

Πρόταση

Με βάση το ΜΤΕ, δεν υπάρχει βέλτιστος ευακμντικός της μη σταθερής παραμετρικής συνάρτησης $g(\theta)$.

Απόδ.

Εστω ότι υπάρχει βέλτιστος ευακμντικός της μη σταθερής συνάρτησης $g(\theta)$ και εστω αυτός T^* .

Αφού, $g(\theta)$ μη σταθερή $\Rightarrow \exists \theta_1 \neq \theta_2 : g(\theta_1) \neq g(\theta_2)$

Ο T^* βέλτιστος με το Μ.Τ.Ε, δηλαδή:

$$E(T^* - g(\theta))^2 \leq E(T - g(\theta))^2, \quad \forall \theta \in \Theta \text{ και } \forall T \quad (1)$$

Εφόσον, η (1) ισχύει $\forall T$ τότε θα ισχύει και για $T_1 = g(\theta_1)$.

$$\text{Αρα, } E(T^* - g(\theta))^2 \leq E(g(\theta_1) - g(\theta))^2, \quad \forall \theta \quad (2)$$

Εφόσον, η (2) ισχύει $\forall \theta$ τότε θα ισχύει και για $\theta = \theta_1$.

$$\text{Αρα, } E(T^* - g(\theta))^2 \leq E(g(\theta_1) - g(\theta_1))^2 = 0$$

Γνωρίζουμε όμως, Y τη μη αμετάβλητη τότε

$$E(Y) = 0 \Rightarrow P(Y=0) = 1. \text{ Αρα, εδω } P(T^* = g(\theta_1)) = 1$$

Παρόμοια ισχύει για την $T_2 = g(\theta_2)$ ή

$$E(T^* - g(\theta_1))^2 \leq E(g(\theta_2) - g(\theta_1))^2$$

Για $\theta = \theta_2$ έχουμε:

$$E(T^* - g(\theta_1))^2 \leq E(g(\theta_2) - g(\theta_2))^2 = 0$$

$$\text{Τότε } P(T^* = g(\theta_2)) = 1$$

Τότε,

$$P(T^* = g(\theta_1)) = P(T^* = g(\theta_2)) = 1$$

$$\text{Άρα, } g(\theta_1) = g(\theta_2) \text{ (άριστο)}$$

Ορισμός (Minimax εκτίμηση)

Ελαχιστομεγιστος ευακμής είναι εκ νέος ο ευακμής που ελαχιστοποιεί τη μέγιστη $J_n(\theta)$.

$$\min_d \max_{\theta \in \Theta} R(\theta, d), \text{ όπου } R(\theta, d): J_n(\theta).$$

Παραδείγματα:

1) Έστω $x_1, x_2, \dots, x_n \sim N(\mu, \sigma^2)$ είναι ένα τ.δ.

σ^2 : αγνωστο και μ : γνωστό

Να βρεθεί ο βέλτιστος ευακμής του σ^2 στην οικογένεια των ευακμής της μορφής

$$d(x) = b \cdot S^2, \quad b > 0 \text{ εαν έχουμε:}$$

i) συνάρτηση απώλειας = Τ.Σ.

$$ii) \quad L(\theta, d) = \frac{d}{\theta} - 1 - \log \frac{d}{\theta}$$

ΛΥΣΗ

$$i) \quad E\sigma^2 = \theta$$

$$\text{Ορίσω να } b\sigma^2 \text{ το } R(\theta, b\sigma^2).$$

$$\begin{aligned} R(\theta, b\sigma^2) &= E(b\sigma^2 - \theta)^2 = \text{Var}(b\sigma^2) + (E(b\sigma^2) - \theta)^2 \\ &= b^2 \text{Var}(\sigma^2) + (b E(\sigma^2) - \theta)^2. \quad (1) \end{aligned}$$

$$\text{Οπότε, } X_1, \dots, X_n \sim N(\mu, \sigma^2 = \theta) \text{ τότε}$$

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi_{n-1}^2 \equiv \Gamma\left(\frac{n-1}{2}, 2\right)$$

$$E\left(\frac{(n-1)S^2}{\sigma^2}\right) = (n-1) \Rightarrow E(S^2) = \sigma^2$$

$$\text{Var}\left(\frac{(n-1)S^2}{\sigma^2}\right) = 2(n-1) \Rightarrow \text{Var}(S^2) = \frac{2\sigma^4}{n-1}$$

