

Στατιστική II

Θεμελίωση Εκτιμητών

Θεωρία και Ασκήσεις

Τμήμα Οικονομικής Επιστήμης, ΟΠΑ

Πρόλογος: Γιατί μελετάμε εκτιμητές

Σε κάθε εμπειρική εφαρμογή των οικονομικών επιστημών — από τον υπολογισμό του δείκτη ανεργίας από την ΕΛΣΤΑΤ μέχρι την αποτίμηση παραγώγων από ένα επενδυτικό τραπέζι — υπάρχει ένα κοινό μοτίβο: *δεν παρατηρούμε ποτέ ολόκληρο τον πληθυσμό*. Έχουμε στη διάθεσή μας ένα πεπερασμένο δείγμα και πρέπει να χρησιμοποιήσουμε αυτό το δείγμα για να βγάλουμε συμπεράσματα για ολόκληρη την οικονομία, αγορά ή πληθυσμό.

Η *σημειακή εκτίμηση* είναι η συστηματική τεχνική για να μετατρέπουμε δεδομένα σε εκτιμήσεις παραμέτρων. Αλλά δεν είναι όλες οι τεχνικές το ίδιο καλές: κάποιες παράγουν συστηματικά λάθος εκτιμήσεις (μεροληπτικές), άλλες δίνουν εκτιμήσεις που «πετάγονται» σε κάθε δείγμα (μεγάλη διασπορά), και άλλες δεν συγκλίνουν ποτέ στην αληθινή τιμή ακόμη και με τεράστια δείγματα (μη συνεπείς).

Σε αυτή την εβδομάδα θα μάθουμε:

1. Τι είναι ένας εκτιμητής (και τι τον διακρίνει από μια εκτίμηση).
2. Ποιες ιδιότητες κάνουν έναν εκτιμητή «καλό»: *αμεροληψία, μικρή διασπορά, χαμηλό MSE, συνέπεια, αποτελεσματικότητα*.
3. Πώς μπορούμε συστηματικά να κατασκευάζουμε εκτιμητές μέσω της *Μεθόδου των Ροπών*.

Ιστορική σημείωση. Η θεωρία εκτίμησης όπως τη γνωρίζουμε σήμερα αναπτύχθηκε κυρίως στο τέλος του 19ου και τις αρχές του 20ού αιώνα. Ο Karl Pearson (1857–1936) εισήγαγε τη Μέθοδο των Ροπών το 1894 ως την πρώτη συστηματική τεχνική κατασκευής εκτιμητών. Ο Sir Ronald Fisher (1890–1962), σε μια σειρά εργασιών κατά τη δεκαετία του 1920, ανέπτυξε το μεγαλύτερο μέρος των ιδιοτήτων εκτιμητών (αμεροληψία, συνέπεια, αποτελεσματικότητα, επάρκεια) μαζί με τη Μέθοδο Μέγιστης Πιθανοφάνειας (που θα δούμε στην Εβδομάδα 2). Η οικονομετρία ως ξεχωριστή πειθαρχία, με τους Frisch, Haavelmo και Koopmans, βασίστηκε εξ ολοκλήρου σε αυτό το πλαίσιο.

Προϋποθέσεις. Πιθανότητες, τυχαίες μεταβλητές, μέση τιμή, διασπορά, βασικές κατανομές (Bernoulli, Binomial, Geometric, Poisson, Εκθετική, Ομοιόμορφη, Κανονική, Γάμμα), βασικές ανισότητες (Markov, Chebyshev).

