

Στατιστική II

Εβδομάδα 2: Πληροφορία Fisher, Cramer-Rao, Επάρκεια, Μέγιστη Πιθανοφάνεια

Θεωρία και Ασκήσεις

Τμήμα Οικονομικής Επιστήμης, ΟΠΑ

Εισαγωγή: από το «καλός εκτιμητής» στο «ο βέλτιστος εκτιμητής»

Στην Εβδομάδα 1 μάθαμε να αξιολογούμε εκτιμητές (αμεροληψία, διασπορά, MSE, συνέπεια) και να τους κατασκευάζουμε με τη Μέθοδο των Ροπών. Όμως μείναμε με δύο ανοιχτά ερωτήματα:

1. Υπάρχει ένα κάτω όριο στη διασπορά κάθε αμερόληπτου εκτιμητή; Δηλαδή, πόσο «καλός» μπορεί θεωρητικά να γίνει ένας αμερόληπτος εκτιμητής; — Η απάντηση είναι η **Ανισότητα Cramer-Rao**.
2. Υπάρχει μια συστηματική, σχεδόν αυτόματη μέθοδος που να δίνει (ασυμπτωτικά) τον βέλτιστο εκτιμητή; — Η απάντηση είναι η **Μέθοδος Μέγιστης Πιθανοφάνειας (MLE)**.

Το θεωρητικό όριο: Πληροφορία Fisher, Ανισότητα Cramer-Rao, και η έννοια της **Επάρκειας** (πότε ένα στατιστικό «συμπυκνώνει» όλη την πληροφορία του δείγματος).

Η κυρίαρχη μέθοδος: Μέθοδος Μέγιστης Πιθανοφάνειας, οι ασυμπτωτικές της ιδιότητες, η ιδιότητα αναλλοιότητας (invariance), η μέθοδος Delta, και μια σύντομη εισαγωγή στη Γενικευμένη Μέθοδο Ροπών (GMM).

Σύνδεση με την Οικονομετρία

Η Μέθοδος Μέγιστης Πιθανοφάνειας είναι ο πυρήνας σχεδόν κάθε σύγχρονου οικονομετρικού μοντέλου: logit & probit (διακριτές επιλογές), Tobit (λογοκριμένα δεδομένα), Poisson regression (δεδομένα μέτρησης), GARCH (μεταβλητότητα χρηματοοικονομικών αποδόσεων), μοντέλα διάρκειας (ανεργία, επιβίωση επιχειρήσεων). Ό,τι κάνετε εδώ θα το ξανασυναντήσετε αυτούσιο.

Προαπαιτούμενα από την Εβδομάδα 1: αμεροληψία, διασπορά, MSE, συνέπεια, σύγκλιση κατά πιθανότητα, νόμος μεγάλων αριθμών, UMVUE.

Περιεχόμενα

Εισαγωγή	1
1 Πληροφορία Fisher, Cramer-Rao, Επάρκεια	2
1.1 Η συνάρτηση πιθανοφάνειας	2
1.2 Συνάρτηση Score και Πληροφορία Fisher	3
1.3 Η Ανισότητα Cramer-Rao	4
1.4 Επάρκεια	5

2	Λυμένες Ασκήσεις	6
2.1	Ομάδα Α: Πληροφορία Fisher και Cramer-Rao	6
2.2	Ομάδα Β: Επάρκεια	8
3	Μέγιστη Πιθανοφάνεια	9
3.1	Η μέθοδος	9
3.2	Ασυμπτωτικές ιδιότητες του MLE	10
3.3	Ιδιότητα Αναλλοιότητας (Invariance)	10
3.4	Η Μέθοδος Delta	11
3.5	Σύντομη επαφή: Γενικευμένη Μέθοδος Ροπών (GMM)	11
4	Λυμένες Ασκήσεις	12
4.1	Ομάδα Γ: Μέγιστη Πιθανοφάνεια	12
4.2	Ομάδα Δ: Αναλλοιότητα και Μέθοδος Delta	16
5	Σύνοψη Τύπων	19

1 Πληροφορία Fisher, Cramer-Rao, Επάρκεια

1.1 Η συνάρτηση πιθανοφάνειας

Όλη η θεωρία αυτής της εβδομάδας στηρίζεται σε μία κεντρική έννοια: τη *συνάρτηση πιθανοφάνειας*.

Ορισμός 2.1 (Πιθανοφάνεια και λογαριθμική πιθανοφάνεια)

Έστω $X_1, \dots, X_n$ τυχαίο δείγμα με κοινή συνάρτηση (πυκνότητας ή μάζας) πιθανότητας $f(x; \theta)$, όπου θ η άγνωστη παράμετρος. Η **συνάρτηση πιθανοφάνειας** (likelihood function) ορίζεται ως

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

Η **λογαριθμική πιθανοφάνεια** (log-likelihood) είναι

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

Πιθανότητα vs Πιθανοφάνεια

Είναι η ίδια μαθηματική έκφραση $\prod f(x_i; \theta)$, αλλά διαβάζεται «ανάποδα»:

- Ως **πιθανότητα**: το θ είναι σταθερό και ρωτάμε πόσο πιθανά είναι τα δεδομένα.
- Ως **πιθανοφάνεια**: τα δεδομένα $x_1, \dots, x_n$ είναι σταθερά (τα παρατηρήσαμε) και ρωτάμε ποιά τιμή του θ τα κάνει πιο εύλογα.

Δουλεύουμε σχεδόν πάντα με τη $\ell(\theta)$ αντί της $L(\theta)$: το γινόμενο γίνεται άθροισμα (ευκολότερη παραγωγή) και ο λογάριθμος είναι γνησίως αύξων, οπότε δεν μετακινεί τη θέση του μεγίστου.

1.2 Συνάρτηση Score και Πληροφορία Fisher

Ορισμός 2.2 (Συνάρτηση score)

Η **συνάρτηση score** είναι η παράγωγος της λογαριθμικής πιθανοφάνειας ως προς την παράμετρο:

$$S(\theta) = \frac{\partial \ell(\theta)}{\partial \theta} = \sum_{i=1}^n \frac{\partial \ln f(X_i; \theta)}{\partial \theta}.$$

Θεώρημα 2.1 (Μηδενική μέση τιμή του score)

Κάτω από συνθήκες ομαλότητας, η μέση τιμή της συνάρτησης score είναι μηδέν:

$$\mathbb{E} \left[\frac{\partial \ln f(X; \theta)}{\partial \theta} \right] = 0.$$

Απόδειξη. Αφού $\int f(x; \theta) dx = 1$ για κάθε θ , παραγωγίζουμε ως προς θ και εναλλάσσουμε παράγωγο-ολοκλήρωμα:

$$0 = \frac{\partial}{\partial \theta} \int f(x; \theta) dx = \int \frac{\partial f(x; \theta)}{\partial \theta} dx = \int \frac{\partial \ln f(x; \theta)}{\partial \theta} f(x; \theta) dx = \mathbb{E} \left[\frac{\partial \ln f(X; \theta)}{\partial \theta} \right].$$

Χρησιμοποιήσαμε την ταυτότητα $\frac{\partial \ln f}{\partial \theta} = \frac{1}{f} \frac{\partial f}{\partial \theta}$. □

Ορισμός 2.3 (Πληροφορία Fisher)

