat{\theta}_{MV} - \theta}{(\hat{\theta}_{MV}+1)}$$

$$1-\alpha = \left[\hat{\theta}_{nv} \pm z_{1-\frac{\alpha}{2}} \sqrt{I_n^{-1}(\hat{\theta}_{nv})} \right]$$

$$= \left[-\frac{m}{\log \pi x_i} - 1 \pm z_{1-\frac{\alpha}{2}} \sqrt{\frac{\left(-\frac{m}{\log \pi x_i} - 1 + 1 \right)^2}{m}} \right] = \left[-\frac{m}{\log \pi x_i} \pm z_{1-\frac{\alpha}{2}} \sqrt{\frac{m}{\log \pi x_i}} \right]$$

(ii) Determino la funzione di verosimiglianza di $\tilde{IC}(X_1)$

$$\mathcal{L}(X_1) = 2z_{1-\frac{\alpha}{2}} \frac{(\hat{\theta}_{nv} + 1)}{\sqrt{m}} \quad \text{e} \quad \text{er}(\vartheta) = \mathbb{E}_{\vartheta}[\mathcal{L}(X_1)] = \mathbb{E}\left[2z_{1-\frac{\alpha}{2}} \frac{(\hat{\theta}_{nv} + 1)}{\sqrt{m}} \right]$$

$$= \frac{2z_{1-\frac{\alpha}{2}}}{\sqrt{m}} \left[\mathbb{E}(\hat{\theta}_{nv}) + 1 \right] = \frac{2z_{1-\frac{\alpha}{2}}}{\sqrt{m}} (\vartheta + 1)$$

utilizzo la distribuzione campionaria di $\hat{\theta}_{nv} \sim N(\vartheta, \text{er}(\vartheta))$
quindi $\mathbb{E}(\hat{\theta}_{nv}) \approx \vartheta$

3) $X_i | \vartheta \sim \text{Bern}(\vartheta)$ $i=1, \dots, m$ ind $\psi = \log \frac{\vartheta}{1-\vartheta}$

il problema $\tilde{IC}_{1-\alpha}$ per ψ :

utilizzo la quantità standard $\frac{g(\hat{\theta}_{nv}) - g(\vartheta)}{\sqrt{g'(\hat{\theta}_{nv})^2 I_n^{-1}(\hat{\theta}_{nv})}} \sim N(0, 1)$

perché $g(\hat{\theta}_{nv}) \sim N(g(\vartheta), \frac{g'(\vartheta)^2}{I_n(\vartheta)})$

so che $\hat{\theta}_{nv} = \bar{x}_n$ e che $I_n^{-1}(\vartheta) = \frac{\vartheta(1-\vartheta)}{m}$ quindi $I_n^{-1}(\hat{\theta}_{nv}) = \frac{\bar{x}_n(1-\bar{x}_n)}{m}$

calcolo all'esercizio 1

il calcolo $g'(\vartheta) = \frac{d}{d\vartheta} \left(\log \frac{\vartheta}{1-\vartheta} \right) = \frac{1-\vartheta}{\vartheta} \frac{1 \times (1-\vartheta) - \vartheta(-1)}{(1-\vartheta)^2} = \frac{1-\vartheta + \vartheta}{\vartheta(1-\vartheta)} = \frac{1}{\vartheta(1-\vartheta)}$

quindi $g'(\vartheta)^2 = \frac{1}{\vartheta^2(1-\vartheta)^2}$ e $g'(\hat{\theta}_{nv})^2 = \frac{1}{\bar{x}_n^2(1-\bar{x}_n)^2}$

quindi $g'(\hat{\theta}_{nv})^2 I_n^{-1}(\hat{\theta}_{nv}) = \frac{1}{\bar{x}_n^2(1-\bar{x}_n)^2} \frac{\bar{x}_n(1-\bar{x}_n)}{m} = \frac{1}{m\bar{x}_n(1-\bar{x}_n)}$

$$\tilde{IC}_{1-\alpha} = \left[g(\hat{\theta}_{nv}) \pm z_{1-\frac{\alpha}{2}} \sqrt{g'(\hat{\theta}_{nv})^2 I_n^{-1}(\hat{\theta}_{nv})} \right]$$

$$= \left[\ln \left[\frac{\bar{x}_n}{1-\bar{x}_n} \right] \pm z_{1-\frac{\alpha}{2}} \frac{1}{\sqrt{m\bar{x}_n(1-\bar{x}_n)}} \right]$$

4) $X_n = (X_1, \dots, X_n)$ con $f_X(x, \theta) = \frac{3}{\theta} \frac{x^2}{\theta^3} \mathbb{1}_{(0, 2\theta)}(x)$ e $\theta > 0$

i) Costruire l'intervallo di confidenza appross. utilizzando lo stimatore dei momenti $\hat{\theta}_n$

$$\begin{aligned} E[X] = \bar{X}_n &\Rightarrow E[X] = \int_0^{2\theta} \frac{3}{\theta} x \frac{x^2}{\theta^3} dx = \int_0^{2\theta} \frac{3}{\theta} \frac{x^3}{\theta^3} dx \\ &= \frac{3}{\theta \theta^3} \int_0^{2\theta} x^3 dx = \frac{3}{\theta \theta^3} \left[\frac{x^4}{4} \right]_0^{2\theta} = \frac{3}{\theta \theta^3} \frac{16\theta^4}{4} = \frac{3}{2} \theta \end{aligned}$$