Περιεχόμενα

Πρόλογος	1
1 Θεωρία	2
1.1 Στατιστική Συμπερασματολογία	2
1.2 Εκτιμητής vs Εκτίμηση και η Δειγματοληπτική Κατανομή	3
1.3 Αμεροληψία	4
1.4 Διασπορά Εκτιμητή και Τυπικό Σφάλμα	6

1.5 Μέσο Τετραγωνικό Σφάλμα (MSE)	6
1.6 Συνέπεια	7
1.7 Σχετική Αποτελεσματικότητα και UMVUE	9
1.8 Μέθοδος των Ροπών	10
2 Λυμένες Ασκήσεις	11
2.1 Ομάδα Α: Ιδιότητες Εκτιμητών	11
2.2 Ομάδα Β: Μέθοδος των Ροπών — Μονοπαραμετρικές Κατανομές	19
2.3 Ομάδα Γ: Μέθοδος των Ροπών — Διπαραμετρικές Κατανομές	23
3 Σύνοψη Τύπων	27
Τι ακολουθεί στην Εβδομάδα 2	28

1 Θεωρία

1.1 Στατιστική Συμπερασματολογία

Ορισμός 1.1 (Στατιστική συμπερασματολογία)

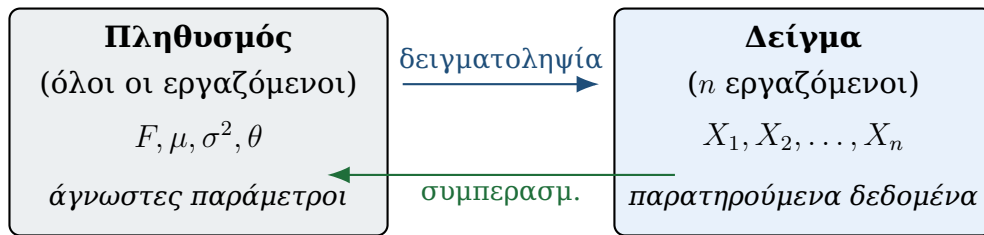
Η **στατιστική συμπερασματολογία** (statistical inference) είναι η διαδικασία με την οποία αντλούμε συμπεράσματα για άγνωστα χαρακτηριστικά ενός **πληθυσμού**, χρησιμοποιώντας πληροφορίες από ένα **δείγμα**.

Βασική ορολογία:

- **Πληθυσμός** (population): το σύνολο όλων των μονάδων ενδιαφέροντος. Στα οικονομικά, ένας «πληθυσμός» μπορεί να είναι όλοι οι μισθωτοί εργαζόμενοι μιας χώρας, όλα τα ασφαλιστήρια συμβόλαια μιας εταιρείας, όλες οι ημερήσιες αποδόσεις μιας μετοχής σε μια χρονική περίοδο, κ.λπ.
- **Παράμετρος** (parameter): ένα άγνωστο σταθερό χαρακτηριστικό του πληθυσμού, που συμβολίζεται συνήθως με θ . Παραδείγματα:
 - ο μέσος μηνιαίος μισθός μ ,
 - η αναλογία p δανείων που πέφτουν σε αθέτηση,
 - η διασπορά σ^2 των ημερήσιων αποδόσεων μιας μετοχής,
 - η ελαστικότητα η της ζήτησης ως προς την τιμή,
 - ο ρυθμός λ απαιτήσεων ασφάλισης ανά μήνα.

Η παράμετρος είναι *σταθερή* (αλλά άγνωστη). Δεν τυχαιοποιείται — αυτό που τυχαιοποιείται είναι το δείγμα μας.

- **Δείγμα** (sample): πεπερασμένο υποσύνολο $X_1, X_2, \dots, X_n$ τυχαίων μεταβλητών από τον πληθυσμό. Όταν αυτές είναι ανεξάρτητες και ισόνομες (independent and identically distributed), λέμε ότι έχουμε *τυχαίο δείγμα μεγέθους n* , και γράφουμε $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} F$, όπου F η κατανομή του πληθυσμού.
- **Στατιστικό** (statistic): μια οποιαδήποτε συνάρτηση $T(X_1, \dots, X_n)$ του δείγματος που *δεν εξαρτάται από άγνωστες παραμέτρους*. Παραδείγματα: ο δειγματικός μέσος $\bar{X}$, η δειγματική διασπορά S^2 , ο δειγματικός μέγιστος $\max_i X_i$, ο δειγματικός διάμεσος.