Αρα, (1) είναι:

$$\begin{aligned} R(\theta, b\sigma^2) &= b^2 \cdot \frac{2\theta^2}{n-1} + (b\theta - \theta)^2 = \\ &= \left\{ \frac{2b^2}{n-1} + (b-1)^2 \right\} \theta^2. \end{aligned}$$

Βέλτιστος, ευχρηστος ευτιμος που ελαχιστοποιεί
τη συνάρτηση: $f(b) = \frac{2b^2}{n-1} + (b-1)^2$

$$f'(b) = 0 \Rightarrow b^* = \frac{n-1}{n+1}$$

$$f''(b) > 0.$$

$$\text{Αρα, } d^* = b^* \cdot S^2 = \frac{n-1}{n+1} \cdot S^2 = \frac{\sum (X_i - \bar{X})^2}{n+1}$$

Είναι ο βέλτιστος ευχρηστος της $d(x)$.

$$\begin{aligned}
 \text{ii)} \quad R(\theta, bS^2) &= E\left(\frac{bS^2}{\theta} - 1 - \log \frac{bS^2}{\theta}\right) = \\
 &= E\left(\frac{bS^2}{\theta}\right) - 1 - E\left(\log \frac{bS^2}{\theta}\right) = \\
 &= \frac{b}{\theta} \cdot \theta - 1 - E(\log b + \log \frac{S^2}{\theta}) = \\
 &= b - 1 - \log b - E(\log \frac{S^2}{\theta}) = \\
 &= b - 1 - \log b + \log \theta - E(\log S^2)
 \end{aligned}$$

Ελαχιστοποιώ τη συνάρτηση

$$f(b) = b - 1 - \log b + c, \text{ όπου } c = \log \theta - E(\log S^2)$$

$$f'(b) = 1 - \frac{1}{b} \rightarrow f'(1) = 0$$

$$f''(b) = +\frac{1}{b^2} > 0$$

Άρα, το σημείο $b^* = 1$ το ελάχιστο

Άρα, $b^* = b^* S^2 = S^2$ ο βέλτιστος επιμετωρ της $d(x)$.

2) Έστω $X_1, X_2, \dots, X_n \sim N(\mu, \sigma^2)$, σ^2 άγνωστο και συνάρτηση απώλειας το τετραγωνικό σφάλμα. Ποια από τις $d_1 = \frac{1}{n} \sum_{i=1}^n (X_i - \mu)^2$ και $d_2 = k \cdot d_1$ με $k > 0$ είναι η καλύτερη απώλωση; ΛΥΣΗ

Εξετάζουμε ποιο από τα $R(\theta, d_1)$ και $R(\theta, d_2)$ είναι μικρότερο από το άλλο

$$R(\theta, d_1) = \text{Var } d_1 + (E d_1 - \theta)^2$$

$$\frac{X_i - \mu}{\sigma} \sim N(0, 1) \rightarrow \frac{(X_i - \mu)^2}{\sigma^2} \sim \chi_1^2 \rightarrow$$

$$\rightarrow \sum_{i=1}^n \frac{(X_i - \mu)^2}{\sigma^2} \sim \chi_{\Sigma}^2 = \chi_n^2 \equiv \Gamma_{\text{αριθ}}\left(\frac{n}{2}, 2\right)$$

Επομένως,

$$E \left[\frac{\sum (x_i - \mu)^2}{\sigma^2} \right] = n, \quad \text{Var} \left[\frac{\sum (x_i - \mu)^2}{\sigma^2} \right] = 2n$$

Ετσι, λοιπόν:

$$\begin{aligned} R(\theta, d_1) &= \frac{1}{n^2} \text{Var} \left(\sum (x_i - \mu)^2 \right) + \left(\frac{1}{n} E \left(\sum (x_i - \mu)^2 \right) - \theta \right)^2 = \\ &= \frac{1}{n^2} \cdot 2n \theta^2 + \left[\frac{1}{n} \cdot n \theta - \theta \right]^2 = \frac{2\theta^2}{n} \end{aligned}$$

