ΠΑΝΕΠΙΣΤΗΜΙΟ ΙΩΑΝΝΙΝΩΝ
ΑΝΟΙΚΤΑ ΑΚΑΔΗΜΑΪΚΑ ΜΑΘΗΜΑΤΑ

Εφαρμοσμένη Στατιστική

Εκτιμητική

Διδάσκων: Επίκουρος Καθηγητής
Κωνσταντίνος Μπλέκας

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης

Άδειες Χρήσης

- Το παρόν εκπαιδευτικό υλικό υπόκειται σε άδειες χρήσης Creative Commons.
- Για εκπαιδευτικό υλικό, όπως εικόνες, που υπόκειται σε άλλου τύπου άδειας χρήσης, η άδεια χρήσης αναφέρεται ρητώς.

Εκτιμητές Μεγίστης Πιθανοφάνειας (Maximum Likelihood Estimators – MLE)

Εστω τ.δ. $X=\{x_1, x_2, \ldots, x_n\}$ με κατανομή με σ.π.π. $f(x;\theta)$.
Η από-κοινού σ.π.π. των n δειγμάτων είναι η συνάρτηση

$$L(\theta) = f(x_1, x_2, \ldots, x_n; \theta) = \prod_{i=1}^{n} f(x_i; \theta)$$

και ονομάζεται **πιθανοφάνεια** (likelihood) του δείγματος.

Ορισμός: Μία εκτιμήτρια $\hat{\theta}$ καλείται εκτιμήτρια μεγίστης πιθανοφάνειας (**MLE**) της παραμέτρου θ αν ισχύει ότι

$$\hat{\theta} = \sup_{\theta} L(\theta) = \sup_{\theta} f(x_1, x_2, \ldots, x_n; \theta)$$

δηλ. η $\hat{\theta}$ είναι η τιμή της παραμέτρου θ που μεγιστοποιεί την συνάρτηση $L(\theta)$.

Εκτιμητές Μεγίστης Πιθανοφάνειας (Maximum Likelihood Estimators – MLE) (συν.)

Συνήθως είναι ευκολότερο και βολικότερο να αναζητούμε το σημείο μεγίστου της λογαριθμικής πιθανοφάνειας (*log-likelihood*):

$$\mathrm{L}(\theta) = \ln L(\theta) = \sum_{i=1}^{n} \ln f(x_i;\theta)$$

Καθώς έχουν το ίδιο σημείο μεγίστου διότι η συνάρτηση *ln* είναι αύξουσα.

Παρατηρήσεις:

- Η μεγιστοποίηση συνήθως επιτυγχάνεται με μεθόδους Διαφορικού Λογισμού
- Υπάρχουν περιπτώσεις με περισσότερο του ενός μέγιστα ή υπάρχουν περιπτώσεις πολλών τοπικών μεγίστων.
- Η μέθοδος *MLE* προσφέρει ευκολία στον τρόπο αναζήτησης και μπορεί να εφαρμοστεί σε ευρεία κλάση προβλημάτων

Εκτιμητές Μεγίστης Πιθανοφάνειας (Maximum Likelihood Estimators – MLE) (συν.)

Μεθοδολογία εύρεσης MLE εκτιμητών

- Εάν $X=\{x_1, x_2, \ldots, x_n\}$ *τ.δ.* προερχόμενο από κατανομή με σ.π.π. $f(x;\theta)$, τότε γράφουμε την συνάρτηση πιθανοφάνειας

$$L(\theta) = f(x_1, x_2, \ldots, x_n; \theta) = \prod_{i=1}^{n} f(x_i; \theta)$$

- Μεγιστοποίηση της $l(\theta) = \ln L(\theta)$ ως προς θ, όπου $\hat{\theta}$ η τιμή για την οποία

$$\frac{\partial}{\partial \theta} l(\theta) = \frac{\partial}{\partial \theta} \ln L(\theta) = 0$$

- Για να είναι μέγιστο το ακρότατο θα πρέπει να ισχύει

$$\left. \frac{\partial^2}{\partial \theta^2} l(\theta) \right|_{\theta=\hat{\theta}} < 0$$

το οποίο και ισχύει για τις περισσότερες γνωστές κατανομές

Μέθοδος των ροπών

- Είναι η απλούστερη μέθοδος εύρεσης εκτιμητών
- Σχεδόν πάντα δίνει εκτιμητή ο οποίος όμως δεν είναι πάντα «ποιοτικός»

Μέθοδος των ροπών:

βρίσκουμε τις δειγματικές ροπές και τις εξισώνουμε με τις αντίστοιχες ροπές της κατανομής του πληθυσμού.

Μέθοδος των ροπών (συν.)

- Έστω τ.δ. X={x1, x2, …, xn} από κατανομή με σ.π.π. f(x;θ), όπου θ=($θ_1$, $θ_2$, …, $θ_k$).
- Ροπές πληθυσμού Δειγματικές Ροπές

$$\mu_1 = E(X) \qquad m_1 = \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$\mu_2 = E(X^2) \qquad m_2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2$$

$$\mu_k = E(X^k) \qquad m_k = \frac{1}{n}\sum_{i=1}^{n} x_i^k$$

- Οι ροπές του πληθυσμού θα είναι συναρτήσεις των *k* παραμέτρων, δηλ. $\mu_j(\theta_1, \theta_2, \ldots, \theta_k)$. Έτσι προκύπτουν *k* εξισώσεις με *k* αγνώστους και οι εκτιμήτριες θα είναι η λύση του συστήματος.