Η συμπερασματολογία χωρίζεται σε δύο μεγάλες κατηγορίες:

1. **Εκτίμηση** (estimation): επιλογή μιας τιμής (ή διαστήματος τιμών) για την άγνωστη παράμετρο.
2. **Έλεγχος υποθέσεων** (hypothesis testing): απόφαση μεταξύ ανταγωνιστικών ισχυρισμών για την παράμετρο (βλ. Εβδομάδα 3).

Η εκτίμηση με τη σειρά της διακρίνεται σε:

- **Σημειακή εκτίμηση** (point estimation): επιστροφή μίας μόνο αριθμητικής τιμής ως εκτίμηση της παραμέτρου.
- **Εκτίμηση με διαστήματα εμπιστοσύνης** (interval estimation): επιστροφή ενός διαστήματος που, με προκαθορισμένη πιθανότητα, καλύπτει την παράμετρο.

Σε όλη την Εβδομάδα 1 ασχολούμαστε αποκλειστικά με τη σημειακή εκτίμηση.

Σύνδεση με την Οικονομετρία

Στην Οικονομετρία I θα δείτε ότι ο εκτιμητής Ελαχίστων Τετραγώνων (OLS) είναι ακριβώς μια σημειακή εκτίμηση των συντελεστών παλινδρόμησης $\beta_0, \beta_1, \dots, \beta_k$ — και όλες οι ιδιότητες που θα μελετήσουμε εδώ (αμεροληψία, διασπορά, συνέπεια, αποτελεσματικότητα) θα ξαναεμφανιστούν σε εκείνο το πλαίσιο, σχεδόν αυτούσιες.

1.2 Εκτιμητής vs Εκτίμηση και η Δειγματοληπτική Κατανομή

Ορισμός 1.2 (Εκτιμητής)

Ένας **εκτιμητής** (estimator) της παραμέτρου θ είναι ένα στατιστικό:

$$\hat{\theta} = T(X_1, X_2, \dots, X_n),$$

δηλαδή μια *τυχαία μεταβλητή* (συνάρτηση του τυχαίου δείγματος). Η συγκεκριμένη αριθμητική τιμή που παίρνει για συγκεκριμένα δεδομένα $x_1, \dots, x_n$ ονομάζεται **εκτίμηση** (estimate) και συμβολίζεται $\hat{\theta}(x_1, \dots, x_n)$ ή απλώς $\hat{\theta}_{\text{obs}}$.

Διάκριση που έχει σημασία

Εκτιμητής = τυχαία μεταβλητή (έχει κατανομή).

Εκτίμηση = αριθμός (μία πραγματοποίηση αυτής της τυχαίας μεταβλητής για συγκεκριμένο δείγμα).

Π.χ. ο εκτιμητής του μέσου μισθού είναι ο τυχαίος $\bar{X} = \frac{1}{n} \sum X_i$. Η εκτίμηση για το συγκεκριμένο δείγμα 30 εργαζομένων που μελέτησε ο αναλυτής μπορεί να είναι $\bar{x} = 1.450 \text{ €}$. Αν ο ίδιος αναλυτής έπαιρνε ένα διαφορετικό τυχαίο δείγμα 30 εργαζομένων, θα έπαιρνε διαφορετική εκτίμηση (π.χ. 1.482 €), παρόλο που η αληθινή παράμετρος μ παραμένει η ίδια.

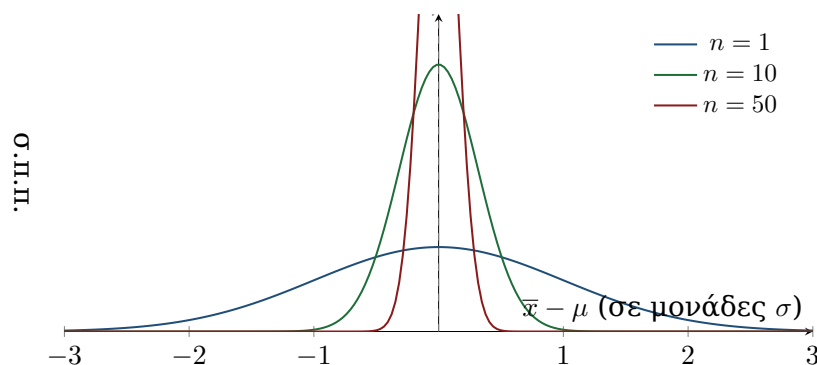
Ορισμός 1.3 (Δειγματοληπτική κατανομή)