To $d_2 = k \cdot d_1 = k \cdot \frac{\sum (x_i - \mu)^2}{n}$

Τότε,

$$\begin{aligned} R(\theta, d_2) &= \text{Var} \left(\frac{k}{n} \cdot \sum (x_i - \mu)^2 \right) + \left(E \left(\frac{k}{n} \cdot \sum (x_i - \mu)^2 \right) - \theta \right)^2 = \\ &= \frac{k^2}{n^2} \cdot 2n \theta^2 + \left(\frac{k}{n} \cdot n \theta - \theta \right)^2 = \\ &= \frac{2k^2}{n} \theta^2 + (k-1)^2 \theta^2 = \left(\frac{2k^2}{n} + (k-1)^2 \right) \theta^2 \end{aligned}$$

Αν πάρουμε τη διαφορά

$$R(\theta, d_2) - R(\theta, d_1) = \left(\frac{2k^2}{n} + (k-1)^2 - \frac{2}{n} \right) \theta^2$$

Οετω $g(k) = \frac{2k^2}{n} + (k-1)^2 - \frac{2}{n}$

$$\Delta = 16 > 0$$

$$k_{1,2} = \frac{1}{2} \left(\frac{n-2}{n+2} \right)$$

$$\left(\frac{n-2}{n+2}, 1 \right) \rightarrow F(k) < 0, \quad d_2 \text{ καλύτερα}$$

$$\text{Εξως των ριζών} \rightarrow F(k) > 0, \quad d_1 \text{ καλύτερα}$$

- 3) Αν d είναι μια απόφαση για το θ με σταθερή σφάλματος κινδύνου και $d^* = \lambda_1 d + \lambda_2$, $\lambda_i, i=1,2$ σταθ. νδσ $R(d^*, \lambda_1 \theta + \lambda_2)$ σταθερή.
(Συνάρτηση απώλειας το τετραγωνικό σφάλμα).

ΛΥΣΗ

$$\begin{aligned} R(d^*, \lambda_1 \theta + \lambda_2) &= E(d^* - \lambda_1 \theta - \lambda_2)^2 = \\ &= E(\lambda_1 d + \lambda_2 - \lambda_1 \theta - \lambda_2)^2 = E(\lambda_1 (d - \theta))^2 = \\ &= \lambda_1^2 \cdot E(d - \theta)^2 = \underbrace{\lambda_1^2}_{\text{σταθ.}} \underbrace{R(\theta, d)}_{\text{σταθ.}} = \text{σταθ.} \end{aligned}$$

- 4) Έστω $x_1, x_2, \dots, x_n$ τδ από την $N(\mu, 1)$
i) $R(\mu, \bar{x}) = ?$ (συνάρτ. απώλ. το τετρ. σφάλμα)
ii) αν $x_1, x_2, \dots, x_{2n}$ τδ από την $N(\mu, 1)$
νδσ $T = \frac{1}{n} \sum_{i=1}^n x_i$ δεν είναι παραδεκτός

ΛΥΣΗ

$$i) R(\mu, \bar{x}) = V_{\mu, \bar{x}} + (E \bar{x} - \mu)^2 = \frac{1}{n} + 0 = \frac{1}{n}$$

- ii) Ανάπτυξη εκφρασεων εκτύπωσης του T .

$$R(\mu, \frac{\sum x_i}{n}) = \frac{1}{n}$$

$$R(\mu, \frac{\sum x_i}{2n}) = \frac{1}{2n}$$

- 5) Έστω $X \sim \mu$ με $B(1, \theta)$
 $P(X=x, \theta) = \theta^x \cdot (1-\theta)^{1-x}$, $x=0,1$
 $\theta \in \Theta = \{ \overset{\theta_1}{\frac{1}{4}}, \overset{\theta_2}{\frac{1}{2}} \}$
Χώρος αποφάσεων $D = \{a_1, a_2\}$
Συνάρτηση απώλειας κάποια ωσφ.

$\theta \backslash d$	a_1	a_2
θ_1	1	4
θ_2	3	2

Να βρεθεί ο ελάχιστος κερδιστός
Ευρήματα του $R(\theta, d)$.

