Teoria statistica delle decisioni – a.a. 23-24 - Esonero 1 - 24-11-23 **[SI CORREGGONO SOLO QUESTI FOGLI]**

Cognome, nome e n. di matricola: SOLUZIONI

1. Siano $X_1, \dots, X_n|\theta$ un campione i.i.d. con funzione di massa di probabilità $p_\theta^X(x) = (1-\theta)^{x-1}\theta$, $x = 1, 2, \dots$ e $\theta \in (0,1)$. Sia $\mathbf{z}_n$ un generico campione osservato. Determinare:
   (a) la funzione di verosimiglianza $\ell(\theta; \mathbf{z}_n)$;
   (b) l'espressione di $d_{mv}(\mathbf{z}_n)$, stima di massima verosimiglianza di θ.

(a) $\ell(\theta) = (1-\theta)^{S_n - n}\,\theta^n, \quad \theta \in (0,1) \qquad S_n = \sum_1^n x_i$

(b) $\ln \ell(\theta) = (S_n - n)\ln(1-\theta) + n \ln\theta$

$\frac{d}{d\theta}[\,''\,] = -\frac{S_n - n}{1-\theta} + \frac{n}{\theta} = 0 \iff -(S_n - n)\theta + n - n\theta = 0$

$\iff \cancel{n\theta} - S_n\theta + n - \cancel{n\theta} = 0 \qquad \theta = n/S_n = 1/\bar{x}_n$

$\frac{d^2}{d\theta}[\,''\,] < 0 \quad \forall\theta \Rightarrow \quad \boxed{d_{mv} = n/S_n}$

2. Con riferimento al precedente esercizio, assumere che $\Theta \sim \text{Beta}(\alpha, \beta)$.
   (a) Determinare $\pi(\theta|\mathbf{z}_n)$ e verificare che il modello Beta è coniugato al modello delle v.a. X_i.
   (b) Determinare l'espressione di $d_B(\mathbf{z}_n) = \mathbb{E}[\Theta|\mathbf{z}_n]$ (valore atteso a posteriori di Θ).

(a) $\pi(\theta|z_n) = c \cdot \ell(\theta)\,\pi(\theta|z_n) = c \cdot \theta^n (1-\theta)^{S_n - n} * \theta^{\alpha-1}(1-\theta)^{\beta-1}$

$= c \cdot \theta^{(\alpha+n)-1}(1-\theta)^{(\beta + S_n - n)-1}$ per $\Theta|z_n \sim \text{Beta}(\bar\alpha, \bar\beta)$

con

(b) $\begin{cases} \bar\alpha = \alpha + n \\ \bar\beta = \beta + S_n - n \end{cases} \Rightarrow \quad d_B = \mathbb{E}[\Theta|z_n] = \frac{\bar\alpha}{\bar\alpha + \bar\beta}$

$\boxed{d_B = \frac{\alpha + n}{\alpha + \beta + S_n}}$

3. Con riferimento ai precedenti due esercizi:
   (a) determinare i valori di α e β tali che $d_{mv}(\mathbf{z}_n) = d_B(\mathbf{z}_n)$;
   (b) dire se la distribuzione a priori risultante è ancora una funzione densità su $\Omega = [0, 1]$.

(a) $d_B = d_{mv} \iff \frac{n}{S_n} = \frac{\alpha + n}{\alpha + \beta + S_n} \iff \alpha = \beta = 0$

(b) $\alpha = \beta = 0 \Rightarrow \pi(\theta) = \frac{1}{\theta(1-\theta)} \quad \theta \in [0,1)$

che è IMPROPRIA in $(0,1)$ $\Rightarrow$ no densità su Ω

RICORDA: la v.a. Beta(α, β) richiede $\boxed{\alpha, \beta > 0}$

4. Come sopra, sia $X|\theta \sim p_\theta^X(x) = (1-\theta)^{x-1}\theta$, $x = 1, 2, \ldots$ e $\theta \in (0,1)$ e $\Theta \sim \text{Beta}(\alpha, \beta)$ **ma assumere ora** $n = 1$. Calcolare $\mathbb{E}(X)$, il valore atteso rispetto alla distribuzione **marginale** di X, sapendo che $\mathbb{E}(X|\theta) = 1/\theta$.
[**Sugg.**: ricordare che $\Gamma(\alpha) = (\alpha-1)\Gamma(\alpha-1)$ e che, in generale, $\mathbb{E}(1/Y) \neq 1/\mathbb{E}(Y)$].