Quindi $\frac{3}{2} \theta = \bar{X}_n \Rightarrow \hat{\theta}_n = \frac{2}{3} \bar{X}_n$

Definisco la distribuzione canonica di $\hat{\theta}_n$ (usando TLC):

$$\begin{aligned} \frac{(\bar{X}_n - E[X])\sqrt{n}}{\sqrt{V(X)}} &\sim N(0,1) \Rightarrow \bar{X}_n \sim N(E[X], \frac{V(X)}{n}) = N(E[\bar{X}_n], V(\bar{X}_n)) \\ &\Rightarrow a\bar{X}_n \sim N(E[a\bar{X}_n], V[a\bar{X}_n]) \end{aligned}$$

Quindi $E[\hat{\theta}_n] = E\left[\frac{2}{3} \bar{X}_n\right] = \frac{2}{3} E[X] = \frac{2}{3} \cdot \frac{3}{2} \theta = \theta$

$$V[\hat{\theta}_n] = V\left[\frac{2}{3} \bar{X}_n\right] = \frac{4}{9} \frac{V[X]}{n} \Rightarrow V[X] = E[X^2] - E[X]^2$$

$$E[X^2] = \int_0^{2\theta} \frac{3}{\theta} x^2 \frac{x^2}{\theta^3} dx = \frac{3}{\theta \theta^3} \left[\frac{x^5}{5} \right]_0^{2\theta} = \frac{3}{\theta \theta^3} \frac{32\theta^5}{5} = \frac{12}{5} \theta^2$$

$$V[X] = \frac{12}{5} \theta^2 - \left[\frac{3}{2} \theta \right]^2 = \frac{12}{5} \theta^2 - \frac{9}{4} \theta^2 = \frac{(48 - 45)\theta^2}{20} = \frac{3}{20} \theta^2 =$$

quindi $V[\hat{\theta}_n] = \frac{4}{9} \frac{V[X]}{n} = \frac{4}{9n} \frac{3}{20} \theta^2 = \frac{\theta^2}{15n}$

Quindi $\hat{\theta}_n \sim N(\theta, \theta^2/15n)$ $\Rightarrow \frac{\hat{\theta}_n - \theta}{\sqrt{\frac{\theta_n^2}{15n}}} = \frac{(\hat{\theta}_n - \theta)\sqrt{15n}}{\hat{\theta}_n} \sim N(0,1)$

Quantità casuale

Quindi $IC_{1-\alpha} = \left[\hat{\theta}_n \pm z_{1-\frac{\alpha}{2}} \frac{\hat{\theta}_n}{\sqrt{15n}} \right] = \left[\frac{2}{3} \bar{X}_n \pm z_{1-\frac{\alpha}{2}} \frac{\frac{2}{3} \bar{X}_n}{\sqrt{15n}} \right]$

ii) Determina $e_n(\theta)$ (lunghezza attesa) e m^* (minimo valore di n t.c. $P_n(\theta_0) \leq \varepsilon$) calcolato per $\varepsilon = 0.1$, $\theta_0 = 2$, $\alpha = 0.2$

$$X_n = 2Z_{1-\frac{\alpha}{2}} \frac{\frac{2}{3} \bar{X}_n}{\sqrt{15n}} = 2Z_{1-\frac{\alpha}{2}} \frac{\hat{\theta}_n}{\sqrt{15n}}$$

$$E_n(\theta) = E[|X_n|] = 2Z_{1-\frac{\alpha}{2}} \frac{E[\hat{\theta}_n]}{\sqrt{15n}} = \frac{2Z_{1-\frac{\alpha}{2}}}{\sqrt{15n}} \theta$$

$$E_n(\theta_0) < \varepsilon \Leftrightarrow \frac{2Z_{1-\frac{\alpha}{2}}}{\sqrt{15n}} \theta_0 < \varepsilon \Leftrightarrow \frac{2Z_{1-\frac{\alpha}{2}} \theta_0}{\sqrt{15} \varepsilon} < \sqrt{n} \Leftrightarrow$$

$$n > \frac{4Z_{1-\frac{\alpha}{2}}^2 \theta_0^2}{15 \varepsilon^2} = \frac{4 \cdot 2.05^2 \times 4}{15 \times 0.1^2} = \frac{4 \times 1.20 \times 4}{15 \times 0.1^2} = 136.53 \Rightarrow n^* = 137$$

↳ numero di
campioni minima che
garantisce una lunghezza
non superiore al valore
 $\varepsilon = 0.1$ fissato