ΛΥΣΗ

ΣΥΝΟΛΟ ΔΥΝΑΤΩΝ ΑΠΟΦΑΣΕΩΝ

$x \backslash d$	d_1	d_2	d_3	d_4
$x=0$	a_1	a_1	a_2	a_2
$x=1$	a_1	a_2	a_1	a_2

$$R(\theta, d) = E(L(\theta, d(x)))$$

$\theta \in \{\theta_1, \theta_2\}$ } 8 ενδογενείς
 $d \in \{d_1, \dots, d_4\}$

$$\begin{aligned}
 \bullet R(\theta_1, d_1) &= E(L(\theta_1, d_1(x))) = \\
 &= \sum_x L(\theta_1, d_1(x)) \cdot P(X=x | \theta=\theta_1) \\
 &= L(\theta_1, d_1(0)) \cdot P(X=0 | \theta=\theta_1) + L(\theta_1, d_1(1)) \cdot P(X=1 | \theta=\theta_1) \\
 &= L(\theta_1, a_1) \cdot P(X=0 | \theta_1) + L(\theta_1, a_1) \cdot P(X=1 | \theta_1) \\
 &= 1 - \theta_1 + \theta_1 = 1 \quad \frac{3}{4} + \frac{1}{4} = 1
 \end{aligned}$$

$$\begin{aligned}
 \bullet R(\theta_1, d_2) &= E(L(\theta_1, d_2(x))) = \\
 &= \sum_x L(\theta_1, d_2(x)) \cdot P(X=x | \theta=\theta_1) \\
 &= L(\theta_1, a_1) \cdot P(X=0 | \theta_1) + L(\theta_1, a_2) \cdot P(X=1 | \theta_1) \\
 &= 1 - \theta_1 + 4 \cdot \theta_1 = 1 \cdot \frac{3}{4} + 4 \cdot \frac{1}{4} = \frac{7}{4}
 \end{aligned}$$

$$\bullet R(\theta_1, d_3) = 4 \cdot \frac{3}{4} + 1 \cdot \frac{1}{4} = \frac{13}{4}$$

$$\bullet R(\theta_1, d_4) = 4 \cdot \frac{3}{4} + 4 \cdot \frac{1}{4} = 4$$

Παραπομπές,

$$\begin{aligned}
 \bullet R(\theta_2, d_1) &= L(\theta_2, a_1) \cdot P(X=0 | \theta_2) + L(\theta_2, a_1) \cdot P(X=1 | \theta_2) \\
 &= 3 \cdot \frac{1}{2} + 3 \cdot \frac{1}{2} = 3
 \end{aligned}$$

$$\bullet R(\theta_2, d_2) = 3 \cdot \frac{1}{2} + 2 \cdot \frac{1}{2} = \frac{3}{2} + 1 = \frac{5}{2}$$

$$\bullet R(\theta_2, d_3) = 2\frac{1}{2} + 3\frac{1}{2} = 1 + \frac{3}{2} = \frac{5}{2}$$

$$\bullet R(\theta_2, d_4) = 2\frac{1}{2} + 2\frac{1}{2} = 2$$

Από τα μεγαλύτερα τέρση, ζητώ το ελάχιστο άρα,

$$\left. \begin{array}{l} R(\theta_2, d_1) > R(\theta_1, d_1) \\ R(\theta_2, d_2) > R(\theta_1, d_2) \\ R(\theta_2, d_3) < R(\theta_1, d_3) \\ R(\theta_2, d_4) < R(\theta_1, d_4) \end{array} \right\} \begin{array}{l} \text{Και η μέγιστη τιμή είναι } \frac{5}{2} \\ \text{Αρα, ο min max επιλέγει} \\ \text{είναι } \begin{cases} x_1, x=0 \\ x_2, x=1 \end{cases} \end{array}$$

BAYES

Η διακρονοποίηση των κλασικών στατιστικών συμπερασματολογιών από την Μπεϋσιανή είναι η ακόλουθη

Κλασ. στατισ. συμπεραση:

$X_1, \dots, X_n$ εδ από πληθυσμό με σπη ή σπ $f(x, \theta)$ όπου θ άγνωστο αλλά σταθερή παράμετρος

Bayes:

$X_1, \dots, X_n$ εδ από πληθυσμό με σπη ή σπ $f(x, \theta)$ όπου θ η τιμή μιας τιμ θ με γνωστή κατανομή

Ορισμός (Εκ των προτέρων κατανομή)

Εκ των προτέρων κατανομή της τιμ θ είναι η κατανομή που εκφράζει την εκ των προτέρων γνώση που έχουμε για τον παράμετρο θ και καθορίζεται πριν τη συλλογή των δεδομένων

Ορισμός (Εκ των υστέρων κατανομή)

Εκ των υστέρων κατανομή της θ είναι αυτή που επιγράφει την ανανεωμένη γνώση που προκύπτει για την παράμ. θ μετά τη συλλογή των δεδομένων