Η **πληροφορία Fisher** μίας παρατήρησης ορίζεται ως η διασπορά της score (που, λόγω μηδενικής μέσης τιμής, ισούται με τη δεύτερη ροπή):

$$I_1(\theta) = \mathbb{E} \left[\left(\frac{\partial \ln f(X; \theta)}{\partial \theta} \right)^2 \right].$$

Κάτω από συνθήκες ομαλότητας, ισχύει η ισοδύναμη (και συνήθως ευκολότερη) μορφή:

$$I_1(\theta) = -\mathbb{E} \left[\frac{\partial^2 \ln f(X; \theta)}{\partial \theta^2} \right].$$

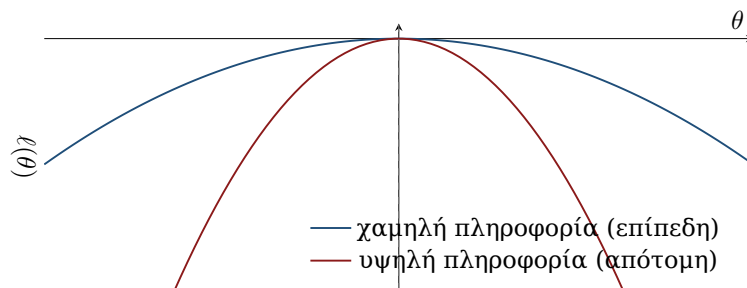
Για τυχαίο δείγμα μεγέθους n , η συνολική πληροφορία είναι *προσθετική*:

$$I_n(\theta) = n I_1(\theta).$$

Διαισθητική ερμηνεία

Η πληροφορία Fisher μετράει την *καμπυλότητα* της λογαριθμικής πιθανοφάνειας γύρω από το μέγιστο. Μια απότομα κορυφωμένη $\ell(\theta)$ (μεγάλη $|\ell''|$) σημαίνει ότι μικρές αλλαγές στο θ μειώνουν δραστικά την πιθανοφάνεια — άρα τα δεδομένα είναι πολύ «πληροφοριακά» για το θ . Αντίθετα, μια επίπεδη $\ell(\theta)$ σημαίνει ότι πολλές τιμές του θ είναι σχεδόν εξίσου εύλογες — λίγη πληροφορία.

Η προσθετικότητα $I_n = n I_1$ είναι η μαθηματική διατύπωση του «περισσότερα δεδομένα = περισσότερη πληροφορία», γραμμικά στο n .



Σχήμα 1: Δύο λογαριθμικές πιθανοφάνειες με το ίδιο μέγιστο $\hat{\theta}$. Η απότομη (κόκκινη) έχει μεγαλύτερη πληροφορία Fisher — τα δεδομένα «δείχνουν» πιο καθαρά πού είναι το θ .

1.3 Η Ανισότητα Cramer-Rao

Θεώρημα 2.2 (Ανισότητα Cramer-Rao)

Έστω $\hat{\theta}$ αμερόληπτος εκτιμητής της παραμέτρου θ , βασισμένος σε τυχαίο δείγμα μεγέθους n . Κάτω από συνθήκες ομαλότητας:

$$\text{Var}(\hat{\theta}) \geq \frac{1}{I_n(\theta)} = \frac{1}{n I_1(\theta)}.$$

Η ποσότητα $\frac{1}{n I_1(\theta)}$ ονομάζεται **κάτω φράγμα Cramer-Rao** (Cramer-Rao Lower Bound — CRLB).

Ορισμός 2.4 (Αποτελεσματικός εκτιμητής)

Ένας αμερόληπτος εκτιμητής που *επιτυγχάνει* το κάτω φράγμα Cramer-Rao, δηλαδή $\text{Var}(\hat{\theta}) = \frac{1}{n I_1(\theta)}$, ονομάζεται **αποτελεσματικός** (efficient). Ένας αποτελεσματικός εκτιμητής είναι αυτομάτως UMVUE.

Γιατί είναι σημαντικό

Το CRLB μάς δίνει ένα *απόλυτο μέτρο σύγκρισης*. Στην Εβδομάδα 1 μπορούσαμε μόνο να συγκρίνουμε δύο εκτιμητές μεταξύ τους (σχετική αποτελεσματικότητα). Τώρα έχουμε ένα θεωρητικό «δάπεδο»: κανένας αμερόληπτος εκτιμητής δεν μπορεί να έχει διασπορά μικρότερη από $1/(n I_1)$. Αν βρούμε έναν εκτιμητή που το πιάνει, ξέρουμε ότι είναι ο καλύτερος δυνατός — δεν χρειάζεται να ψάξουμε άλλο.

Ορισμός 2.5 (Αποδοτικότητα εκτιμητή)

Η **αποδοτικότητα** (efficiency) ενός αμερόληπτου εκτιμητή $\hat{\theta}$ ορίζεται ως

$$\text{eff}(\hat{\theta}) = \frac{\text{CRLB}}{\text{Var}(\hat{\theta})} = \frac{1/(n I_1(\theta))}{\text{Var}(\hat{\theta})} \in (0, 1].$$

Τιμή 1 σημαίνει πλήρως αποδοτικός (επιτυγχάνει το CRLB).

1.4 Επάρκεια

Η έννοια της *επάρκειας* απαντά στο ερώτημα: «ποιό στατιστικό περιέχει *όλη* την πληροφορία του δείγματος για την παράμετρο θ ;»

Ορισμός 2.6 (Επαρκές στατιστικό)

Ένα στατιστικό $T = T(X_1, \dots, X_n)$ ονομάζεται **επαρκές** (sufficient) για την παράμετρο θ αν η δεσμευμένη κατανομή του δείγματος $(X_1, \dots, X_n)$ *δοθέντος* του $T = t$ **δεν εξαρτάται από το θ** .

Διαισθητικά

Αν το T είναι επαρκές, τότε μόλις γνωρίζουμε την τιμή του T , οι αρχικές παρατηρήσεις δεν προσφέρουν *καμία* επιπλέον πληροφορία για το θ . Όλη η «χρήσιμη πληροφορία» έχει συμπυκνωθεί στο T . Π.χ., για Bernoulli δείγμα, το πλήθος επιτυχιών $\sum X_i$ είναι επαρκές — η σειρά με την οποία ήρθαν οι επιτυχίες/αποτυχίες είναι άσχετη για την εκτίμηση του p .

Ο ορισμός είναι δύσχρηστος στην πράξη (απαιτεί υπολογισμό δεσμευμένων κατανομών). Ευτυχώς υπάρχει ένα απλό κριτήριο:

Θεώρημα 2.3 (Θεώρημα Παραγοντοποίησης Fisher-Neyman)

Το στατιστικό $T(X_1, \dots, X_n)$ είναι επαρκές για το θ *αν και μόνο αν* η συνάρτηση πιθανοφάνειας μπορεί να γραφεί ως

$$L(\theta) = g(T(x_1, \dots, x_n); \theta) \cdot h(x_1, \dots, x_n),$$

όπου:

- η g εξαρτάται από τα δεδομένα *μόνο* μέσω του T και μπορεί να εξαρτάται από το θ ,
- η h *δεν* εξαρτάται από το θ (μπορεί να εξαρτάται ελεύθερα από τα δεδομένα).

