Teoria statistica delle decisioni – a.a. 24-25 - **Parte 2** - 08-01-25 [**SI CORREGGONO SOLO QUESTI FOGLI**]

Sia $\mathbf{Z}_n = (X_1, \dots, X_n)$ un campione i.i.d. da Pois(θ), con $\theta > 0$ incognito. Siano inoltre d_{mv} e $d_1(\mathbf{Z}_n) = \frac{n-1}{n} \cdot d_{mv}$ due stimatori puntuali di θ, dove d_{mv} indica lo stimatore di massima verosimiglianza per θ. Si consideri la perdita quadratica $\mathbb{L}(\theta, d) = (\theta - d)^2$.

1. Calcolare valore atteso e varianza di d_{mv}.
2. Determinare $R\big(\theta, d_{mv}\big)$ e stabilire se d_{mv} sia UMVUE di θ nel modello in esame.
3. Verificare che $R(\theta, d_1) = a_n\theta + b_n\theta^2$, dove $a_n = \frac{(n-1)^2}{n^3}$ e $b_n = \frac{1}{n^2}$; studiare la consistenza di d_1.
4. Determinare per quali valori di $\theta > 0$ si ha $R\big(\theta, d_1\big) < R\big(\theta, d_{mv}\big)$, esprimendo la condizione in funzione di n.
5. Determinare $K_{va}[R(\Theta, d_1)] = r_\pi(d_1)$, supponendo che $\Theta \sim \text{Gamma}(\alpha,\ \texttt{rate} = \beta)$.
6. Determinare l'approssimazione normale per la distribuzione campionaria di $d_{mv}|\theta$ e trovare l'intervallo di confidenza asintotico $\widetilde{C}$ di livello $1 - \gamma$.
7. Determinare l'espressione di $R(\theta, \tilde{C}) = \mathbb{E}_\theta\big[\mathbb{L}(\theta, \widetilde{C})\big]$, assumendo che $\mathbb{L}(\theta, \widetilde{C})$ coincida con il quadrato della lunghezza dell'intervallo.
8. Verificare che il modello in esame ha rapporto delle verosimiglianze monotono in d_{mv}.
9. Si considerino le ipotesi $H_0 : \theta = \theta_0$ vs. $H_1 : \theta > \theta_0$. Determinare la funzione di potenza $\eta(\theta)$ del test esatto (**non asintotico**) di Karlin-Rubin ad ampiezza γ basato su d_{mv}, espressa in termini della funzione di ripartizione della statistica sufficiente per il modello.
10. Si consideri la funzione di perdita quadratica ponderata $\mathbb{L}(\theta, a) = \frac{1}{\theta}(\theta - a)^2$.

    a) Determinare l'espressione della perdita finale attesa $\rho_\pi(a; \mathbf{z}_n)$ in funzione di a e di valori attesi a posteriori di opportune funzioni di Θ, considerando una distribuzione a priori generica.
    b) Determinare l'espressione dell'azione $a^\star$ che minimizza la funzione $\rho_\pi(a; \mathbf{z}_n)$ trovata al punto *a*).
    c) Determinare l'espressione di $a^\star$ che si ottiene assumendo che $\Theta \sim \text{Gamma}(\alpha, \beta)$[1].

[1]Ricordare che, se $X \sim \text{InvGa}(\alpha, \beta)$, allora $\mathbb{E}[X] = \beta/(\alpha - 1),\ \alpha > 1$.

Soluzione

(1)

Semplici calcoli conducono a $d_{mv} = \frac{1}{n}\sum_{i=1}^n X_i$. Ricordando che $\sum_{i=1}^n X_i \sim \text{Pois}(n\theta)$, si ha $\mathbb{E}_\theta(d_{mv}) = \frac{1}{n} n\theta = \theta$ e $\mathbb{V}_\theta(d_{mv}) = \frac{1}{n^2} n\theta = \frac{\theta}{n}$.

(2)

Poiché d_{mv} è uno stimatore corretto di θ nel modello in esame, la funzione di rischio $R\big(\theta, d_{mv}\big)$ è pari a $\mathbb{V}_\theta(d_{mv}) = \frac{\theta}{n}$. Inoltre, essendo d_{mv} funzione di statistica sufficiente e completa per il modello, fatto garantito dall'appartenenza della distribuzione di Poisson alla famiglia esponenziale, allora tale stimatore sarà anche UMVUE per il Teorema di Lehmann–Scheffé.

(3)

Notiamo che $d_1(\mathbf{Z}_n)$ è uno stimatore distorto con valore atteso $\mathbb{E}_\theta(d_1) = \frac{n-1}{n}\theta$. Pertanto, data la funzione di perdita quadratica, $R(\theta, d_1) = \mathbb{V}_\theta(d_1) + B_\theta^2(d_1)$, dove

$$\mathbb{V}_\theta(d_1) = \left(\frac{n-1}{n}\right)^2 \cdot \mathbb{V}_\theta(d_{mv}) = \left(\frac{n-1}{n}\right)^2 \frac{\theta}{n} = \frac{(n-1)^2}{n^3}\theta = a_n\theta$$

$$B_\theta^2(d_1) = \left(\theta - \frac{n-1}{n}\theta\right)^2 = \frac{\theta^2}{n^2} = b_n\theta^2$$

Infine, è immediato verificare che $\lim_{n\to\infty} R(\theta, d_1) = 0$.

