

Θεωρία Αποφάσεων Bayes.

Σχετική ιστοσελίδα www.math.uoi.gr/~abatsidis/731.html.
Γραφείο 303α.

e-mail: varappas@uoi.gr / rappas.v@gmail.com

Συγγραμμά * κ. Φερεντίνος Εκθετική σιμωχένεια κατανομήν
Θεωρία Bayes (Παν/κες εκδόσεις)

Βοηθητικά συγγραμμάτα: (ηλεκτρονική μορφή)

1) Κατζέρη Μ (2012) Σημειώσεις για το μάθημα Θ.Α - Bayes

2) Δελλαπόρτας Π κ' Τσιαμυρτζής Π (2004)

Σημειώσεις Μαθημάτος "Στατιστική κατά Bayes".

3) Παπαϊωάννου κ Φερεντίνος : Μαθηματική Στατιστική
(εκδόσεις Σταμούλη) Βιβλιοθήκη.

→ Πιθανή πρόοδος (Ιανουάριο)

Ορισμός Η Θ.Α μελετά τα προβλήματα συμπερασματολογίας (εκτιμητική, έλεγχος υποθέσεων) κατά τη διαίρεση των οποίων η λήψη απόφασης καθορίζεται με θεμελιωμένα κριτήρια

Σε έναν πληθυσμό (σύνολο οντοτήτων του οποίου ένα ή περισσότερα χαρακτηριστικά γνωρίσματα μελετούμε) που μεταβάλλονται από μέλος σε μέλος του πληθυσμού, υπάρχει κάποιο χαρακτηριστικό πχ μέσο μηνιαίο εισόδημα μέσος αριθμός πελατών στη μονάδα του χρόνου, με μοναδική άγνωστη τιμή, έστω θ .

Το θ καλείται άγνωστη παράμετρος και ανήκει σε ένα γνωστό πεπερασμένο ή μη σύνολο Θ .

θ κεφαλαίο

Το Θ λέγεται παραμετρικός χώρος.

πχ $N(\mu, \sigma^2)$

$$\mu \in \mathbb{R} \quad \sigma^2 > 0. \in (0, +\infty)$$

Ο παραμετρικός χώρος Θ περιέχει όλες τις δυνατές (πιθανές) τιμές της άγνωστης παραμέτρου θ .

ΠΑΡΑΤΗΡΗΣΗ: Την πραγματική τιμή δεν την ξέρουμε, αλλά θέλουμε να την εκτιμήσουμε.

Παραδείγματα:

• Έστω $X \sim B(n, p)$ $\rightarrow$ πιθανότητα $p \in [0, 1] \equiv \Theta$

• Έστω $X \sim \text{Exp}(\theta)$ $\theta \in (0, +\infty) \equiv \Theta$

• Έστω $X \sim N(\mu, \sigma^2)$ άγνωστα τότε $\Theta = (\theta_1, \theta_2)$ με $\theta_1 = \mu$
 $\mu \in \mathbb{R}$
 $\sigma^2 \in (0, +\infty)$
 $\theta_2 = \sigma^2$

Άρα $\Theta \equiv \mathbb{R} \times (0, +\infty) = \mathbb{R} \times \mathbb{R}^+$.

Στη θεωρία Αποφάσεων στη Διοίκηση μας έχουμε για την εκτίμηση και για τη λήψη γενικότερα αποφάσεων τα δεδομένα:

$X = (x_1, x_2, \dots, x_n)$, όπου είναι οι παρατηρούμενες τιμές των τ.μ $x_1, x_2, \dots, x_n$

Ο αριθμός των δεδομένων n , καλείται μέγεθος του δείγματος και όταν οι τ.μ $x_1, x_2, \dots, x_n$ είναι ανεξάρτητες (δηλ δεν σχετίζονται) και ισόνομες (δηλ έχουν την ίδια κατανομή) σ.π. ή σ.π. λέμε ότι έχουμε τδ (i.i.d)

Αν $x_i \sim f(x_i | \theta)$ ανεξάρτητες τότε η από κοινού κατανομή ($i=1, \dots, n$) του θα είναι:

$$f(x, \theta) = f(x_1, \dots, x_n | \theta) = \prod_{i=1}^n f(x_i, \theta)$$

λόγω ανεξ.