Πρακτική μέθοδος

Για να βρούμε επαρκές στατιστικό:

1. Γράφουμε την πιθανοφάνεια $L(\theta) = \prod f(x_i; \theta)$.
2. Την «μαζεύουμε» αλγεβρικά ώστε η εξάρτηση από το θ να εμφανίζεται μόνο μέσα από μία ποσότητα T των δεδομένων.
3. Ό,τι περισσεύει και δεν εξαρτάται από το θ το βάζουμε στην h .
4. Το T είναι το επαρκές στατιστικό.

Σύνδεση με την Οικονομετρία

Η επάρκεια εξηγεί *γιατί* στην οικονομετρία αναφέρουμε μόνο συγκεντρωτικά στατιστικά ($\sum x_i$, $\sum x_i^2$, $\sum x_i y_i$) και όχι ολόκληρα τα δεδομένα: για τα κλασικά μοντέλα, αυτά τα αθροίσματα είναι επαρκή στατιστικά. Ολόκληρος ο πίνακας δεδομένων «συμπυκνώνεται» χωρίς απώλεια πληροφορίας στα $X'X$ και $X'y$.

2 Λυμένες Ασκήσεις

2.1 Ομάδα Α: Πληροφορία Fisher και Cramer-Rao

Άσκηση 1: CRLB για Bernoulli — η αναλογία αθέτησης είναι αποτελεσματική

Μια τράπεζα εκτιμά την αναλογία p δανείων σε αθέτηση. Τυχαιο δείγμα $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} \text{Bernoulli}(p)$, με σ.μ.π.

$$f(x; p) = p^x(1-p)^{1-x}, \quad x \in \{0, 1\}.$$

(α') Υπολογίστε την πληροφορία Fisher $I_1(p)$ μίας παρατήρησης.

(β') Γράψτε το κάτω φράγμα Cramer-Rao για κάθε αμερόληπτο εκτιμητή του p .

(γ') Δείξτε ότι ο εκτιμητής $\hat{p} = \bar{X}$ επιτυγχάνει το CRLB. Συμπεράνετε ότι είναι αποτελεσματικός (UMVUE).

Λύση

(α) Λογάριθμος της σ.μ.π.:

$$\ln f(x; p) = x \ln p + (1-x) \ln(1-p).$$

Πρώτη παράγωγος ως προς p (συνάρτηση score):

$$\frac{\partial \ln f}{\partial p} = \frac{x}{p} - \frac{1-x}{1-p}.$$

Δεύτερη παράγωγος:

$$\frac{\partial^2 \ln f}{\partial p^2} = -\frac{x}{p^2} - \frac{1-x}{(1-p)^2}.$$

Παίρνουμε μέση τιμή και αλλάζουμε πρόσημο, χρησιμοποιώντας $\mathbb{E}(X) = p$:

$$I_1(p) = -\mathbb{E}\left[\frac{\partial^2 \ln f}{\partial p^2}\right] = \frac{\mathbb{E}(X)}{p^2} + \frac{1-\mathbb{E}(X)}{(1-p)^2} = \frac{p}{p^2} + \frac{1-p}{(1-p)^2} = \frac{1}{p} + \frac{1}{1-p}.$$

Συνδυάζοντας τα κλάσματα:

$$I_1(p) = \frac{(1-p) + p}{p(1-p)} = \frac{1}{p(1-p)}.$$

(β) Η συνολική πληροφορία είναι $I_n(p) = nI_1(p) = \frac{n}{p(1-p)}$. Άρα

$$\text{CRLB} = \frac{1}{I_n(p)} = \frac{p(1-p)}{n}.$$

(γ) Από την Εβδομάδα 1 (Άσκηση 3) γνωρίζουμε ότι ο $\hat{p} = \bar{X}$ είναι αμερόληπτος με

$$\text{Var}(\hat{p}) = \frac{p(1-p)}{n}.$$

Αυτό είναι ακριβώς το CRLB. Άρα ο $\hat{p}$ επιτυγχάνει το κάτω φράγμα, οπότε είναι **αποτελεσματικός** και άρα **UMVUE**: κανένας αμερόληπτος εκτιμητής του p δεν μπορεί να έχει μικρότερη διασπορά.

Έλεγχος

Αποδοτικότητα: $\text{eff}(\hat{p}) = \frac{\text{CRLB}}{\text{Var}(\hat{p})} = \frac{p(1-p)/n}{p(1-p)/n} = 1$. Πλήρως αποδοτικός. ✓

Άσκηση 2: CRLB για Poisson — αξιώσεις ασφάλισης

Ο αριθμός αξιώσεων ανά μήνα είναι $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} \text{Poisson}(\lambda)$, με

$$f(x; \lambda) = \frac{e^{-\lambda} \lambda^x}{x!}, \quad x = 0, 1, 2, \dots$$

(α') Υπολογίστε την πληροφορία Fisher $I_1(\lambda)$.

(β') Δώστε το CRLB και ελέγξτε αν ο $\hat{\lambda} = \bar{X}$ είναι αποτελεσματικός.

(γ') Γράψτε την ασυμπτωτική κατανομή του $\hat{\lambda} = \bar{X}$.

Λύση

(α)

$$\ln f(x; \lambda) = -\lambda + x \ln \lambda - \ln(x!).$$

$$\frac{\partial \ln f}{\partial \lambda} = -1 + \frac{x}{\lambda}, \quad \frac{\partial^2 \ln f}{\partial \lambda^2} = -\frac{x}{\lambda^2}.$$

Με $\mathbb{E}(X) = \lambda$:

$$I_1(\lambda) = -\mathbb{E}\left[-\frac{X}{\lambda^2}\right] = \frac{\mathbb{E}(X)}{\lambda^2} = \frac{\lambda}{\lambda^2} = \boxed{\frac{1}{\lambda}}.$$

(β) $I_n(\lambda) = n/\lambda$, άρα

$$\text{CRLB} = \frac{\lambda}{n}.$$

Από την Εβδομάδα 1, ο $\hat{\lambda} = \bar{X}$ είναι αμερόληπτος με $\text{Var}(\bar{X}) = \text{Var}(X)/n = \lambda/n$ (αφού για Poisson $\text{Var}(X) = \lambda$). Αυτό ισούται με το CRLB → **αποτελεσματικός, UMVUE**.

(γ) Από το Κεντρικό Οριακό Θεώρημα, $\sqrt{n}(\bar{X} - \lambda) \xrightarrow{d} N(0, \text{Var}(X)) = N(0, \lambda)$. Άρα

$$\boxed{\sqrt{n}(\hat{\lambda} - \lambda) \xrightarrow{d} N(0, \lambda)}.$$

Ισοδύναμα, για μεγάλα n : $\hat{\lambda} \approx N(\lambda, \lambda/n)$ — η ασυμπτωτική διασπορά ισούται με το CRLB, όπως αναμένεται για αποτελεσματικό εκτιμητή.

Άσκηση 3: CRLB για τη μέση τιμή κανονικής (γνωστή διασπορά)

Έστω $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} N(\mu, \sigma^2)$ με σ^2 γνωστή και μ άγνωστη. (Π.χ. καθημερινές αποδόσεις με γνωστή ιστορική μεταβλητότητα.) Η σ.π.π. είναι

$$f(x; \mu) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right).$$

(α') Υπολογίστε $I_1(\mu)$.

(β') Δώστε το CRLB και ελέγξτε αν ο $\bar{X}$ το επιτυγχάνει.