Εκτιμητές Bayes

- Στη κλασική στατιστική όπου η παράμετρος θ θεωρείται άγνωστη αλλά σταθερή ποσότητα που πρέπει να εκτιμηθεί.
- Η **Μπεϋζιανή (Bayesian) στατιστική** θεωρεί ότι η παράμετρος θ είναι μία τ.μ. η οποία ακολουθεί μία γνωστή (συγκεκριμένη) κατανομή.
- Η κατανομή αυτή ονομάζεται *εκ των προτέρων κατανομή* ή (*a-priori distribution*) και εκφράζει την αρχική μας θεώρηση ή τους περιορισμούς που θέτουμε για την παράμετρο θ.
- Η εκ των προτέρων κατανομή είναι *υποκειμενική* και θα αναπροσαρμοστεί με βάση τις παρατηρήσεις του δείγματος.

Εκτιμητές Bayes (συν.)

- Η αναπροσαρμοσμένη εκ των προτέρων κατανομή ονομάζεται εκ των υστέρων κατανομή ή a-posteriori distribution και προκύπτει από τον κανόνα του Bayes.
- Έστω π(θ) η εκ των προτέρων κατανομή και f(X;θ) δειγματική κατανομή. Τότε **η εκ των υστέρων κατανομή** δίνεται ως εξής:

$$p(\theta \mid X) = \frac{f(X \mid \theta)\pi(\theta)}{p(X)} = \frac{f(X \mid \theta)\pi(\theta)}{\int f(X \mid \theta)\pi(\theta)d\theta} \propto L(\theta)\pi(\theta)$$

- Η p(X) είναι η **περιθώρια** συνάρτηση του δείγματος X. Η f(X;θ) είναι η γνωστή μας συνάρτηση **πιθανοφάνειας**.
- Για την εκτίμηση της παραμέτρου θ μπορούμε να χρησιμοποιήσουμε χαρακτηριστικά της εκ των υστέρων κατανομής του θ, p(θ|X), όπως ο μέσος ή η διάμεσος.

Εκτιμητές Bayes (συν.)

- Ειδική περίπτωση αποτελεί η ομοιόμορφη κατανομή ως εκ των προτέρων κατανομή π(θ) της παραμέτρου θ. Τότε, καθώς η ομοιομορφία δεν προσφέρει κάποια ειδική γνώση ή περιορισμό για την παράμετρο θ, η εκ των υστέρων κατανομή συμπίπτει με την πιθανοφάνεια, δηλ.

$$p(\theta \mid X) \propto L(\theta)$$

  και επομένως ο εκτιμητής Bayes είναι ο MLE εκτιμητής.

- **Συζυγείς κατανομές (conjugate family)**:

  Μία οικογένεια Π των εκ των προτέρων κατανομών π(θ) ονομάζεται **συζυγής** αν η εκ των υστέρων κατανομή p(θ|X) του θ ανήκει στην ίδια οικογένεια Π.

Εκτιμητές Ελαχίστων Τετραγώνων (Least Squares)

Εστω τ.δ. $X=\{x_1, x_2, \ldots, x_n\}$ με κατανομή με σ.π.π. $f(x;\theta)$.

Ενας **εκτιμητής ελαχίστων τετραγώνων** μιας παραμετρικής συνάρτησης g(θ) προκύπτει από την ελαχιστοποίηση της παράστασης:

$$\sum_{i=1}^{n}\left(x_i - g(\theta)\right)^2$$

η οποία είναι η **συνάρτηση τετραγωνικού σφάλματος**.
Στο παραπάνω πρόβλημα ελαχιστοποίησης μπορούμε εναλλακτικά να χρησιμοποιηθεί πρώτερη γνώση του θ με την μορφή περιορισμών, δηλ. **πρόβλημα ελαχιστοποίησης με περιορισμούς**

Τέλος Ενότητας

Χρηματοδότηση

- Το παρόν εκπαιδευτικό υλικό έχει αναπτυχθεί στα πλαίσια του εκπαιδευτικού έργου του διδάσκοντα.
- Το έργο «**Ανοικτά Ακαδημαϊκά Μαθήματα στο Πανεπιστήμιο Ιωαννίνων**» έχει χρηματοδοτήσει μόνο τη αναδιαμόρφωση του εκπαιδευτικού υλικού.
- Το έργο υλοποιείται στο πλαίσιο του Επιχειρησιακού Προγράμματος «Εκπαίδευση και Δια Βίου Μάθηση» και συγχρηματοδοτείται από την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο) και από εθνικούς πόρους.

Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο
ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΔΙΑ ΒΙΟΥ ΜΑΘΗΣΗ
επένδυση στην κοινωνία της γνώσης
ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ
ΕΣΠΑ 2007-2013
πρόγραμμα για την ανάπτυξη
ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ
Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης

Σημειώματα

Σημείωμα Ιστορικού Εκδόσεων Έργου

Το παρόν έργο αποτελεί την έκδοση 1.0.

Έχουν προηγηθεί οι κάτωθι εκδόσεις:

- Έκδοση 1.0 διαθέσιμη εδώ. http://ecourse.uoi.gr/course/view.php?id=1052.

Σημείωμα Αναφοράς

Copyright Πανεπιστήμιο Ιωαννίνων, Διδάσκων: Επίκουρος Καθηγητής Κωνσταντίνος Μπλέκας. «Εφαρμοσμένη Στατιστική. Εκτιμητική». Έκδοση: 1.0. Ιωάννινα 2014. Διαθέσιμο από τη δικτυακή διεύθυνση: http://ecourse.uoi.gr/course/view.php?id=1052.

Σημείωμα Αδειοδότησης