ή από κοινού

είναι το γινόμενο των περιθωρίων

Άλλωστε και από τη θεωρία πιθανοτήτων

ισχύει

$$P(A_1 \cap A_2 \cap \dots \cap A_n) \stackrel{\text{ανεξ}}{=} \prod_{i=1}^n P(A_i).$$

ΠΡΟΣΟΧΗ

Η εκτίμηση της άγνωστης παραμέτρου θ ή μιας άγνωστης συνάρτησης αυτής $g(\theta)$ βασίζεται σε συνιστες των δεδομένων $x_1, \dots, x_n$ και γνωστών ποσοτήτων που δεν εξαρτώνται, από την παράμετρο θ .

Παραδείγματα

- Έστω $N(\mu, \sigma^2)$ με μ άγνωστό
τότε μπορεί να υποτεθεί ότι η ποσότητα $\frac{\sum_{i=1}^n (x_i - \mu)^2}{n}$

Είναι στατιστική συνλση.

- ~~BC~~
- Όμως για $N(\mu, \sigma^2)$ με μ, σ^2 άγνωστη, δηλ $\theta = (\theta_1, \theta_2) = (\mu, \sigma^2)$
η ποσότητα $\sum_{i=1}^n \frac{(x_i - \theta_1)^2}{n}$ δεν είναι στατιστική συνλση
~~Είναι~~
διότι περιέχει την άγνωστο παράμετρο $\theta_1 = \mu$

Μια στατιστική συνλση (σ.σ) που χρησιμοποιείται για την
επίμνηση μιας άγνωστης παραμέτρου θ καλείται επιμνητής
Ενώ η τιμή που προκύπτει όταν αντικατασταθούν τα
 $x_1, \dots, x_n$ από τις τιμές τους λέγεται επίμνηση της άγνωστης
παραμέτρου.

- Χώρος αποφάσεων: συμβολίζεται με D και είναι το σύνολο
όλων των δυνατών ενεργειών - πράξεων - αποφάσεων d
που μπορεί κανείς να λάβει.
- Στην εκτιμητική ο χώρος των αποφάσεων είναι ο ίδιος με τον
παραμετρικό χώρο Θ (δηλ. $D \equiv \Theta$)
- Ενώ, στον έλεγχο υποθέσεων, ο χώρος των αποφάσεων είναι
το σύνολο $D = \{ \text{"απορρίπτω την } H_0, \text{"} \cup \{ \text{"δεν απορρ την } H_0 \} \}$, όπου
 H_0 μηδενική υπόθεση.

• Συνάρτηση Απόφασης - Απόφαση (decision function)

Είναι μια μέτρησιμη συνάρτηση d , η οποία ορίζεται στο χώρο $\mathbb{R}^n$ και παίρνει τιμές στο χώρο D , δηλ ισχύει ότι $d: \mathbb{R}^n \rightarrow D$

Η τιμή $d(x)$ στο σημείο x ονομάζεται απόφαση

• Συνάρτηση ζημίας ή απώλειας (Loss function). ($L(\theta, d)$) - Συμβολισμός

Είναι η συνλση που εκφράζει την απώλεια-ζημία που θα έχουμε όταν παίρνουμε την απόφαση $d(x)$, τη συχνή που η πραγματική τιμή που θέλουμε να ευτιμήσουμε είναι θ .

Απόφαση
 $d(x)$

Φύση-πραγματικότητα.
 θ

Από τώρα και στο εξής θα συμβολίζουμε με $d(x) = d$ επομένως για τη συνλση ζημίας ή απώλειας θα ισχύει:

$L(\theta, d) : \Theta \times D \rightarrow \mathbb{R}$ και πληρεί τις εξής 2 ιδιότητες

① $L(\theta, d) \geq 0 \quad \forall \theta$

② $L(\theta, \theta) = 0$

Κατά τριάδα (Θ, D, L) ορίζει ένα στατιστικό πρόβλημα απόφασης

Παραδείγματα συνόσεων ανώλειας.

1) Απόλυτο σφάλμα $L(\theta, d) = |d - \theta|$

2) Τετραγωνικό σφάλμα $L(\theta, d) = (d - \theta)^2$

3) Γενίευση των περιπτώσεων 1 ή 2 : $L(\theta, d) = |d - \theta|^k, k > 0.$

4) LINEX (linear exponential function) : $L(\theta, d) = e^{k(d-\theta)} - k(d-\theta) - 1$
 $k = \text{σταθερά.}$

5) Stein-function : $L(\theta, d) = \frac{d}{\theta} - 1 - \log \frac{d}{\theta}$

6) Διαφορές συνόσεις "διαβαίθμησης"

πχ $L(\theta, d) = \begin{cases} (d - \theta)^2 & \text{οταν } d < \theta \text{ (υποεπιτίμηση)} \\ 10(d - \theta)^2 & \text{οταν } d \geq \theta \text{ (υπερεπιτίμηση)} \end{cases}$

↑
Με αυτόν τον τρόπο τιμωρείται

πιο αυστηρά η υπερεπιτίμηση της
αξιμίας παραμέτρου θ .