Λύση**(α)**

$$\ln f(x; \mu) = -\frac{1}{2} \ln(2\pi\sigma^2) - \frac{(x - \mu)^2}{2\sigma^2}.$$

$$\frac{\partial \ln f}{\partial \mu} = \frac{x - \mu}{\sigma^2}, \quad \frac{\partial^2 \ln f}{\partial \mu^2} = -\frac{1}{\sigma^2}.$$

Η δεύτερη παράγωγος είναι σταθερή (δεν εξαρτάται από x), οπότε

$$I_1(\mu) = -\mathbb{E}\left[-\frac{1}{\sigma^2}\right] = \boxed{\frac{1}{\sigma^2}}.$$

(β) $I_n(\mu) = n/\sigma^2$, άρα

$$\text{CRLB} = \frac{\sigma^2}{n}.$$

Ο $\bar{X}$ είναι αμερόληπτος με $\text{Var}(\bar{X}) = \sigma^2/n = \text{CRLB} \rightarrow$ **αποτελεσματικός**.

Παρατήρηση

Παρατηρήστε το μοτίβο: στις τρεις παραπάνω ασκήσεις (Bernoulli, Poisson, Κανονική) ο δειγματικός μέσος είναι πάντα αποτελεσματικός. Αυτό δεν είναι σύμπτωση: και οι τρεις κατανομές ανήκουν στην **εκθετική οικογένεια** (exponential family), για την οποία ο MLE της φυσικής παραμέτρου επιτυγχάνει πάντα το CRLB. Στην οικονομετρία, τα Generalized Linear Models (GLM) — που περιλαμβάνουν OLS, logit, Poisson regression — βασίζονται ακριβώς σε αυτή την οικογένεια.

2.2 Ομάδα Β: Επάρκεια**Άσκηση 4: Επαρκή στατιστικά για Εκθετική και Poisson**

Με χρήση του θεωρήματος παραγοντοποίησης Fisher-Neyman, βρείτε ένα επαρκές στατιστικό για:

(α') $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} \text{Exp}(\lambda)$ (rate λ), με $f(x; \lambda) = \lambda e^{-\lambda x}$, $x > 0$. (Διάρκεια ζωής δανείων.)

(β') $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} \text{Poisson}(\lambda)$. (Αξιώσεις ασφάλισης.)

Λύση**(α) Εκθετική.** Η πιθανοφάνεια:

$$L(\lambda) = \prod_{i=1}^n \lambda e^{-\lambda x_i} = \lambda^n \exp\left(-\lambda \sum_{i=1}^n x_i\right).$$

Αυτό εξαρτάται από τα δεδομένα μόνο μέσω του $T = \sum_{i=1}^n x_i$. Θέτουμε:

$$g(T; \lambda) = \lambda^n e^{-\lambda T}, \quad h(x_1, \dots, x_n) = 1.$$

Η g εξαρτάται από τα δεδομένα μόνο μέσω του T , η h δεν εξαρτάται από το λ .

Άρα από το θεώρημα παραγοντοποίησης,

$$T = \sum_{i=1}^n X_i \text{ (ισοδύναμα } \bar{X}) \text{ είναι επαρκές για το } \lambda.$$

(β) Poisson.

$$L(\lambda) = \prod_{i=1}^n \frac{e^{-\lambda} \lambda^{x_i}}{x_i!} = e^{-n\lambda} \lambda^{\sum x_i} \cdot \frac{1}{\underbrace{\prod_{i=1}^n x_i!}_{\text{δεν εξαρτάται από } \lambda}}.$$

Θέτουμε $T = \sum x_i$, και

$$g(T; \lambda) = e^{-n\lambda} \lambda^T, \quad h(x_1, \dots, x_n) = \frac{1}{\prod x_i!}.$$

Άρα

$$T = \sum_{i=1}^n X_i \text{ είναι επαρκές για το } \lambda.$$

Παρατήρηση

Σε αμφότερες τις περιπτώσεις το επαρκές στατιστικό είναι το $\sum X_i$ — ισοδύναμα ο $\bar{X}$. Αυτό συνδέεται άμεσα με τις προηγούμενες ασκήσεις: ο εκτιμητής που επιτυγχάνει το CRLB είναι σχεδόν πάντα συνάρτηση του επαρκούς στατιστικού. Αυτό δεν είναι τυχαίο — είναι το περιεχόμενο του θεωρήματος Rao-Blackwell, που σας λέει ότι ο βέλτιστος αμερόληπτος εκτιμητής πρέπει να εξαρτάται από τα δεδομένα μόνο μέσω του επαρκούς στατιστικού.

3 Μέγιστη Πιθανοφάνεια

3.1 Η μέθοδος

Ορισμός 2.7 (Εκτιμητής Μέγιστης Πιθανοφάνειας)

Ο **εκτιμητής μέγιστης πιθανοφάνειας** (Maximum Likelihood Estimator — MLE) είναι η τιμή της παραμέτρου που μεγιστοποιεί τη συνάρτηση πιθανοφάνειας:

$$\hat{\theta}_{\text{ML}} = \arg \max_{\theta} L(\theta) = \arg \max_{\theta} \ell(\theta).$$

Διαδικασία βήμα-προς-βήμα:

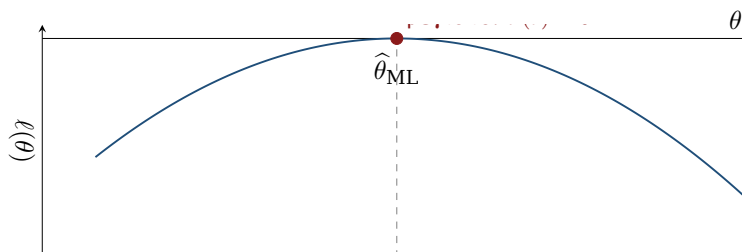
1. Γράφουμε την πιθανοφάνεια $L(\theta) = \prod f(x_i; \theta)$.
2. Παίρνουμε λογάριθμο: $\ell(\theta) = \sum \ln f(x_i; \theta)$.
3. Παραγωγίζουμε και θέτουμε την παράγωγο ίση με μηδέν (συνθήκη πρώτης τάξης / likelihood equation):

$$\frac{\partial \ell(\theta)}{\partial \theta} = 0.$$

4. Λύνουμε ως προς θ .
5. Επαληθεύουμε ότι πρόκειται για μέγιστο (συνθήκη δεύτερης τάξης: $\ell''(\hat{\theta}) < 0$).

Προσοχή: η μέθοδος FOC δεν δουλεύει πάντα

Η συνθήκη πρώτης τάξης $\ell'(\theta) = 0$ υποθέτει ότι το μέγιστο βρίσκεται στο εσωτερικό του παραμετρικού χώρου και ότι η ℓ είναι παραγωγίσιμη εκεί. Σε κατανομές με «σύνορο» που εξαρτάται από την παράμετρο (π.χ. $U(0, \theta)$), το μέγιστο βρίσκεται στο άκρο και πρέπει να εργαστούμε με μονοτονία αντί για παράγωγο (βλ. Άσκηση 8).



Σχήμα 2: Ο MLE είναι το σημείο όπου η λογαριθμική πιθανοφάνεια μεγιστοποιείται (η εφαπτομένη είναι οριζόντια).

3.2 Ασυμπτωτικές ιδιότητες του MLE

Η δύναμη του MLE βρίσκεται στις ασυμπτωτικές του ιδιότητες — γι' αυτό κυριαρχεί στην οικονομετρία.

