

ΠΡΟΤΑΣΗ: Δεν υπάρχει βέλτιστος (in particular) εκτιμητής της μη σταθερής παραμετρικής συνάρτησης $g(\theta)$ με βάση το Μεσο Τετραγωνικό Σφάλμα (MSE)

Απόδειξη:

Έστω ότι $\exists$ βέλτιστος εκτιμητής της μη σταθερής παραμετρικής συνάρτησης $g(\theta)$ και είναι ο T^*

Αυτό σημαίνει ότι $E(T^* - g(\theta))^2 \leq E(T - g(\theta))^2$ για $\forall \theta \in \Theta$ και για κάθε άλλο εκτιμητή T (1)

Αγού $g(\theta)$ είναι μη σταθερή, σημαίνει $\exists \theta_1, \theta_2 \in \Theta$ $\mu \epsilon \theta_1 \neq \theta_2$ τ.ω $g(\theta_1) \neq g(\theta_2)$

Αγού η (1) ισχύει $\forall T$ και $\forall \theta \in \Theta$, επιλέγω $T_1 = g(\theta_1)$ τότε $E(T^* - g(\theta)) \leq E(g(\theta_1) - g(\theta))^2$
 $\forall \theta$

Επιλέγω $\theta = \theta_1$, τότε $E(T^* - g(\theta_1)) \leq 0 \Rightarrow P(T^* = g(\theta_1)) = 1$

Αντίστοιχα

Αγού η (1) ισχύει $\forall T$ και $\forall \theta \in \Theta$, επιλέγω $T_2 = g(\theta_2)$ τότε $E(T^* - g(\theta)) \leq E(g(\theta_2) - g(\theta))^2$ $\forall \theta$
για $\theta = \theta_2$ τότε $E(T^* - g(\theta_2)) \leq 0 \Rightarrow P(T^* = g(\theta_2)) = 1$

Άρα $g(\theta_1) = g(\theta_2)$ (Απορία)

□

ΠΑΡΑΔΕΙΓΜΑΤΑ

① $X_1, X_2, \dots, X_n$ τυχαίο δείγμα από $N(\mu, \sigma^2)$, σ^2 άγνωστη διακύμανση

α) Να βρεθεί ο βέλτιστος εκτιμητής του σ^2 στην οικογένεια των εκτιμητών της μορφής

$$d(\underline{x}) = b s^2 = b \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2, \quad b > 0$$

1) Συνάρτηση Απώλειας (Σ.Α.) το Τετραγωνικό Σφάλμα (Τ.Σ.)

2) Συνάρτηση Απώλειας: $L(\theta, d) = \frac{d}{\theta} - 1 - \log \frac{d}{\theta}$ (Stein)

$$R(\theta, d) = E(d - \theta)^2 = \text{Var } d + (E d - \theta)^2 \xrightarrow[\substack{\theta = \sigma^2 \\ d = b s^2}]{\substack{\theta = \sigma^2 \\ d = b s^2}} E(b s^2 - \sigma^2)^2 = \text{Var}(b s^2) + (E b s^2 - \sigma^2)^2 =$$

$$= b^2 \text{Var } s^2 + (b E s^2 - \sigma^2)^2 = b^2 \frac{2\sigma^4}{n-1} + (b-1)^2 \sigma^4$$

$$*: \frac{(n-1)s^2}{\sigma^2} \sim \chi_{n-1}^2$$

$$E\left[\frac{(n-1)s^2}{\sigma^2}\right] = n-1$$

$$\frac{n-1}{\sigma^2} E s^2 = n-1$$

$$\text{Var}\left[\frac{n-1}{\sigma^2} s^2\right] = 2(n-1)$$

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

$$\Rightarrow \text{Var } s^2 = \frac{2\sigma^4}{n-1}$$

$$R(\theta, d) = \frac{2b^2 + (b-1)^2(n-1)}{n-1} \sigma^4 \leadsto \text{Αρκει να ελαχιστοποιήσουμε τον Αριθμητή}$$

Η εύρεση του βέλτιστου εκτιμητή στην οικογένεια της μορφής bs^2 ανάγεται στην ελαχιστοποίηση της συνάρτησης διακινδύνευσης $f(b) = 2b^2 + (b-1)^2(n-1)$

Το ελάχιστο επιτυγχάνεται στο $b^* = \frac{n-1}{n+1}$

Άρα ο βέλτιστος εκτιμητής: $d^* = \frac{s^2(n-1)}{n+1}$

Σύμφωνα με τη Πρόταση δεν υπάρχει βέλτιστος εκτιμητής για τη σταθερή συνάρτηση $g(\theta)$. Σε μια οικογένεια Gauss (περιορισμένος χώρος) υπάρχει!

