

Θεωρία Αποφάσεων Bayes.Παράδειγμα 1 (από προηγ. μαθ)Έστω $X_1, \dots, X_n \sim \text{i.i.d. } N(\mu, \sigma^2)$ δάχτυλοΝα βρεθεί ο βέλτιστος εκτιμητής του σ^2 στην οικογένεια των εκτιμητών της μορφής $d(X) = bS^2$, $b > 0$ ii) όταν η συνολική απώλεια είναι συνολική Stein ($L(\theta, d) = \frac{d}{\theta} - 1 - \log d$)Λύση

$$\begin{aligned}
 \text{ii) Έχουμε } R(\theta, bS^2) &= E \left[\frac{bS^2}{\theta} - 1 - \log d \right] \\
 &= \frac{b}{\theta} E(S^2) - 1 - E(\log b + \log \frac{S^2}{\theta}) = \\
 &= \frac{b}{\theta} E(S^2) - 1 - E(\log b) - E(\log S^2 - \log \theta) \\
 &= \frac{b}{\theta} \theta - 1 - \log b - \underbrace{E(\log S^2)}_{\text{ποσότητα ανεξαρτητή του } b} + E(\log \theta)
 \end{aligned}$$

Θεωρούμε $g(b) = b - 1 - \log b$ και είναι $g'(b) = 1 - \frac{1}{b} = \frac{b-1}{b} = 0 \Rightarrow b^* = 1$ } Άρα για $b^* = 1$ ελάχιστο.

Επίσης $g''(b) = \frac{1}{b^2} > 0$

Επομένως ο βέλτιστος εκτιμητής θα είναι ο $b^*S^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$

Ορισμός Ένας εκτιμητής θα καλείται ελαχιστομέγιστος (minimax) όταν ελαχιστοποιεί τη μέγιστη J ημία, δηλ. ικανοποιεί τη σχέση $\min_d (\max_{\theta \in \Theta} R(\theta, d))$.

Παράδειγμα 2 Έστω $\text{i.i.d. } X_1, X_2, \dots, X_n$ από $N(\mu, \sigma^2 = 9)$ δάχτυλο

Έχοντας ως συνολική απώλεια το T.E ποια από τις ακόλουθες αποφάσεις (εκτιμητές) είναι καλύτερες και πότε;

$$d_1 = \frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2 \text{ και } d_2 = k d_1, k > 0.$$

ΛΥΣΗ

Θα συγκρίνουμε τη συνολική κινδύνου για τις 2 αποφάσεις, δηλ. την $R(\theta, d_1)$ με τη $R(\theta, d_2)$.

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

Όταν $X_1, X_2, \dots, X_n$ τ.δ από $N(\mu, \sigma^2)$ τότε $\frac{X_i - \mu}{\sigma} \sim N(0, 1)$ και $\left(\frac{X_i - \mu}{\sigma}\right)^2 \sim \chi_1^2$ άρα $\sum_{i=1}^n \frac{(X_i - \mu)^2}{\sigma^2} \sim \chi_{2n}^2 = \chi_n^2 \equiv G\left(\frac{n}{2}, 2\right)$.

Επομένως

$$E\left(\frac{1}{\sigma^2} \sum_{i=1}^n (X_i - \mu)^2\right) = n \quad \text{και} \quad \text{Var}\left(\frac{1}{\sigma^2} \sum_{i=1}^n (X_i - \mu)^2\right) = 2n$$

$$\Rightarrow \text{Var}\left(\sum_{i=1}^n (X_i - \mu)^2\right) = 2n\sigma^4.$$

$$\begin{aligned} \text{Οπότε θα είναι } R(\theta, d_1) &= \text{Var}\left(\frac{1}{n} \sum_{i=1}^n (X_i - \mu)^2\right) + \left[E\left(\frac{1}{n} \sum_{i=1}^n (X_i - \mu)^2\right) - \theta\right]^2 \\ &= \frac{1}{n^2} 2n\sigma^2 + \left[\frac{1}{n} n\sigma^2 - \theta\right]^2 = \frac{2\sigma^2}{n} \end{aligned}$$