Θεώρημα 2.4 (Ασυμπτωτικές ιδιότητες MLE)

Κάτω από συνθήκες ομαλότητας, ο MLE $\hat{\theta}_{ML}$ από τυχαίο δείγμα μεγέθους n έχει τις εξής ιδιότητες:

1. **Συνέπεια:** $\hat{\theta}_{ML} \xrightarrow{\mathbb{P}} \theta$.
2. **Ασυμπτωτική κανονικότητα:**

$$\sqrt{n}(\hat{\theta}_{ML} - \theta) \xrightarrow{d} N\left(0, \frac{1}{I_1(\theta)}\right).$$

3. **Ασυμπτωτική αποτελεσματικότητα:** η ασυμπτωτική διασπορά $\frac{1}{nI_1(\theta)}$ ισούται με το CRLB. Δηλαδή ο MLE είναι ασυμπτωτικά ο βέλτιστος δυνατός εκτιμητής.

Γιατί ο MLE είναι «βασιλιάς»

Καμία άλλη γενική μέθοδος δεν προσφέρει αυτόματα και τις τρεις ιδιότητες. Ο MLE μπορεί να είναι μεροληπτικός σε μικρά δείγματα, αλλά ασυμπτωτικά είναι αμερόληπτος, συνεπής, και αποτελεσματικός (πάνει το CRLB). Γι' αυτό σχεδόν κάθε σύγχρονο οικονομετρικό μοντέλο εκτιμάται με MLE.

3.3 Ιδιότητα Αναλλοιότητας (Invariance)**Θεώρημα 2.5 (Αναλλοιότητα του MLE)**

Αν $\hat{\theta}_{ML}$ είναι ο MLE του θ , τότε για κάθε συνάρτηση g , ο MLE της παραμέτρου $\eta = g(\theta)$ είναι $g(\hat{\theta}_{ML})$:

$$\hat{\eta}_{ML} = g(\hat{\theta}_{ML}).$$

Παρατήρηση

Πρακτικά: για να εκτιμήσουμε μια *συνάρτηση* της παραμέτρου (π.χ. την πιθανότητα μηδενικών αξιώσεων $e^{-\lambda}$ σε Poisson, ή την ελαστικότητα σε ένα μοντέλο), δεν χρειάζεται να ξαναλύσουμε το πρόβλημα μεγιστοποίησης — απλώς εφαρμόζουμε τη συνάρτηση στον υπάρχοντα MLE. Η ιδιότητα αυτή είναι μοναδική του MLE (η MoM δεν την έχει γενικά).

3.4 Η Μέθοδος Delta

Η ιδιότητα αναλλοιότητας μάς δίνει την *εκτίμηση* μιας συνάρτησης $g(\theta)$. Αλλά πώς βρίσκουμε την *κατανομή* (άρα και το τυπικό σφάλμα) του $g(\hat{\theta})$; Η απάντηση είναι η μέθοδος Delta.

Θεώρημα 2.6 (Μέθοδος Delta)

Έστω εκτιμητής $\hat{\theta}$ με

$$\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{d} N(0, \sigma_\theta^2).$$

Αν g είναι παραγωγίσιμη συνάρτηση με $g'(\theta) \neq 0$, τότε

$$\sqrt{n}(g(\hat{\theta}) - g(\theta)) \xrightarrow{d} N(0, [g'(\theta)]^2 \sigma_\theta^2).$$

Διαισθητικά (γραμμικοποίηση Taylor)

Πρώτης τάξης ανάπτυγμα Taylor γύρω από το θ :

$$g(\hat{\theta}) \approx g(\theta) + g'(\theta)(\hat{\theta} - \theta).$$

Άρα το $g(\hat{\theta})$ είναι (περίπου) γραμμικός μετασχηματισμός του $\hat{\theta}$ · πολλαπλασιάζοντας μια κανονική τ.μ. με τη σταθερά $g'(\theta)$, η διασπορά πολλαπλασιάζεται με $[g'(\theta)]^2$.

Σύνδεση με την Οικονομετρία

Η μέθοδος Delta είναι παντού στην εφαρμοσμένη οικονομετρία: τυπικά σφάλματα για ελαστικότητες (μη γραμμικές συναρτήσεις των $\hat{\beta}$), για αναλογίες συντελεστών, για marginal effects σε logit/probit, για τον υπολογισμό του long-run multiplier σε δυναμικά μοντέλα. Η εντολή nlcom στο Stata και η deltamethod στην R υλοποιούν ακριβώς αυτό το θεώρημα.

3.5 Σύντομη επαφή: Γενικευμένη Μέθοδος Ροπών (GMM)

Κλείνουμε με μια γέφυρα προς την Οικονομετρία II. Η **Γενικευμένη Μέθοδος Ροπών** (Generalized Method of Moments — GMM) γενικεύει τη Μέθοδο Ροπών της Εβδομάδας 1.

Ορισμός 2.8 (Συνθήκες ροπών και GMM)

Έστω ότι το οικονομικό μοντέλο συνεπάγεται **συνθήκες ροπών** (moment conditions):

$$\mathbb{E}[m(X_i, \theta_0)] = 0,$$

όπου θ_0 η αληθινή τιμή και m ένα γνωστό διάνυσμα συναρτήσεων. Ο εκτιμητής GMM ορίζεται ως

$$\hat{\theta}_{\text{GMM}} = \operatorname{argmin}_{\theta} \bar{m}(\theta)' W \bar{m}(\theta), \quad \bar{m}(\theta) = \frac{1}{n} \sum_{i=1}^n m(X_i, \theta),$$

όπου W ένας θετικά ορισμένος πίνακας στάθμισης.

Πώς συνδέεται με όσα ξέρετε

- Αν οι συνθήκες ροπών είναι ακριβώς όσες οι παράμετροι (*just-identified*), τότε $\bar{m}(\hat{\theta}) = 0$ και η GMM ταυτίζεται με την κλασική Μέθοδο Ροπών της Εβδομάδας 1.
- Αν οι συνθήκες είναι περισσότερες από τις παραμέτρους (*over-identified*), δεν μπορούμε να τις μηδενίσουμε όλες — τις «ζυγίζουμε» με τον W και ελαχιστοποιούμε.
- Ο MLE είναι ειδική περίπτωση GMM με συνθήκη ροπής τη score: $\mathbb{E}[S(\theta_0)] = 0$ (Θεώρημα 2.1).

Σύνδεση με την Οικονομετρία

Η GMM είναι το κυρίαρχο πλαίσιο στη σύγχρονη μακροοικονομετρία και χρηματοοικονομική οικονομετρία (εκτίμηση Euler equations, asset pricing models, instrumental variables, dynamic panel — Arellano-Bond). Ο Lars Peter Hansen τιμήθηκε με το Νόμπελ Οικονομίας 2013 για την ανάπτυξη της (1982). Στην Οικονομετρία II θα τη δείτε αναλυτικά — αλλά η αρχή της αντιστοιχίας που μάθατε στην Εβδομάδα 1 είναι ακριβώς η ίδια.

4 Λυμένες Ασκήσεις

4.1 Ομάδα Γ: Μέγιστη Πιθανοφάνεια

Άσκηση 5: MLE για Bernoulli — πλήρης εξαγωγή + ασυμπτωτική

Τυχαιο δείγμα $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} \text{Bernoulli}(p)$ (επιτυχία πώλησης).