$$2) R(\theta, d) = E L(\theta, d) = E \left(\frac{d}{\theta} - 1 - \log \frac{d}{\theta} \right) =$$

$$\begin{aligned} \frac{d=bs^2}{\theta=\sigma^2} E \left[\frac{bs^2}{\sigma^2} - 1 - \log \frac{bs^2}{\sigma^2} \right] &= \frac{b}{\sigma^2} \sigma^2 - 1 - E \left[\log \frac{bs^2}{\sigma^2} \right] = \\ &= b - 1 - E \left[\log(bs^2) - \log \sigma^2 \right] = b - 1 - E \left[\log b + \log s^2 - \log \sigma^2 \right] = \\ &= b - 1 - \log b + E(\log s^2) + \log \sigma^2 \end{aligned}$$

Αρκει να ελαχιστοποιήσω τον: $b - 1 - \log b$

Θέλω να βρω $\min R(\theta, d)$ ως προς b

Γραφίστε ότι $Q = \frac{(n-1)s^2}{\sigma^2} \sim \chi_{n-1}^2$

Βρείτε την κατανομή του $Y = S^2$

Ξέρω την κατανομή του Q και ψάχνω την κατανομή του Y

$$\boxed{Y = \frac{\sigma^2}{n-1} Q} \leftarrow h(Q)$$

Χρησιμοποιώ τη 2η μέθοδο για $\min$

$$\boxed{Q = \frac{n-1}{\sigma^2} Y} \leftarrow h^{-1}(Y)$$

$$f_Y^*(y) = f_Q \left(\frac{n-1}{\sigma^2} y \right) \left| \frac{n-1}{\sigma^2} \right|$$

$$\frac{dQ}{dy}$$

Πρέπει να βρω τον f_Q
Την ξέρω αφού f_Q

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

$$\frac{1}{2^{\frac{n-1}{2}} \Gamma(\frac{n-1}{2})} q^{\frac{n-1}{2}-1} e^{-\frac{q}{2}}, q > 0$$

Από Βιβλίο
Πιθανοτήτων-Συμμετρίας
(μεθοδολογίες για
εύρεση κατανομής)

$$f_Y(y) = \frac{\left(\frac{n-1}{\sigma^2} y\right)^{\frac{n-1}{2}-1} e^{-\frac{n-1}{\sigma^2} \frac{y}{2}}}{2^{\frac{n-1}{2}} \Gamma\left(\frac{n-1}{2}\right)} \cdot \frac{n-1}{\sigma^2}$$

$$Y \sim \text{Gamma}\left(\frac{n-1}{2}, \frac{2\sigma^2}{n-1}\right)$$

$$f_Y(y) = \frac{y^{\frac{n-1}{2}-1} e^{-\frac{y}{2\sigma^2}(n-1)}}{\left(\frac{2\sigma^2}{n-1}\right)^{\frac{n-1}{2}} \Gamma\left(\frac{n-1}{2}\right)}, y > 0$$

$$E \log S^2 = \int \log y f_Y(y) dy$$

Χρησιμοποιώ τη 3η μέθοδο:

$$m_Q(t) = E[e^{tQ}]$$

$$E Q^k = \left. \frac{d^k m_Q(t)}{dt^k} \right|_{t=0}$$

$$Y = \frac{\sigma^2}{n-1} Q \stackrel{c=\frac{\sigma^2}{n-1}}{=} cQ$$

$$m_Y(t) = E(e^{tcQ}) = m_Q(ct)$$

$$m_Q(t) = (1-2t)^{-\frac{n-1}{2}}, t < \frac{1}{2}$$

$$m_Y(t) = (1-2ct)^{-\frac{n-1}{2}}, tc < \frac{1}{2}$$

$$m_Y(t) = (1-\underbrace{2c}_{\theta} \underbrace{t}_{\alpha})^{-\frac{n-1}{2}}, t < \frac{1}{2c} \Rightarrow Y \sim \text{Gamma}\left(\frac{n-1}{2}, 2c\right)$$