$\Rightarrow 1) L(\theta, \theta) = 0 \Rightarrow |\theta - \theta| = 0$

2) $L(d, \theta) = |d - \theta| > 0$

Εφαρμογή: Νόμο η συνολική LINEX είναι συνολική ανώλειας

4) $L(\theta, d) = e^{k(d-\theta)} - k(d-\theta) - 1$, θα πρέπει να ισχύουν οι 2 προϋποθέσεις

① $L(\theta, \theta) = e^{k(\theta-\theta)} - k(\theta-\theta) - 1 = 0$

② Θέλουμε να είναι $L(\theta, d) \geq 0 \quad \forall \theta.$

$e^{k(d-\theta)} - k(d-\theta) - 1 \geq 0 \text{ (σπινι)}$

• Θέτω $y = e^{k(d-\theta)}$ παραχγο...

Μειονεκτήματα σωλons απώλειας.

1) $\theta = \text{αίχμωτο}$

2) $d = d(x)$
 $\uparrow$ τυχαϊότητα

Ερώτημα: Πως μπορούμε να αποφασίσουμε για την καταλληλότητα μιας στατιστικής σωλons (σ.σ) $T(x)$ για την εκτίμηση της αίχμωτης ποσότητας θ ;

Αν

το απόλυτο σφάλμα $|T(x) - \theta|$ ή οποιαδήποτε άλλη αντίστοιχη παράσταση που προκύπτει με την υπόθεση οποιαδήποτε άλλης παρόμοιας ποσότητας παραστάωσης-σωλons απώλειας θα οδηγούσε στην αναγνώριση του εφής προβλήματος:

Οι παραπάνω παραστάσεις έχουν και αίχμωτη και τυχαία

τίμη

Αίχμωτη γιατί περιέχουν την αίχμωτη παράμετρο θ .

και Τυχαία επειδή είναι τ.μ ως σωλon των $x_1, \dots, x_n$

Επομένως δεν εξυπηρετεί ως κριτήριο για να αξιολογήσουμε τον εκτιμητή $T(x)$

Για να καταλήξουμε σε ένα χρήσιμο κριτήριο θα πρέπει "να διαχειριστούμε" την τυχαϊότητα του σφάλματος εκτίμησης. Ένας τρόπος να γίνει αυτό είναι να θεωρήσουμε την κατά μέσο όρο τιμή της παράστάωσης-σωλons απώλειας ως προς την κατανομή του x .

- Σύνολο κινδύνου - Διακινδυνεύσεις (Risk function)

Συμβολισμός $R(\theta, d)$

Η συνολική κινδύνου ορίζεται ως η μέση τιμή ως προς X της συνολικής απώλειας, δηλ ισχύει $R(\theta, d) = E[L(\theta, d)]$.

ΠΡΟΤΑΣΗ: Όταν η συνολική απώλεια είναι το τετραγωνικό σφάλμα $L(\theta, d) = (d - \theta)^2$, τότε η συνολική κινδύνου μπορεί να γραφεί ως εξής

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

Απόδ

Έχουμε $R(\theta, d) \stackrel{\text{ορισμός}}{=} E[L(\theta, d)] = E[(d - \theta)^2]$

ΥΠΕΝΟΧΥΜΙΣΗ: Γνωρίζω ότι $\text{Var}(X) = E(X^2) - (EX)^2$

Εφαρμόζουμε για $X = d - \theta$ και έχουμε

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

όπως $\text{Var}(ax + b) = a^2 \text{Var} x$ για $a = b = 0$
 $E(ax + b) = aE(x) + b$

Επομένως $R(\theta, d) = \text{Var } d + [E(d) - \theta]^2$

$\downarrow$
απόφαση

$\downarrow$
Μεροληψία.

ΠΑΡΑΤΗΡΗΣΗ:

Η ποσότητα $E(d) - \theta$ ονομάζεται ^{ποσό} μεροληψίας

Όταν $E(d) < \theta$ έχουμε υποεπιτίμηση, ενώ

όταν $E(d) > \theta$ έχουμε υπερεπιτίμηση

Όταν $E(d) = \theta$ τότε λέμε ότι ο επιτιμητής μας είναι αμερόληπτος επιτιμητής της άγνωστης παραμέτρου θ .