$\boxed{n=1}$

$$\mathbb{E}_X[X] = \mathbb{E}_\Theta\left[\mathbb{E}_{X|\theta}(X)\right] = \mathbb{E}_\Theta\left[1/\Theta\right]$$

$$= \int_0^1 \frac{1}{\theta}\frac{1}{B(\alpha,\beta)}\theta^{\alpha-1}(1-\theta)^{\beta-1}d\theta = \frac{1}{B(\alpha,\beta)}\int_0^1 \theta^{(\alpha-1)-1}(1-\theta)^{\beta-1}d\theta$$

$$= \frac{1}{B(\alpha,\beta)} * B(\alpha-1,\beta) = \frac{\Gamma(\alpha-1)\cancel{\Gamma(\beta)}}{\Gamma(\alpha+\beta-1)} \cdot \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\cancel{\Gamma(\beta)}} = \boxed{\frac{(\alpha+\beta-1)}{(\alpha-1)}}$$

5. Sia $X|\theta \sim p_\theta^X(x) = \theta x^{\theta-1}$, $x \in (0,1)$ e $\theta > 0$. Considerare: $H_0 : \theta = \theta_0$ e $H_1 : \theta = \theta_1$, con $\theta_0 > \theta_1 > 0$.
(a) Determinare l'espressione di $B_{01}(x)$, fattore di Bayes per il confronto tra le ipotesi in esame basato su una singola osservazione x.
(b) Verificare che $B_{01}(x) > 1 \iff x > A$ e determinare l'espressione di A (in funzione di θ_0 e θ_1).

(a) $B_{01}(x) = \frac{\theta_0}{\theta_1} x^{\theta_0-\theta_1}$ $\qquad \theta_0 > \theta_1$

(b) $B_{01}(x) > 1 \iff x^{\theta_0-\theta_1} > \frac{\theta_1}{\theta_0} \iff \boxed{x > A}$

con $\boxed{A = \left(\frac{\theta_1}{\theta_0}\right)^{\frac{1}{\theta_0-\theta_1}}}$ $\qquad$ ok perchè $\theta_0 - \theta_1 > 0$

6. Con riferimento al precedente esercizio, sfruttando la relazione del punto (b):
(a) determinare la probabilità $\mathbb{P}_{\theta_0}[B_{01}(X) > 1]$ [**Sugg.**: usare la distribuzione di $X|\theta_0$];
(b) calcolare il valore numerico che si ottiene ponendo $\theta_0 = 2$ e $\theta_1 = 1$.

(a) $$\mathbb{P}_{\theta_0}[B_{01}(X) > 1] = \mathbb{P}_{\theta_0}[X > A] = \int_A^1 \theta_0 x^{\theta_0-1}dx$$

$$= \theta_0 \left.\frac{x^{\theta_0}}{\theta_0}\right|_A^1 = \left. x^{\theta_0}\right|_A^1 = 1 - A^{\theta_0}$$

(b) $\theta_0 = 2$, $\theta_1 = 1 \Rightarrow A = \frac{1}{2} \Rightarrow 1 - A^{\theta_0} = 1 - \frac{1}{4} = 3/4$