2) $X_1, X_2, \dots, X_n$ τυχαίο δείγμα από κανονικό με $N(\mu, \sigma^2 = \theta)$, μ γνωστό
Συνάρτηση Απώλειας το Τετραγωνικό Σφάλμα (Τ.Σ)

$$d_1 = \frac{1}{n} \sum_{i=1}^n (X_i - \mu)^2$$

$$d_2 = \kappa d_1, \kappa > 0$$

Είναι καλύτερη; και πότε;

$$R(\theta, d_1) \stackrel{T.S.}{=} \text{Var} d_1 + (E d_1 - \theta)^2 = \frac{2\theta^2}{n}$$

$$R(\theta, d_2) \stackrel{T.S.}{=} \text{Var}(k d_1) + (E k d_1 - \theta)^2 = k^2 \text{Var} d_1 + (k E d_1 - \theta)^2$$

$$\frac{X_i - \mu}{\theta} \sim N(0, 1) \Rightarrow \left(\frac{X_i - \mu}{\theta}\right)^2 \sim \chi_1^2$$

$$\sum_{i=1}^n \left(\frac{X_i - \mu}{\theta}\right)^2 \sim \chi_n^2 \rightarrow E\left[\sum_{i=1}^n \left(\frac{X_i - \mu}{\theta}\right)^2\right] = n$$

$$\text{Var}\left[\sum_{i=1}^n \left(\frac{X_i - \mu}{\theta}\right)^2\right] = 2n$$

$$E d_1 = E\left(\frac{1}{n} \sum_{i=1}^n (X_i - \mu)^2\right) = \frac{1}{n} E \sum_{i=1}^n (X_i - \mu)^2 = \frac{n\theta^2}{n} = \theta$$

$$\text{Var} d_1 = \text{Var}\left(\frac{1}{n} \sum_{i=1}^n (X_i - \mu)^2\right) = \frac{1}{n^2} \text{Var} \sum_{i=1}^n (X_i - \mu)^2 = \frac{1}{n^2} 2n\theta^2 = \frac{2\theta^2}{n}$$

Αρα

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

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

Ο d_2 καλύτερος όταν:

$$F(k, n) < 0$$

Ο d_1 καλύτερος όταν: $F(k, n) > 0$

$$F(k, n) = \frac{2k^2}{n} + (k-1)^2 - \frac{2}{n} = \frac{2k^2 + n(k-1)^2 - 2}{n}$$

$$\Delta = 16, \text{ ~~και~~ } k_{1,2} = \begin{cases} 1 \\ \frac{n-2}{n+2} \end{cases} \quad + | - | + \quad \text{Αρα } F(k, n) < 0 \rightarrow \left(\frac{n-2}{n+2}, 1\right)$$

$$F(k, n) > 0 \rightarrow \text{εκτός του πρώτου}$$

ΟΡΙΣΜΟΣ: (Ελάχιστο ^{μέγιστος} ~~μέγιστος~~ εκτίμησης)

Είναι εκείνος ο εκτιμητής που ελαχιστοποιεί το μέγιστο $\int \eta f d\mu$

$$\min_d \max_{\theta} R(\theta, d)$$

Λέγεται και minimax

ΠΑΡΑΔΕΙΓΜΑ

- ③ $X_1, X_2, \dots, X_n$ τυχαία δείγματα από κανονική $\mu \in N(\mu, 1)$ και Σχέση Αντιθέσεως το ΤΣ
- (i) Να υπολογιστεί $R(\mu, \bar{x})$ (άγνωστο το μ)
- (ii) Αν $x_1, x_2, \dots, x_n$ από $N(\mu, 1)$

ΝΔΟ $T = \frac{1}{n} \sum X_i$ δεν είναι παραδεκτός.

(i) $R(\mu, \bar{x}) = R(\mu, \frac{\sum x_i}{n}) = \text{Var } \bar{x} + (\bar{x} - \mu)^2$

$\text{Var } \bar{x} = \frac{\sigma^2}{n}$

$E \bar{x} = \mu$ Άρα $R(\mu, \bar{x}) = \frac{\sigma^2}{n}$

- (ii) Δεν είναι παραδεκτός, δηλαδή $\exists T_1$ τ.ω. $R(\mu, T_1) < R(\mu, T) = \frac{\sigma^2}{n}$

$T_1 = \frac{1}{2n} \sum_{i=1}^n x_i$ $E T_1 = \mu$ $\text{Var } T_1 = \frac{1}{4n^2} \sum \text{Var } x_i = \frac{1}{4n^2} \cdot 2 \sigma^2 = \frac{\sigma^2}{2n}$

④ Από σημειώσεις σε 23 (ΠΑΡΑΔΕΙΓΜΑ)

Έστω X τ.μ. $\mu \in B(1, \theta)$

$P(X=x, \theta) = \theta^x (1-\theta)^{1-x}$, $x=0, 1$

Ο χώρος αποφάσεων $D = \{a_1, a_2\}$

$L(\theta, d)$:

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

Να βρεθεί ο minimax εκτίμητής

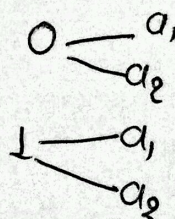
$\min_d \max_{\theta} R(\theta, d)$

Δίνεται $\theta_1 = \frac{1}{4}$, $\theta_2 = \frac{1}{2}$

Θέλω μέγιστο ως προς θ τα $R(\theta, d)$ και ελάχιστο ως προς d .