- Γράψτε την πιθανοφάνεια και τη λογαριθμική πιθανοφάνεια.
- Βρείτε τον MLE $\hat{p}_{\text{ML}}$ μέσω της συνθήκης πρώτης τάξης.
- Επαληθεύστε με τη συνθήκη δεύτερης τάξης ότι είναι μέγιστο.
- Γράψτε την ασυμπτωτική κατανομή του $\hat{p}_{\text{ML}}$ χρησιμοποιώντας την πληροφορία Fisher (Άσκηση 1).

Λύση

(α)

$$L(p) = \prod_{i=1}^n p^{x_i} (1-p)^{1-x_i} = p^{\sum x_i} (1-p)^{n-\sum x_i}.$$

$$\ell(p) = \left(\sum x_i\right) \ln p + \left(n - \sum x_i\right) \ln(1-p).$$

(β) Συνθήκη πρώτης τάξης:

$$\ell'(p) = \frac{\sum x_i}{p} - \frac{n - \sum x_i}{1-p} = 0.$$

Πολλαπλασιάζουμε με $p(1-p)$:

$$(1-p) \sum x_i - p \left(n - \sum x_i \right) = 0 \implies \sum x_i - p \sum x_i - pn + p \sum x_i = 0 \implies \sum x_i = pn.$$

Άρα

$$\hat{p}_{\text{ML}} = \frac{\sum x_i}{n} = \bar{X}.$$

(γ) Δεύτερη παράγωγος:

$$\ell''(p) = -\frac{\sum x_i}{p^2} - \frac{n - \sum x_i}{(1-p)^2} < 0 \quad \text{για κάθε } p \in (0, 1).$$

Αρνητική παντού $\rightarrow$ πρόκειται για (ολικό) μέγιστο. ✓

(δ) Από την Άσκηση 1, $I_1(p) = \frac{1}{p(1-p)}$. Από το Θεώρημα 2.4:

$$\sqrt{n}(\hat{p}_{\text{ML}} - p) \xrightarrow{d} N\left(0, \frac{1}{I_1(p)}\right) = N(0, p(1-p)).$$

Δηλαδή, για μεγάλα n : $\hat{p}_{\text{ML}} \approx N\left(p, \frac{p(1-p)}{n}\right)$.

Παρατήρηση

Ο MLE της Bernoulli ταυτίζεται με τον MoM εκτιμητή (Εβδομάδα 1, Άσκηση 8): και οι δύο δίνουν $\bar{X}$. Αυτό συμβαίνει συχνά στις «απλές» κατανομές, αλλά δεν είναι κανόνας — στην Άσκηση 8 θα δούμε ένα παράδειγμα όπου MLE και MoM διαφέρουν δραματικά.

Άσκηση 6: MLE για Εκθετική — διάρκεια αποπληρωμής δανείου

Η διάρκεια (έτη) μέχρι την αποπληρωμή δανείου είναι $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} \text{Exp}(\lambda)$, $f(x; \lambda) = \lambda e^{-\lambda x}$, $x > 0$.

(α') Βρείτε τον MLE $\hat{\lambda}_{\text{ML}}$.

(β') Υπολογίστε την πληροφορία Fisher $I_1(\lambda)$ και γράψτε την ασυμπτωτική κατανομή του MLE.

(γ') Σε δείγμα $n = 40$ δανείων, ο μέσος χρόνος αποπληρωμής βρέθηκε $\bar{x} = 5$ έτη. Δώστε την εκτίμηση και ένα κατά προσέγγιση τυπικό σφάλμα.

Λύση

(α)

$$L(\lambda) = \prod_{i=1}^n \lambda e^{-\lambda x_i} = \lambda^n e^{-\lambda \sum x_i}, \quad \ell(\lambda) = n \ln \lambda - \lambda \sum x_i.$$

Συνθήκη πρώτης τάξης:

$$\ell'(\lambda) = \frac{n}{\lambda} - \sum x_i = 0 \implies \hat{\lambda}_{\text{ML}} = \frac{n}{\sum x_i} = \frac{1}{\bar{X}}.$$

Συνθήκη δεύτερης τάξης: $\ell''(\lambda) = -n/\lambda^2 < 0 \rightarrow$ μέγιστο. ✓

(β) Πληροφορία Fisher:

$$\ln f = \ln \lambda - \lambda x, \quad \frac{\partial \ln f}{\partial \lambda} = \frac{1}{\lambda} - x, \quad \frac{\partial^2 \ln f}{\partial \lambda^2} = -\frac{1}{\lambda^2}.$$

$$I_1(\lambda) = -\mathbb{E}\left[-\frac{1}{\lambda^2}\right] = \frac{1}{\lambda^2}.$$

Άρα από το Θεώρημα 2.4:

$$\sqrt{n}(\hat{\lambda}_{\text{ML}} - \lambda) \xrightarrow{d} N(0, \lambda^2).$$

Ισοδύναμα, $\hat{\lambda}_{\text{ML}} \approx N(\lambda, \lambda^2/n)$ για μεγάλα n .

(γ) Εκτίμηση:

$$\hat{\lambda}_{\text{ML}} = \frac{1}{\bar{x}} = \frac{1}{5} = 0,20 \text{ ανά έτος.}$$

Κατά προσέγγιση τυπικό σφάλμα (αντικαθιστώντας λ με $\hat{\lambda}$):

$$\widehat{\text{SE}}(\hat{\lambda}_{\text{ML}}) \approx \frac{\hat{\lambda}}{\sqrt{n}} = \frac{0,20}{\sqrt{40}} \approx \frac{0,20}{6,325} \approx 0,0316.$$

Παρατήρηση

Ο MLE $\hat{\lambda} = 1/\bar{X}$ είναι ίδιος με τον MoM εκτιμητή της Εβδομάδας 1 (Άσκηση 10). Όπως είδαμε εκεί, *δεν είναι αμερόληπτος* (Jensen), αλλά είναι συνεπής και ασυμπτωτικά αποτελεσματικός — η ασυμπτωτική του διασπορά λ^2/n ισούται με το CRLB $1/(nI_1) = \lambda^2/n$.

Άσκηση 7: MLE για Κανονική — δύο παράμετροι

Έστω $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} N(\mu, \sigma^2)$ με αμφότερες παραμέτρους άγνωστες (αποδόσεις χαρτοφυλακίου).

(α') Γράψτε τη λογαριθμική πιθανοφάνεια.

(β') Βρείτε τους MLE $\hat{\mu}_{\text{ML}}$ και $\hat{\sigma}_{\text{ML}}^2$ λύνοντας το σύστημα των δύο συνθηκών πρώτης τάξης.