$\Rightarrow \boxed{\mathbb{P}_{\theta_0}[B_{01}(X) > 1] = 3/4}$

7. Siano $X_1, \dots, X_n$ le risposte quantitative (aleatorie) a un trattamento medico somministrato a n pazienti e sia θ l'*effetto del trattamento* medio nella popolazione da cui provengono i dati. Supporre che, condizionatamente alla conoscenza di θ, le risposte al trattamento siano v.a. indipendenti con distribuzione **normale** centrata sul valore vero dell'effetto del trattamento e varianza unitaria. Si dispone di informazione pre-sperimentale e sperimentale sull'*effetto del trattamento*.
Informazione sperimentale: su 10 osservazioni sperimentali, la somma delle risposte al trattamento è uguale a 20 (in opportuna unità di misura).
Informazione pre-sperimentale: basandoci su dati storici è e lecito supporre che il vero valore del trattamento θ abbia distribuzione **normale** con valore atteso pari a 1/2 e varianza pari a 1/4.
Quesito. Formalizzare opportunamente il problema per mezzo di un modello bayesiano, ovvero: indicare quali sono le distribuzioni di $X_i|\theta$ e di Θ da utilizzare e i valori da associare ai parametri noti e agli e iperparametri dei modelli considerati.

$$\begin{cases} X_i|\theta \sim N(\theta, \sigma^2) \text{ iid } i=1\cdots n \\ \Theta \sim N(\mu_0, \sigma^2/n_0) \end{cases} \text{ con } \begin{cases} n = 10 \\ \bar{x}_n = 20/10 = 2 \\ \sigma^2 = 1 \\ n_0 = 4 \\ \mu_0 = 1/2 \end{cases}$$

- n_0: da $\frac{\sigma^2}{n_0} = \frac{1}{4} \Leftrightarrow \frac{1}{n_0} = \frac{1}{4}$
- $\bar{x}_n = \frac{\sum^n x_i}{n} = \frac{20}{10} = 2 \Leftrightarrow \boxed{n_0 = 4}$

quindi $\boxed{\begin{matrix} X_i|\theta \sim N(\theta, 1) \\ \Theta \sim N(1/2, 1/4) \end{matrix}}$

8. Con riferimento al precedente esercizio: scrivere le espressioni generiche (formule) e specifiche (sostituendo i valori numerici) della distribuzione a posteriori di Θ.

Da (7) quindi

$$\Theta | \underline{z}_n \sim N(\bar{\mu}_p, \bar{\sigma}_p^2) \text{ con}$$

$$\bar{\mu}_p = \frac{n_0 \mu_0 + n \bar{x}_n}{n_0 + n} = \frac{4(1/2) + 10(2)}{4 + 10} = \frac{2 + 20}{14} = \frac{11}{7} = 1.57$$

$$\bar{\sigma}_p^2 = \sigma^2/(n_0+n) = \frac{1}{14} \text{ quindi } \boxed{\Theta | \underline{z}_n \sim N\left(\frac{11}{7}, \frac{1}{14}\right)}$$

9. Con riferimento ai precedenti due esercizi, fornire una stima puntuale e una intervallare di θ basata sulla distribuzione a posteriori di Θ [riportare espressioni generiche (formule) e specifiche (sostituendo i valori numerici)]. Per la stima intervallare utilizzare un livello di credibilità $\gamma = 0.95$ e ricordare che $z_{0.975} = 1.96$.

$$\hat{\theta}_B = \bar{\mu}_p = 11/7 = 1.57$$

$$C_{0.95} = \bar{\mu}_p \pm \bar{\sigma}_p z_{1-\frac{\gamma}{2}} = \frac{11}{7} \pm (1.96)\frac{1}{\sqrt{14}}$$

$$= (1.04, 2.09)$$

$1-4p \geq 0 \quad 4p \leq 1 \quad p \leq 1/4$
$1-4p \leq 1 \quad 4p \geq 0 \quad p \geq 0$

10. Sia $\Omega = \{\omega_1, \omega_2, \omega_3\}$. Per due decisioni δ_1 e δ_2 si considerino le perdite e la distribuzione $p(\omega)$ seguenti:

	W_{δ_1}	W_{δ_2}	$p(\omega)$
ω_1	0	2	p
ω_2	2	1	$3p$
ω_3	1	0	$1-4p$

(a) Determinare l'insieme di valori che p può assumere affinché $p(\omega)$ sia una funzione di massa di probabilità su Ω.