Θέλω τα $d(x)$

$x \backslash d$	d_1	d_2	d_3	d_4
θ	a_1, a_1	a_2	a_2	
1	a_1, a_2	a_1	a_2	



$P(X=x, \theta)$

$P(X=0, \theta) = 1-\theta$

$P(X=1, \theta) = \theta$

Βρίσκω πρώτα τα max:

$d_1: \max\{R(\theta_1, d_1), R(\theta_2, d_1)\}$

$d_2: \max\{R(\theta_1, d_2), R(\theta_2, d_2)\}$

$d_3: \max\{R(\theta_1, d_3), R(\theta_2, d_3)\}$

$d_4: \max\{R(\theta_1, d_4), R(\theta_2, d_4)\}$

Συνδυάζω κάθε περίπτωση για να προκύψουν οι τιμές

$$\begin{aligned}
 R(\theta_1, d_1(x)) &= E L(\theta_1, d_1(x)) = \sum_{x=0,1} L(\theta_1, d_1(x)) P(X=x, \theta_1) = \\
 &= L(\theta_1, d_1(0)) P(X=0, \theta_1) + L(\theta_1, d_1(1)) P(X=1, \theta_1) = \\
 &= L(\theta_1, a_1)(1-\theta_1) + L(\theta_1, a_1)\theta_1 = 1(1-\theta_1) + 1 \cdot \theta_1 = 1
 \end{aligned}$$

Άρα $R(\theta_1, d_1) = 1$

$$\begin{aligned}
 R(\theta_2, d_1(x)) &= E L(\theta_2, d_1(x)) = \sum_{x=0,1} L(\theta_2, d_1(x)) P(X=x, \theta_2) = \\
 &= L(\theta_2, d_1(0)) P(X=0, \theta_2) + L(\theta_2, d_1(1)) P(X=1, \theta_2) = \\
 &= L(\theta_2, a_1)(1-\theta_2) + L(\theta_2, a_1)\theta_2 = 3(1-\theta_2) + 3\theta_2 = 3
 \end{aligned}$$

Για $R(\theta_1, d_3)$

$$\begin{aligned}
 R(\theta_1, d_3(x)) &= E L(\theta_1, d_3(x)) = \sum_{x=0,1} L(\theta_1, d_3(x)) \cdot P(X=x, \theta_1) = \\
 &= L(\theta_1, d_3(0)) P(X=0, \theta_1) + L(\theta_1, d_3(1)) P(X=1, \theta_1) = \\
 &= L(\theta_1, a_2)(1-\theta_1) + L(\theta_1, a_1)\theta_1 = 4(1-\theta_1) + 1\theta_1 = 4-3\theta_1 = 4-3\frac{1}{4} = \frac{13}{4}
 \end{aligned}$$

Άρα βρεθούν όλα τα αποτελέσματα έχουμε:

$$d_1 = \max\{1, 3\} = 3$$

$$d_2 = \max\{\frac{1}{4}, \frac{5}{2}\} = \frac{5}{2}$$

$$d_3 = \max\{\frac{13}{4}, \frac{5}{2}\} = \frac{13}{4}$$

$$d_4 = \max\{4, 2\} = 4$$

Ποιο είναι το μικρότερο από αυτά;

- Είναι το $\frac{5}{2}$, αντιστοιχεί στην απόφαση d_2

Ο minimax εκτιμητής είναι ο d_2 όταν $d_2 = \begin{cases} a_1, & x=0 \\ a_2, & x=1 \end{cases}$

Τέλος κεφαλαίου, σελ. 62

ΝΑ ΛΥΘΟΥΝ ΓΙΑ ΤΟ ΕΠΟΜΕΝΟ ΜΑΘΗΜΑ: • ΑΣΚΗΣΕΙΣ 1 και 3

• ΝΑ ΒΡΕΘΟΥΝ ΟΛΑ ΤΑ ΑΠΟΤΕΛΕΣΜΑΤΑ $R(\theta, d(x))$ ΓΙΑ ΤΟ ΠΑΡΑΚΕΙΜΕΝΟ ΠΑΡΑΤΗΛΟ