(γ') Συγκρίνετε τον $\hat{\sigma}_{\text{ML}}^2$ με τον αμερόληπτο S^2 της Εβδομάδας 1. Είναι ο MLE αμερόληπτος;

Λύση

(α)

$$\ell(\mu, \sigma^2) = -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln \sigma^2 - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2.$$

(β) Ως προς μ :

$$\frac{\partial \ell}{\partial \mu} = \frac{1}{\sigma^2} \sum_{i=1}^n (x_i - \mu) = 0 \implies \sum x_i = n\mu \implies \boxed{\hat{\mu}_{\text{ML}} = \bar{X}}.$$

Ως προς σ^2 (θεωρώντας σ^2 ως μία μεταβλητή):

$$\frac{\partial \ell}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_{i=1}^n (x_i - \mu)^2 = 0.$$

Πολλαπλασιάζουμε με $2\sigma^4$ και αντικαθιστούμε $\mu = \hat{\mu}_{\text{ML}} = \bar{X}$:

$$-n\sigma^2 + \sum (x_i - \bar{X})^2 = 0 \implies \boxed{\hat{\sigma}_{\text{ML}}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{X})^2}.$$

(γ) Ο MLE της διασποράς διαιρεί με n (όχι με $n-1$). Από την Εβδομάδα 1 (Άσκηση 2):

$$\mathbb{E}(\hat{\sigma}_{\text{ML}}^2) = \frac{n-1}{n} \sigma^2 \neq \sigma^2.$$

Άρα ο MLE της διασποράς είναι μεροληπτικός (υποεκτιμά τη σ^2 με $\text{bias} = -\sigma^2/n$). Όμως είναι ασυμπτωτικά αμερόληπτος ($\mathbb{E} \rightarrow \sigma^2$) και συνεπής. Ο διορθωμένος $S^2 = \frac{n}{n-1} \hat{\sigma}_{\text{ML}}^2$ είναι αμερόληπτος.

Παρατήρηση

Αυτή είναι μια κομβική παρατήρηση: **ο MLE δεν είναι γενικά αμερόληπτος σε πεπερασμένα δείγματα**. Εδώ ο MLE της σ^2 ταυτίζεται με τον MoM εκτιμητή (Εβδομάδα 1, Άσκηση 13) και είναι ο μεροληπτικός $\hat{\sigma}^2 = \frac{1}{n} \sum (X_i - \bar{X})^2$. Στην οικονομετρία, η εκτίμηση της διασποράς των καταλοίπων με MLE δίνει τον μεροληπτικό εκτιμητή· γι' αυτό τα λογισμικά εφαρμόζουν τη διόρθωση βαθμών ελευθερίας.

Άσκηση 8: MLE για Ομοιόμορφη

Χρόνος συναλλαγής $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} U(0, \theta)$, $\theta > 0$ άγνωστο. Η σ.π.π. είναι $f(x; \theta) = \frac{1}{\theta}$ για $0 \leq x \leq \theta$ (αλλιώς 0).

- (α') Γράψτε την πιθανοφάνεια προσεκτικά, λαμβάνοντας υπόψη τον περιορισμό $0 \leq x_i \leq \theta$.
- (β') Δείξτε ότι η συνθήκη πρώτης τάξης $\ell'(\theta) = 0$ δεν έχει λύση. Βρείτε τον MLE με επιχείρημα μονοτονίας.
- (γ') Συγκρίνετε με τον MoM εκτιμητή $2\bar{X}$ της Εβδομάδας 1 (Άσκηση 12). Ποιος είναι πιο «λογικός»;

Λύση

(α) Κάθε όρος είναι $1/\theta$ μόνο αν $0 \leq x_i \leq \theta$. Με δείκτρια συνάρτηση:

$$L(\theta) = \prod_{i=1}^n \frac{1}{\theta} \mathbb{1}(0 \leq x_i \leq \theta) = \frac{1}{\theta^n} \mathbb{1}\left(\theta \geq \max_i x_i\right).$$

(Όλα τα $x_i \leq \theta$ ισοδυναμεί με $\max_i x_i \leq \theta$ · υποθέτουμε $x_i \geq 0$.)

(β) Για $\theta \geq \max_i x_i$, η πιθανοφάνεια είναι $L(\theta) = \theta^{-n}$, που είναι γνησίως φθίνουσα στο θ . Για $\theta < \max_i x_i$, είναι $L(\theta) = 0$. Άρα η L μεγιστοποιείται στο μικρότερο επιτρεπτό θ , δηλαδή:

$$\hat{\theta}_{\text{ML}} = \max_{1 \leq i \leq n} X_i = X_{(n)}.$$

Η συνθήκη πρώτης τάξης αποτυγχάνει: για $\theta > \max x_i$, $\ell'(\theta) = -n/\theta \neq 0$ — δεν μηδενίζεται ποτέ. Το μέγιστο είναι στο άκρο του επιτρεπτού χώρου, όχι σε στάσιμο σημείο.

(γ) Ο MoM εκτιμητής ήταν $2\bar{X}$, που (όπως είδαμε στην Εβδομάδα 1) μπορεί να βγει μικρότερος από $\max x_i$ — δηλαδή να δηλώνει ότι μια παρατήρηση που πράγματι συνέβη ήταν «αδύνατη». Ο MLE $X_{(n)}$ είναι πάντοτε $\geq \max x_i$, άρα πάντα λογικός.

Σημείωση για μεροληψία: ο $\hat{\theta}_{\text{ML}} = X_{(n)}$ είναι μεροληπτικός: μπορεί να αποδειχθεί ότι $\mathbb{E}(X_{(n)}) = \frac{n}{n+1}\theta < \theta$ (πάντα υποεκτιμά, αφού ο μέγιστος του δείγματος δεν μπορεί να ξεπεράσει το αληθινό θ). Είναι όμως συνεπής, και ο διορθωμένος $\frac{n+1}{n}X_{(n)}$ είναι αμερόληπτος.

Παρατήρηση

Αυτή η άσκηση «κλειδώνει» δύο μαθήματα: (1) η μέθοδος FOC δεν είναι πανάκεια — πρέπει πάντα να ελέγχουμε τη φύση του παραμετρικού χώρου· (2) ο MLE σέβεται τους λογικούς περιορισμούς του μοντέλου, ενώ η MoM όχι πάντα. Η $U(0, \theta)$ είναι το κλασικό «παθολογικό» παράδειγμα και εμφανίζεται σχεδόν σε κάθε εξεταστική.

4.2 Ομάδα Δ: Αναλλοιότητα και Μέθοδος Delta

Άσκηση 9: Αναλλοιότητα — πιθανότητα μηδενικών αξιώσεων

Αξιώσεις ασφάλισης ανά μήνα: $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} \text{Poisson}(\lambda)$. Ο ασφαλιστής ενδιαφέρεται για την πιθανότητα ενός μήνα χωρίς καμία αξίωση, δηλαδή

$$\eta = \mathbb{P}(X = 0) = e^{-\lambda}.$$

(α') Χρησιμοποιώντας την ιδιότητα αναλλοιότητας, βρείτε τον MLE του η .

(β') Σε δείγμα $n = 24$ μηνών βρέθηκε $\bar{x} = 1,5$. Δώστε την εκτίμηση του η .

Λύση

(α) Από την Εβδομάδα 1 / Άσκηση 2 (αυτής της εβδομάδας), ο MLE του λ για Poisson είναι $\hat{\lambda}_{\text{ML}} = \bar{X}$. Η συνάρτηση είναι $g(\lambda) = e^{-\lambda}$. Από το Θεώρημα 2.5 (αναλλοιότητα):

$$\hat{\eta}_{\text{ML}} = g(\hat{\lambda}_{\text{ML}}) = e^{-\bar{X}}.$$

(β) Με $\bar{x} = 1,5$:

$$\hat{\eta}_{\text{ML}} = e^{-1,5} \approx 0,2231.$$

Δηλαδή εκτιμάται ότι περίπου 22,3% των μηνών δεν θα έχουν καμία αξίωση.