Επομένως $\boxed{R(\theta, d_1) = \frac{2\sigma^2}{n}}$

Αντιστοίχα θα ισχύει

$$\begin{aligned} R(\theta, d_2) &= \text{Var } d_2 + (E(d_2) - \theta)^2 \\ &= \text{Var}\left(\frac{k}{n} \sum_{i=1}^n (X_i - \mu)^2\right) + \left[E\left(\frac{k}{n} \sum_{i=1}^n (X_i - \mu)^2\right) - \theta\right]^2 \\ &= \frac{k^2}{n^2} 2n\sigma^2 + \left[\frac{k}{n} n\sigma^2 - \theta\right]^2 \\ &= \frac{2k^2}{n} \sigma^2 + (k-1)^2 \theta^2 = \left[\frac{2k^2}{n} + (k-1)^2\right] \theta^2. \end{aligned}$$

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

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

Θέτουμε $F(k, n)$ 2^{ου} βαθμού πολυώνυμο
μεγισ. όρου ως προς k με συντελ. $(2+n)$

Καίνοιν πρώτες και βρίσκω $\Delta = 16$

$$=1 \quad K_{1,2} = \begin{cases} 1 \\ \frac{n-2}{n+2} \end{cases}$$

Επομένως θα είναι:

$$\begin{array}{c} + \quad | \quad - \quad | \quad + \\ \hline \frac{n-2}{n+2} \quad \quad \quad 1 \end{array}$$

οπότε στο $(\frac{n-2}{n+2}, 1)$ ισχύει $F(k, n) < 0 \Rightarrow$

$\cdot d_2$ καλύτερος ευτιμητής (απόφαση)

Ενώ εκτός των ριγών ισχύει ότι $F(k, n) > 0 \Rightarrow$

$\cdot d_1$ καλύτερος ευτιμητής (απόφαση).

Παράδειγμα 3 Αν d είναι μια απόφαση για την παράμετρο θ , με σταθερή συνάρτηση κινδύνου και $d_i^* = \lambda_1 d + \lambda_2$, $\lambda_i \quad i=1,2$ σταθερές. Να δείξετε ότι η $R(d_i^*, \lambda_1 \theta + \lambda_2)$ είναι σταθερή, όταν ως συνολική απώλεια χρησιμοποιείται το Τ.Σ.

Λύση

$$\begin{aligned} R(d_i^*, \lambda_1 \theta + \lambda_2) &= E(d_i^* - \lambda_1 \theta - \lambda_2)^2 \\ &= E(\lambda_1 d + \lambda_2 - \lambda_1 \theta - \lambda_2)^2 = E(\lambda_1 (d - \theta))^2 \\ &= \lambda_1^2 E(d - \theta)^2 = \lambda_1^2 R(\theta, d) \end{aligned}$$

Αρα αποδείχθηκε το ζητούμενο. σταθερά.

Παράδειγμα 4 Έστω τ.δ $X_1, \dots, X_n$ από $N(\mu, 1)$

i) Να υπολογιστεί η $R(\mu, \bar{X})$, εάν ως συνολική απώλεια θεωρούμε το Τ.Σ.
ii) Αν $X_1, \dots, X_n$ τ.δ από την $N(\mu, 1)$

Νόο η σ.σ $T = \frac{1}{n} \sum_{i=1}^n X_i$ ΔΕΝ είναι παραδειγμας ευτιμητής.

Λύση

i) Θα είναι $R(\mu, \bar{X}) = \text{Var}(\bar{X}) + (E(\bar{X}) - \mu)^2$ όμως $X_1, X_2, \dots, X_n$ τ.δ από $N(\mu, 1)$

άρα $\bar{X} \sim N(\mu, 1/n)$ Επομένως $E(\bar{X}) = \mu$ και $\text{Var}(\bar{X}) = \frac{1}{n}$

οπότε $R(\mu, \bar{X}) = \frac{1}{n} + (\mu - \mu)^2 = \frac{1}{n} \Rightarrow R(\mu, \bar{X}) = \frac{1}{n}$

ii) Γνωρίζουμε ότι

$$R(\theta, d_1) = \frac{1}{n} \quad \text{Από i)}$$

$$\left[\text{όπου } d_1 = T = \frac{1}{n} \sum_{i=1}^n (x_i) \right]$$

και $\theta = \mu$

$$\text{ωστόσο } \exists T^* = \frac{1}{2n} \sum_{i=1}^{2n} x_i \text{ τ.ω. } R(\mu, T^*) = \frac{1}{2n} < \frac{1}{n} = R(\mu, T)$$

Επομένως ο $T = \frac{1}{n} \sum_{i=1}^n x_i$ δεν είναι παραδειγματίς διότι
βρήκαμε ότι υπάρχει καλύτερος, ο T^* .