(4)

$R(\theta, d_1) < R(\theta, d_{mv}) \iff \frac{(n-1)^2}{n^3}\theta + \frac{\theta^2}{n^2} < \frac{\theta}{n} \iff \frac{(n-1)^2}{n^3}\theta + \frac{\theta^2}{n^2} - \frac{\theta}{n} < 0 \iff \cdots \iff 0 < \theta < \frac{2n-1}{n}$

(5)

Ricordiamo anzitutto che, se $\Theta \sim \text{Gamma}(\alpha,\ \texttt{rate} = \beta)$, allora $\mathbb{E}[\Theta] = \alpha/\beta$, $\mathbb{V}[\Theta] = \alpha/\beta^2$ e $\mathbb{E}[\Theta^2] = \mathbb{V}[\Theta] + \mathbb{E}[\Theta]^2 = \alpha(\alpha+1)/\beta^2$. Dunque $r_\pi(d_1) = \mathbb{E}[R(\Theta, d_1)]$ sarà pari a

$$r_\pi(d_1) = \frac{(n-1)^2}{n^3}\mathbb{E}[\Theta] + \frac{1}{n^2}\mathbb{E}\big[\Theta^2\big] = \frac{(n-1)^2}{n^3}\frac{\alpha}{\beta} + \frac{1}{n^2}\frac{\alpha(\alpha+1)}{\beta^2}$$

(6)

La varianza di $d_{mv}|\theta$ coincide con il limite inferiore di Cramer-Rao, pari all' inverso dell' informazione attesa di Fisher. Pertanto, dato che $d_{mv} \ \dot\sim\ \mathcal{N}\big(\theta,\ I_n(\theta)^{-1}\big)$, si avrà varianza asintotica pari a $\frac{\theta}{n}$. Da ciò segue che $\tilde{C} = d_{mv} \ \pm\ z_{1-\gamma/2}\sqrt{\frac{d_{mv}}{n}}$ con lunghezza $\mathcal{L}_\alpha = 2z_{1-\gamma/2}\sqrt{\frac{d_{mv}}{n}}$.

(7)

Con riferimento al punto precedente, $\mathcal{L}_\gamma^2 = 4z_{1-\gamma/2}^2\frac{d_{mv}}{n}$ e pertanto

$$R(\theta, \tilde{C}) = 4z_{1-\gamma/2}^2\frac{\mathbb{E}[d_{mv}]}{n} = 4z_{1-\gamma/2}^2\frac{\theta}{n}$$

(8)

Indicando per semplicità con $L(\theta)$ la funzione di verosimiglianza per il modello in esame valutata in θ, si avrà

$$\frac{L(\theta_1)}{L(\theta_0)} = \left(\frac{\theta_1}{\theta_0}\right)^{S_n} \exp\big\{-n\big(\theta_1 - \theta_0\big)\big\} = \left(\frac{\theta_1}{\theta_0}\right)^{nd_{mv}} \exp\big\{-n\big(\theta_1 - \theta_0\big)\big\}$$

che è monotono in d_{mv}. In particolare, crescente per $\theta_1 > \theta_0$ e viceversa se $\theta_1 < \theta_0$.

(9)

Per il teorema di Karlin-Rubin, si rifiuta H_0 se la somma campionaria S_n è maggiore di una soglia k o, alternativamente, se $d_{mv} = \frac{S_n}{n} > k' = \frac{k}{n}$. Da ciò segue che

$$\eta(\theta) = \mathbb{P}_\theta\big(S_n > k\big) = 1 - \mathbb{F}_\theta\big(k\big)$$

dove $\mathbb{F}_\theta(\cdot)$ indica la funzione di ripartizione della distribuzione di Poisson con parametro $n\theta$. Si ottiene la funzione di potenza per il test di ampiezza γ ponendo $k = q_{1-\gamma}$, con $q_{1-\gamma}$ quantile di livello $1-\gamma$ della v.a. Poisson di parametro $n\theta_0$. La risposta al quesito è quindi

$$\eta_\gamma(\theta) = 1 - \mathbb{F}_\theta\big(q_{1-\gamma}\big).$$

(10)

(a) Riguardo al primo punto, sviluppando il quadrato e riscalandolo per θ^{-1}, si ha

$$\rho_\pi(a; \mathbf{z}_n) = \mathbb{E}\big[\Theta|z_n\big] - 2a + a^2\mathbb{E}\left[\frac{1}{\Theta}\middle|z_n\right]$$

(b) L'azione ottima $a^\star$ si ottiene derivando tale quantità rispetto alla generica a, ovvero calcolando la derivata $\frac{\partial\rho}{\partial a}$, imponendola pari a 0 e verificando che il valore ottenuto sia un punto di minimo globale. Semplici passaggi restituiscono

$$\frac{\partial\rho}{\partial a} = 0 \iff a = \mathbb{E}\left[\frac{1}{\Theta}\middle|z_n\right]^{-1}$$

che coincide con un punto di minimo.

(c) Assumendo infine che $\Theta \sim \text{Gamma}(\alpha, \beta)$, segue che $\Theta^{-1} \sim \text{InvGa}(\alpha, \beta)$ e dunque $a^\star = (\alpha - 1)/\beta$ in virtù della nota a pie' di pagina.