Επειδή ο εκτιμητής $\hat{\theta}$ είναι τυχαία μεταβλητή, έχει δική του κατανομή πιθανότητας, η οποία προκύπτει από την κατανομή του δείγματος. Αυτή ονομάζεται **δειγματοληπτική κατανομή** (sampling distribution) του εκτιμητή.

Παράδειγμα — αναλυτικά: Αν $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} N(\mu, \sigma^2)$, τότε

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right).$$

Η δειγματοληπτική κατανομή του $\bar{X}$ είναι λοιπόν κανονική με κέντρο μ και διασπορά σ^2/n . Όσο μεγαλώνει το n , η κατανομή «μαζεύεται» γύρω από τη μ — αυτό είναι η οπτική διαίσθηση για τη συνέπεια (παρ. 1.6).



Σχήμα 1: Δειγματοληπτική κατανομή του $\bar{X}$ για $n = 1, 10, 50$ από κανονικό πληθυσμό. Όσο μεγαλώνει το n , η κατανομή συγκεντρώνεται γύρω από τη μ .

Σύνδεση με την Οικονομετρία

Στην Οικονομετρία, η δειγματοληπτική κατανομή του OLS εκτιμητή $\hat{\beta}$ είναι αυτή που νομιμοποιεί ολόκληρη τη συμπερασματολογία: τα τυπικά σφάλματα, τα διαστήματα εμπιστοσύνης, τα t -στατιστικά, τα p -values — όλα αντλούν νόημα από την κατανομή που έχει το $\hat{\beta}$ ως τυχαία μεταβλητή.

1.3 Αμεροληψία**Ορισμός 1.4 (Αμεροληψία)**

Ο εκτιμητής $\hat{\theta}$ ονομάζεται **αμερόληπτος** (unbiased) για την παράμετρο θ αν

$$\mathbb{E}(\hat{\theta}) = \theta, \quad \text{για κάθε τιμή του } \theta.$$

Η ποσότητα

$$\text{bias}(\hat{\theta}) = \mathbb{E}(\hat{\theta}) - \theta$$

ονομάζεται **σφάλμα μεροληψίας** (bias) του εκτιμητή.

- Αν $\text{bias}(\hat{\theta}) > 0$, ο εκτιμητής *υπερεκτιμά* την παράμετρο κατά μέσο όρο.
- Αν $\text{bias}(\hat{\theta}) < 0$, ο εκτιμητής *υποεκτιμά* την παράμετρο κατά μέσο όρο.
- Αν $\text{bias}(\hat{\theta}) = 0$, ο εκτιμητής είναι αμερόληπτος.

Διαισθητική ερμηνεία

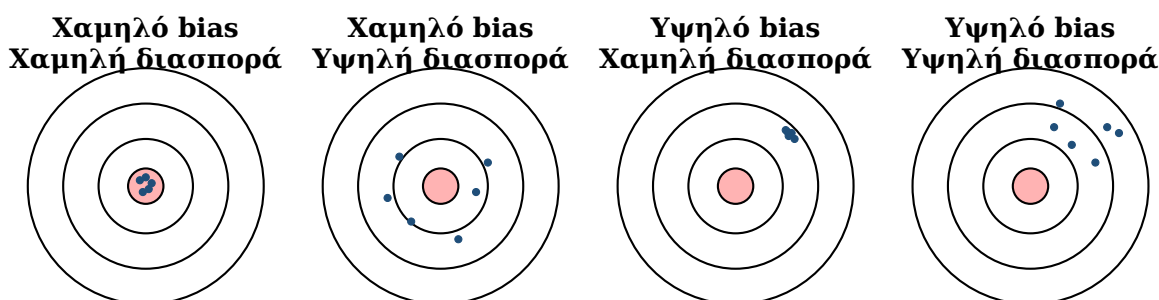
Αν επαναλάβουμε νοητά τη δειγματοληψία πάρα πολλές φορές και κάθε φορά υπολογίσουμε την τιμή $\hat{\theta}$ για το δείγμα μας, ο μέσος όρος αυτών των τιμών συγκλίνει στην $\mathbb{E}(\hat{\theta})$. Αν ο εκτιμητής είναι αμερόληπτος, αυτός ο μέσος όρος ταυτίζεται με την αληθινή τιμή της παραμέτρου.