Παράδειγμα 5 Έστω X τ.μ με $X \sim B\left(\overset{n}{1}, \overset{p}{\theta}\right)$, δηλ. $P(X=x, \theta) = \theta^x (1-\theta)^{1-x}$

$$\text{και } \theta \in \Theta = \left\{ \underset{\theta_1}{\frac{1}{4}}, \underset{\theta_2}{\frac{1}{2}} \right\}$$

↓ Δυνατές ενέργειες

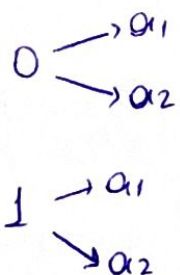
Επίσης ο χώρος αποφάσεων $D = \{a_1, a_2\}$

και η συνολική απώλεια ορίζεται ως εξής:

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

Να βρεθεί ο minimax ευτητητής του θ .

Το ΛΥΣΗ
Το σύνολο των δυνατών αποφάσεων d για τις τιμές του θ θα είναι



Βάση του πολλαπλασιαστικού κανόνα, προκύπτουν
 $2 \times 2 = 4$ δυνατές αποφάσεις d_1, d_2, d_3, d_4

Επομένως το σύνολο των δυνατών αυτών αποφάσεων μπορεί να γραφεί
ως εξής:

$\theta \backslash d$	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

Ζητάμε να υπολογίσουμε την ποσότητα $\min_{d \in \Theta} \max_{\theta \in \Theta} R(\theta, d)$

Αρχικά θα υπολογίσουμε την συνιστητική κινδύνου $R(\theta, d)$

$$\text{Θα είναι } R(\theta, d(x)) = E(L(\theta, d(x)))$$

$$\theta \in \{\theta_1, \theta_2\}$$

$$\text{και } d \in \{d_1, d_2, d_3, d_4\} \quad \left. \begin{array}{l} \\ \end{array} \right\} \begin{array}{l} \Rightarrow \text{Βαση πολιτικού κανόνα προκίντουν} \\ 2 \cdot 4 = 8 \text{ επιλογές} \end{array}$$

Επομένως θα είναι

$$R(\theta_1, d_1) = E(L(\theta_1, d_1(x))) = \sum_{\substack{x: \text{δυνατές τιμές} \\ x: 0, 1 \Rightarrow x \text{ διακριτή}}} L(\theta_1, d_1(x)) P(x=x/\theta_1) = L(\theta_1, d_1(0)) P(x=0/\theta_1) + L(\theta_1, d_1(1)) P(x=1/\theta_1)$$

$$= L(\theta_1, \alpha_1) P(x=0/\theta_1) + L(\theta_1, \alpha_1) P(x=1/\theta_1) \Rightarrow R(\theta_1, d_1) = 1 \cdot \frac{3}{4} + 1 \cdot \frac{1}{4} = 1$$

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

$$\begin{aligned} R(\theta_1, d_2) &= E(L(\theta_1, d_2(x))) = \sum_x L(\theta_1, d_2(x)) P(x=x/\theta_1) \\ &= L(\theta_1, d_2(0)) P(x=0, \theta_1) + L(\theta_1, d_2(1)) P(x=1, \theta_1) \\ &= L(\theta_1, \alpha_1) P(x=0/\theta_1) + L(\theta_1, \alpha_2) P(x=1/\theta_1) \\ &= 1 \cdot \frac{3}{4} + 4 \cdot \frac{1}{4} = \frac{7}{4} \end{aligned}$$