Άσκηση 10: Μέθοδος Delta — ασυμπτωτική κατανομή του $e^{-\hat{\lambda}}$

Συνεχίζοντας την Άσκηση 9, θέλουμε όχι μόνο την εκτίμηση αλλά και την ασυμπτωτική κατανομή (άρα και τυπικό σφάλμα) του $\hat{\eta}_{ML} = e^{-\bar{X}}$.

(α') Γράψτε την ασυμπτωτική κατανομή του $\hat{\lambda}_{ML} = \bar{X}$.

(β') Εφαρμόστε τη μέθοδο Delta με $g(\lambda) = e^{-\lambda}$ για να βρείτε την ασυμπτωτική κατανομή του $\hat{\eta}_{ML}$.

(γ') Με $n = 24$, $\bar{x} = 1,5$, υπολογίστε ένα κατά προσέγγιση τυπικό σφάλμα για το $\hat{\eta}_{ML}$.

Λύση

(α) Από την Άσκηση 2 (πληροφορία Fisher Poisson $I_1 = 1/\lambda$, ή απευθείας από ΚΟΘ αφού $\text{Var}(X) = \lambda$):

$$\sqrt{n}(\hat{\lambda}_{ML} - \lambda) \xrightarrow{d} N(0, \lambda).$$

Δηλαδή $\sigma_{\hat{\lambda}}^2 = \lambda$.

(β) Η συνάρτηση είναι $g(\lambda) = e^{-\lambda}$, με παράγωγο $g'(\lambda) = -e^{-\lambda}$. Από το Θεώρημα 2.6 (Delta):

$$\sqrt{n}(e^{-\bar{X}} - e^{-\lambda}) \xrightarrow{d} N(0, [g'(\lambda)]^2 \sigma_{\hat{\lambda}}^2) = N(0, (e^{-\lambda})^2 \cdot \lambda).$$

Άρα

$$\sqrt{n}(\hat{\eta}_{ML} - e^{-\lambda}) \xrightarrow{d} N(0, \lambda e^{-2\lambda}).$$

Ισοδύναμα, η ασυμπτωτική διασπορά του $\hat{\eta}_{ML}$ είναι $\frac{\lambda e^{-2\lambda}}{n}$.

(γ) Αντικαθιστούμε λ με $\hat{\lambda} = \bar{x} = 1,5$ και $n = 24$:

$$\widehat{\text{Var}}(\hat{\eta}_{ML}) \approx \frac{\hat{\lambda} e^{-2\hat{\lambda}}}{n} = \frac{1,5 \cdot e^{-3}}{24} = \frac{1,5 \cdot 0,049787}{24} \approx \frac{0,074681}{24} \approx 0,0031117.$$

$$\widehat{SE}(\hat{\eta}_{ML}) \approx \sqrt{0,0031117} \approx 0,0558.$$

Δηλαδή $\hat{\eta}_{ML} \approx 0,223$ με τυπικό σφάλμα περίπου 0,056 (ένα κατά προσέγγιση 95% διάστημα θα ήταν περίπου $0,223 \pm 1,96 \times 0,056 \approx [0,113, 0,333]$).

Έλεγχος

Διαστατικός έλεγχος: το $e^{-\lambda}$ είναι πιθανότητα (αδιάστατο), λ είναι θετικό, οπότε η ασυμπτωτική διασπορά $\lambda e^{-2\lambda}/n > 0$ — λογικό. Επίσης για $\lambda \rightarrow 0$ ή $\lambda \rightarrow \infty$ η διασπορά $\rightarrow 0$ (στα άκρα η πιθανότητα $e^{-\lambda}$ τείνει σε 1 ή 0, δηλαδή σε σταθερά — μικρή αβεβαιότητα). Μέγιστη αβεβαιότητα γύρω από $\lambda = 0,5$. ✓

Σύνδεση με την Οικονομετρία

Οι Ασκήσεις 9-10 μαζί είναι το ακριβές μοτίβο που θα χρησιμοποιείτε στην Οικονομετρία: (1) εκτιμάτε μια παράμετρο με MLE, (2) η ποσότητα ενδιαφέρο-

ντος είναι μια μη γραμμική συνάρτηση (πιθανότητα, ελαστικότητα, marginal effect), (3) η αναλλοiotτητα δίνει την εκτίμηση, (4) η μέθοδος Delta δίνει το τυπικό σφάλμα για συμπερασματολογία.

5 Σύνοψη Τύπων

Σύνοψη Εβδομάδας 2

Πιθανοφάνεια:

$$L(\theta) = \prod_{i=1}^n f(x_i; \theta), \quad \ell(\theta) = \sum_{i=1}^n \ln f(x_i; \theta).$$

Score: $S(\theta) = \ell'(\theta)$, με $\mathbb{E}[\partial \ln f / \partial \theta] = 0$.

Πληροφορία Fisher (μία παρατήρηση):

$$I_1(\theta) = \mathbb{E} \left[\left(\frac{\partial \ln f}{\partial \theta} \right)^2 \right] = -\mathbb{E} \left[\frac{\partial^2 \ln f}{\partial \theta^2} \right], \quad I_n(\theta) = n I_1(\theta).$$

Cramer-Rao: για κάθε αμερόληπτο $\hat{\theta}$:

$$\text{Var}(\hat{\theta}) \geq \frac{1}{n I_1(\theta)} = \text{CRLB}.$$

Αποδοτικότητα: $\text{eff}(\hat{\theta}) = \text{CRLB} / \text{Var}(\hat{\theta}) \in (0, 1]$.

Επάρκεια (Fisher-Neyman): T επαρκές $\iff L(\theta) = g(T; \theta) h(\mathbf{x})$.

MLE: $\hat{\theta}_{\text{ML}} = \arg \max_{\theta} \ell(\theta)$. FOC: $\ell'(\theta) = 0$, SOC: $\ell''(\hat{\theta}) < 0$.

Ασυμπτωτικές ιδιότητες MLE:

$$\hat{\theta}_{\text{ML}} \xrightarrow{\mathbb{P}} \theta, \quad \sqrt{n}(\hat{\theta}_{\text{ML}} - \theta) \xrightarrow{d} N(0, 1/I_1(\theta)).$$

Αναλλοιότητα: MLE του $g(\theta)$ είναι $g(\hat{\theta}_{\text{ML}})$.

Μέθοδος Delta: αν $\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{d} N(0, \sigma_{\theta}^2)$, τότε

$$\sqrt{n}(g(\hat{\theta}) - g(\theta)) \xrightarrow{d} N(0, [g'(\theta)]^2 \sigma_{\theta}^2).$$

GMM: $\hat{\theta}_{\text{GMM}} =_{\theta} \bar{m}(\theta)' W \bar{m}(\theta)$, με $\mathbb{E}[m(X, \theta_0)] = 0$. MoM = just-identified GMM.

Πίνακας: Πληροφορία Fisher & MLE βασικών κατανομών

Κατανομή	$I_1(\theta)$	MLE	Ασυμπτ. διασπ. $\times n$
Bernoulli(p)	$\frac{1}{p(1-p)}$	$\bar{X}$	$p(1-p)$
Poisson(λ)	$\frac{1}{\lambda}$	$\bar{X}$	λ
Exp(rate λ)	$\frac{1}{\lambda^2}$	$1/\bar{X}$	λ^2
$N(\mu, \sigma^2)$, σ^2 γνωστή	$\frac{1}{\sigma^2}$	$\bar{X}$	σ^2
$U(0, \theta)$	— (FOC αποτυγχάνει)	$X_{(n)} = \max X_i$	—