Προσοχή: αμεροληψία **δεν** σημαίνει ότι η εκτίμηση από το συγκεκριμένο δείγμα που έχουμε είναι ίση με θ . Σημαίνει ότι «κατά μέσο όρο σε επαναλαμβανόμενα δείγματα» είναι σωστή.

Κλασικά παραδείγματα:

- Ο δειγματικός μέσος $\bar{X} = \frac{1}{n} \sum X_i$ είναι πάντοτε αμερόληπτος για το $\mu = \mathbb{E}(X_i)$ — αρκεί η μέση τιμή του πληθυσμού να υπάρχει.
- Η διορθωμένη δειγματική διασπορά $S^2 = \frac{1}{n-1} \sum (X_i - \bar{X})^2$ είναι αμερόληπτη για το $\sigma^2 = \text{Var}(X_i)$.
- Η μη διορθωμένη δειγματική διασπορά $\hat{\sigma}^2 = \frac{1}{n} \sum (X_i - \bar{X})^2$ είναι *μεροληπτική*: $\mathbb{E}(\hat{\sigma}^2) = \frac{n-1}{n} \sigma^2 < \sigma^2$.
- Η δειγματική τυπική απόκλιση $S = \sqrt{S^2}$ είναι *μεροληπτική* για το σ — ακόμη και όταν το S^2 είναι αμερόληπτο. Αυτό προκύπτει από την ανισότητα Jensen ($\sqrt{\cdot}$ είναι κοίλη).

Η αναλογία του στόχου: Η πιο διαισθητική απεικόνιση του ζεύγους *μεροληψία-διασπορά* είναι ο σκοπευτικός στόχος (dartboard):



Σχήμα 2: Η αναλογία του σκοπευτικού στόχου. Το κέντρο = αληθινή τιμή θ . Κάθε βολή = μία εκτίμηση από διαφορετικό δείγμα. *Bias* = πόσο μακριά είναι ο μέσος όρος των βολών από το κέντρο. *Διασπορά* = πόσο σκορπισμένες είναι μεταξύ τους.

Δεν είναι όλα τα bias κακά

Η αμεροληψία είναι σημαντική, αλλά δεν είναι το μόνο κριτήριο. Όπως θα δούμε στην Άσκηση 2, υπάρχουν περιπτώσεις όπου ένας *μεροληπτικός* εκτιμητής έχει μικρότερο συνολικό σφάλμα από έναν αμερόληπτο. Αυτή η ιδέα είναι η καρδιά της σύγχρονης Στατιστικής Μάθησης (Machine Learning) και της κανονικοποιημένης παλινδρόμησης (Ridge, Lasso) στην οικονομετρία.

1.4 Διασπορά Εκτιμητή και Τυπικό Σφάλμα

Ορισμός 1.5 (Διασπορά εκτιμητή)

Η **διασπορά** του εκτιμητή είναι

$$\text{Var}(\hat{\theta}) = \mathbb{E}[(\hat{\theta} - \mathbb{E}(\hat{\theta}))^2].$$

Η τετραγωνική της ρίζα ονομάζεται **τυπικό σφάλμα** (standard error):

$$\text{SE}(\hat{\theta}) = \sqrt{\text{Var}(\hat{\theta})}.$$

Γιατί δεν αρκεί η αμεροληψία

Δύο εκτιμητές μπορεί να είναι αμρόληπτοι, αλλά ο ένας να έχει πολύ μεγαλύτερη διασπορά. Ο εκτιμητής με τη *μικρότερη* διασπορά είναι προτιμότερος, γιατί δίνει εκτιμήσεις που «πέφτουν» πιο κοντά στην πραγματική τιμή της παραμέτρου σε επαναλαμβανόμενα δείγματα.