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

ομοίως

$$\begin{aligned} R(\theta_1, d_3) &= L(\theta_1, \alpha_2) P(x=0, \theta_1) + L(\theta_1, \alpha_1) P(x=1, \theta_1) \\ &= 4 \cdot \frac{3}{4} + 1 \cdot \frac{1}{4} = \frac{13}{4} \end{aligned}$$

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

Αντίστοιχα υπολογίζονται οι υπόλοιπες τιμές της συνιστητικής κινδύνου.

$$R(\theta_2, d_1) = L(\theta_2, \alpha_1) P(x=0, \theta_1) + L(\theta_2, \alpha_1) P(x=1, \theta_2) = 3$$

$$R(\theta_2, d_2) = L(\theta_2, \alpha_1) P(x=0, \theta_2) + L(\theta_2, \alpha_2) P(x=1, \theta_2) = 5/2$$

$$R(\theta_2, d_3) = L(\theta_2, \alpha_2) P(x=0, \theta_2) + L(\theta_2, \alpha_1) P(x=1, \theta_2) = 5/2$$

$$R(\theta_2, d_4) = L(\theta_2, \alpha_2) P(x=0, \theta_2) + L(\theta_2, \alpha_2) P(x=1, \theta_2) = 2$$

$R(\theta, d)$	θ_1	θ_2
d_1	1	(3)
d_2	$7/4$	($5/2$)
d_3	($13/4$)	$5/2$
d_4	(4)	2

- ΒΗΜΑΤΑ
- ① Σταθεροποιούμε το d και επιλέγουμε το θ που μεγιστοποιεί την $R(\theta, d)$
 - ② Βρίσκουμε το μέγιστο για κάθε σταθεροποίηση του d και επιλέγουμε το ελάχιστο από τα μέγιστα.

Επομένως ο ελαχιστομέγιστος ευτιμής θα είναι

$$d_2(x) = \begin{cases} a_1, & x=0 \\ a_2, & x=1 \end{cases}$$

ΘΕΩΡΙΑ BAYES

Η Διαφοροποίηση της κλασικής στατιστικής συμπερασματολογίας (Κ.Σ.Σ) από τη Μπεϋζιανή είναι η ακόλουθη.

• ΚΖΣ: Έστω $x_1, x_2, \dots, x_n$ τ.δ από έναν πληθυσμό με σ.π. ή σ.π. $f(x|\theta)$ όπου θ : άγνωστη παράμετρος αλλά σταθερή ποσότητα.

• Bayes: Έστω $x_1, \dots, x_n$ τ.δ από έναν πληθυσμό με σ.π. ή σ.π. $f(x, \theta)$ όπου θ η τιμή της τ.μ. θ με γνωστή κατανομή.

Ορισμοί:

prior distribution
• Εκ των προτέρων κατανομή της τ.μ. θ είναι η κατανομή που εκφράζει την εκ των προτέρων γνώση που έχουμε για την παράμετρο θ και καθορίζεται πριν την συλλογή των δεδομένων

posterior distribution
• Εκ των υστέρων κατανομή της τ.μ. θ είναι αυτή που εκφράζει την αναμενόμενη γνώση που προκύπτει για τη θ μετά τη συλλογή των δεδομένων

Στην Μπεϋζιανή προσέγγιση υπάρχουν 4 βήματα κλειδιά:

- 1^ο Καθορισμός της $f(x|\theta)$, δηλ. της κατανομής της τ.μ X .
όταν δίνεται η παράμετρος θ
- 2^ο Καθορισμός της εκ των προτέρων κατανομής $f(\theta)$ ή $\pi(\theta)$, δηλ της κατανομής που εκφράζει την εκ των προτέρων γνώση για το θ .
- 3^ο Υπολογισμός της εκ των υστέρων κατανομής $f(\theta|x)$, δηλ της κατανομής που εκφράζει την πληροφορία για το θ όταν δίνεται η τιμή x της τ.μ X .
- 4^ο Εξαγωγή συμπερασμάτων που στηρίζονται στην εκ των υστέρων κατανομή.