(b) Determinare, se esiste, il valore di p tale da rendere δ_1 e δ_2 equivalenti rispetto al criterio del valore atteso.

(a) da $0 \leq 1-4p \leq 1 \Rightarrow \begin{cases} 1-4p \geq 0 & p \leq 1/4 \\ p \geq 0 \end{cases} \Rightarrow p \in [0, 1/4]$

(b) $K_{VA}[W\delta_1] = 2(3p) + 1(1-4p) = 6p + 1 - 4p = 1+2p$

$K_{VA}[W\delta_2] = 2(p) + 1(3p) = 5p$

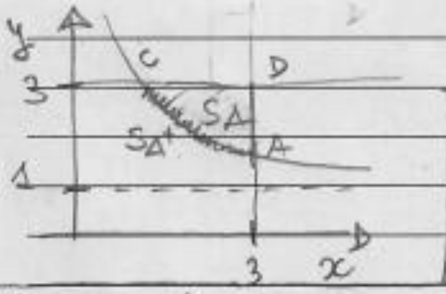
quindi $1+2p = 4p \iff p = 1/3$ ma $p \in [0, 1/4] \Rightarrow \nexists\, p$ cercato

11. Sia S_Δ la porzione del primo quadrante cartesiano limitato inferiormente dalla curva γ_1 di equazione $y = \frac{2}{x} + 1$, $x > 0$ e dalle rette $x = 3$, $y = 3$.

(a) Rappresentare graficamente: S_Δ e S_{Δ^+}.

(b) Determinare le coordinate della decisione ottima per il criterio del valore atteso che si ottiene assumendo per gli stati di natura le probabilità $p_1 = p_2 = 1/2$.

(a)



S_Δ = delimit da CD, AD, CA

S_Δ^+ = arco CA

(b) $p_1 = p_2 = 1/2 \Rightarrow \boxed{y = -x + K}$ $K \in \mathbb{R}$ FASCIO IMPR.

Intersezione FASCIO con γ_1 quindi

$\begin{cases} y = 2/x + 1 \\ y = -x + K \end{cases} \iff 2/x + 1 = -x + K \iff 2 + x = -x^2 + Kx \iff x^2 + (1-K)x + 2 = 0$

Condiz di TANGENZA: $(1-K)^2 - 8 = 0 \iff K^2 - 2K - 7 = 0$

$\iff K = 1 \pm \sqrt{1+7} = 1 \pm \sqrt{8} = 1 \pm 2\sqrt{2} \Rightarrow \boxed{K = 1 + \sqrt{2}}$

$\Rightarrow r_T : \boxed{y = -x + 1 + 2\sqrt{2}}$; p.to tg è (x_T, y_T) con $\boxed{x_T = \sqrt{2} \quad y_T = 2/\sqrt{2} + 1 = \sqrt{2}+1}$

12. Con riferimento al precedente esercizio, sia A il punto di intersezione tra la curva γ_1 e la retta $x = 3$.

(a) Determinare l'equazione della retta del fascio improprio passante per il punto-decisione A e rappresentare graficamente il luogo dei punti di S_Δ equivalenti alla decisione A.

(b) Calcolare il criterio del valore atteso della decisione A.

(a) $A = r \cap \gamma_1$ quindi $\begin{cases} y = 2/x + 1 \\ x = 3 \end{cases} \iff \begin{cases} x = 3 \\ y = \frac{2}{3} + 1 \end{cases} = \boxed{\begin{cases} x = 3 \\ y = 5/3 \end{cases} = A}$

retta di $y = -x + K$ pass per $(3, 5/3)$

$5/3 = -3 + K \Rightarrow K = 3 + 5/3 = 14/3$

$\Rightarrow \boxed{y = -x + 14/3}$

(b) K_{VA} in $\left(3, \frac{5}{3}\right)$: $3\left(\frac{1}{2}\right) + \frac{5}{3}\left(\frac{1}{2}\right) = \frac{3}{2} + \frac{5}{6} = \frac{9+5}{6} = \frac{14}{6} = \boxed{7/3}$