ΥΠΕΝΘΥΜΙΣΗ

Έστω τιμ X τότε ισχύει ότι

$$E(h(x)) = \begin{cases} \sum h(x) P(X=x) & \text{αν } X = \text{διακριτή} \\ \int_{n_0}^x h(x) f_X(x) dx & \text{αν } X = \text{συνεχής} \end{cases}$$

Παράδειγμα

I] Έστω $X_1, \dots, X_n$ τ.δ από $N(\theta, \sigma^2)$ με θ άγνωστη παράμετρο με $\theta \in \mathbb{R}$ και σ^2 γνωστή θετική ποσότητα $\sigma^2 > 0$

Να υπολογιστεί: α) το μέσο τετραγωνικό σφάλμα (ΜΤΣ) του $\bar{X}$.

β) Η συνολική κίνδυνος με συνολική απώλεια το ~~απώ~~ απόλυτο σφάλμα (Α.Σ) του $\bar{X}$.

Λύση

α) Το ΜΤΣ του $\bar{X}$

$$E(\bar{X} - \theta)^2 \stackrel{\text{ΠΡΟΤΑΣΗ}}{=} \text{Var}(\bar{X}) + (E(\bar{X}) - \theta)^2$$

όμως ισχύει ότι

1^ο ΤΡΟΠΟΣ.

$$\text{Var}(\bar{X}) = \text{Var}\left(\frac{\sum X_i}{n}\right) \stackrel{\text{ ανεξ. }}{=} \frac{1}{n^2} \sum \text{Var}(X_i) = \frac{1}{n^2} \sum \sigma^2 = \sigma^2/n$$

$$E(\bar{X}) = E\left(\frac{\sum X_i}{n}\right) = \frac{1}{n} \sum E(X_i) = \frac{1}{n} \sum \theta = \theta$$

Επομένως θα είναι:

$$E(\bar{X}-\theta)^2 = \text{Var}(\bar{X}) + (E(\bar{X})-\theta)^2 = \frac{\sigma^2}{n} + (\theta-\theta)^2 = \frac{\sigma^2}{n}$$

2^{ος} ΤΡΟΠΟΣ

Γνωρίζουμε ότι αν $X_1, \dots, X_n$ εἶναι από $N(\theta, \sigma^2)$

τότε $\bar{X} \sim N(\theta, \sigma^2/n)$

$$E(\bar{X}) = \theta \text{ και } \text{Var}(\bar{X}) = \frac{\sigma^2}{n}.$$

β) Θα έχουμε ότι

$$E[|\bar{X}-\theta|] = \int |\bar{x}-\theta| * (\text{σπν του } \bar{x}) d\bar{x}$$

δηλ $E(h(\bar{x}))$ όπου $h(\bar{x}) = |\bar{x}-\theta|$

$$\text{θα είναι } L(\theta, d) = L(\theta, \bar{x}) = |\bar{x}-\theta|$$

$$\text{Θέτουμε } \bar{x} = y \Rightarrow d\bar{x} = dy$$

$$\overset{\text{άρα}}{R(\theta, \bar{x})} = E[L(\theta, \bar{x})] = E|\bar{x}-\theta| = \int_{-\infty}^{+\infty} |y-\theta| \left(2n \frac{\sigma^2}{n}\right)^{-1/2} e^{-\frac{1}{2\sigma^2/n}(y-\theta)^2} dy$$

$$\begin{aligned} \underline{y-\theta=z} \\ \underline{dy=dz} \end{aligned} \quad \int_{-\infty}^{+\infty} |z| \underbrace{\left(2n \frac{\sigma^2}{n}\right)^{-1/2} e^{-n/2\sigma^2}}_{f(z)} dz$$

$$= 2 \int_0^{+\infty} z \left(2n \frac{\sigma^2}{n}\right)^{-1/2} e^{-n/2\sigma^2} z^2 dz$$

αλλα
συν
n f(z)
f(-z)=f(z)

Παρατηρούμε ότι $\left(e^{-n/2\sigma^2 z^2} \right)' = -2z \frac{n}{2\sigma^2} e^{-n/2\sigma^2 z^2}$

Επομένως με το lemma

θα έχουμε:

$$= R(\theta, d) = R(\theta, \bar{x}) = E|\bar{x} - \theta| = \frac{\sqrt{2}}{\sqrt{nn}} \sigma.$$