Όπως αναφέρθηκε η $f(x|\theta)$ υπολογίζεται μέσω του θ . Bayes για τυχαιές μεταβλητές

Υπενθύμιση από τη Θεωρία Πιθανοτήτων
Έστω A, B δύο ενδεχόμενα του $\mathcal{S}$ τ.ω

$$P(A/B) = \frac{P(A \cap B)}{P(B) \neq 0} = P(A \cap B) = \frac{P(A/B)P(B)}{P(B/A)P(A)} \left\{ \Rightarrow P(A/B) = \frac{P(B/A)P(A)}{P(B)} \right.$$

Γενικότερα έστω $H_1, \dots, H_k$ μια διαμέριση του $\mathcal{S}$, δηλ.

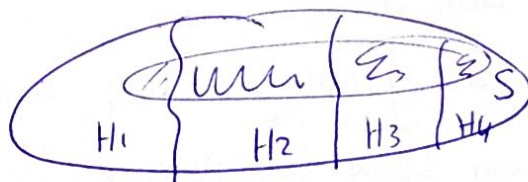
$$H_i \subseteq \mathcal{S}, \forall i$$

$$H_i \cap H_j = \emptyset \quad \forall i \neq j$$

$$\text{και } H_1 \cup H_2 \cup \dots \cup H_k = \mathcal{S}$$

Έστω A ένα ενδεχόμενο του $\mathcal{S}$ τότε $P(A) = \sum_{i=1}^k P(A/H_i) P(H_i)$

Επιπλέον $P(H_i/A) = \frac{P(A/H_i)P(H_i)}{\sum P(A/H_i)P(H_i)}$ αφού $P(H_i/A) = \frac{P(H_i \cap A)}{P(A)}$



Παράδειγμα

Έστω τα ακόλουθα ενδεχόμενα

$\Pi = \{ \text{Ένα τυχαίο επιλεγμένο άτομο παίζει από μια αθλήτρια} \}$

$\Theta = \{ \text{Το τεστ στο οποίο υποβλήθηκε ήταν θετικό} \}$

και γνωρίζουμε ότι $P(\Theta/\Pi) = 0,95$ $P(\Theta/\Pi^c) = 0,05$

$P(\Pi) = 0,12$. Να βρεθεί $P(\Pi/\Theta) = ?$

Λύση

$$\begin{aligned} P(\Pi/\Theta) &= \frac{P(\Theta/\Pi) P(\Pi)}{P(\Theta)} = \frac{P(\Theta/\Pi) P(\Pi)}{P(\Theta/\Pi) P(\Pi) + P(\Theta/\Pi^c) P(\Pi^c)} \quad \swarrow \text{δεν παίζει} \\ &= \frac{0,95 \cdot 0,12}{0,95 \cdot 0,12 + 0,05 \cdot 0,88} = \dots \end{aligned}$$

Το Θεώρημα Bayes σε όρους τ.μ. εκφράζεται ως εξής:

$$f(\Theta/x) = \frac{f(x, \Theta)}{f(x)} = \frac{f(x/\Theta) f(\Theta)}{\int_{\Omega} f(x/\Theta) f(\Theta) d\Theta}$$

Ω = πεδίο ορισμού που προκύπτει
από την τωμή $f(x, \Theta)$ κ $f(\Theta)$ για το x .

ΠΑΡΑΤΗΡΗΣΕΙΣ

1) Η $f(\Theta/x)$ είναι σπη

2) Ολοκληρώσεως $\int f(x/\Theta) f(\Theta) d\Theta$, ο παρονομαστής στο Θ. Bayes είναι σπλην μόνο ως προς x

Συνεπώς για δεδομένες παρατηρήσεις ο παρονομαστής είναι σταθερά και αφομόζεται σταθερά κανονικοποίησης.

Οπότε θα μπορούσε να γραφεί σε $f(\Theta/x) \propto f(x/\Theta) f(\Theta)$

δηλ. η εκ των υστέρων κατανομή

είναι ανάλογη του γινόμενου της εκ των προτέρων κατανομής με την σπλην $f(x/\Theta)$

3) Αν έχουμε τ.σ. $X_1, \dots, X_n$ τότε $f(\theta/x) = \frac{f(\theta) \prod_{i=1}^n f(x_i/\theta)}{\int (\prod_{i=1}^n f(x_i/\theta)) f(\theta) d\theta}$