Με όρους του Σχήματος 2: ανάμεσα στα δύο αμερόληπτα διαγράμματα (πρώτο και δεύτερο), προτιμάμε το πρώτο γιατί έχει σφιχτότερη ομαδοποίηση βολών γύρω από το κέντρο.

Παράδειγμα 1.5.1 — δύο αμερόληπτοι εκτιμητές με διαφορετική διασπορά. Έστω X_1, X_2, X_3 τυχαίο δείγμα από κατανομή με μέση τιμή μ και διασπορά σ^2 . Τόσο ο

$$\hat{\mu}_1 = \bar{X} = \frac{X_1 + X_2 + X_3}{3}$$

όσο και ο

$$\hat{\mu}_2 = \frac{X_1 + X_3}{2}$$

είναι αμερόληπτοι για το μ (επαλήθευση: $\mathbb{E}(\hat{\mu}_1) = \mathbb{E}(\hat{\mu}_2) = \mu$). Όμως

$$\text{Var}(\hat{\mu}_1) = \frac{\sigma^2}{3} \approx 0,333\sigma^2, \quad \text{Var}(\hat{\mu}_2) = \frac{\sigma^2}{2} = 0,5\sigma^2.$$

Ο $\hat{\mu}_1$ έχει 33% μικρότερη διασπορά. Διαισθητικά, ο $\hat{\mu}_1$ χρησιμοποιεί *όλη* τη διαθέσιμη πληροφορία, ενώ ο $\hat{\mu}_2$ «πετάει» την παρατήρηση X_2 .

1.5 Μέσο Τετραγωνικό Σφάλμα (MSE)

Το *Μέσο Τετραγωνικό Σφάλμα* ενοποιεί τα δύο κριτήρια (αμεροληψία & διασπορά) σε ένα μέγεθος.

Ορισμός 1.6 (MSE)

Το **μέσο τετραγωνικό σφάλμα** του $\hat{\theta}$ ορίζεται ως

$$\text{MSE}(\hat{\theta}) = \mathbb{E}[(\hat{\theta} - \theta)^2].$$

Θεώρημα 1.1 (Αποσύνθεση Bias-Variance)

Για κάθε εκτιμητή $\hat{\theta}$ ισχύει:

$$\text{MSE}(\hat{\theta}) = \text{Var}(\hat{\theta}) + [\text{bias}(\hat{\theta})]^2.$$

Απόδειξη. Θέτουμε $\mu_{\hat{\theta}} = \mathbb{E}(\hat{\theta})$. Προσθαφαιρούμε στον τύπο του MSE:

$$\begin{aligned}\text{MSE}(\hat{\theta}) &= \mathbb{E}[(\hat{\theta} - \theta)^2] \\ &= \mathbb{E}[(\hat{\theta} - \mu_{\hat{\theta}} + (\mu_{\hat{\theta}} - \theta))^2] \\ &= \mathbb{E}[(\hat{\theta} - \mu_{\hat{\theta}})^2] + 2\mathbb{E}[(\hat{\theta} - \mu_{\hat{\theta}})(\mu_{\hat{\theta}} - \theta)] + (\mu_{\hat{\theta}} - \theta)^2.\end{aligned}$$

Ο πρώτος όρος είναι η $\text{Var}(\hat{\theta})$. Ο μεσαίος όρος μηδενίζεται γιατί η ποσότητα $\mu_{\hat{\theta}} - \theta$ είναι σταθερή και $\mathbb{E}(\hat{\theta} - \mu_{\hat{\theta}}) = 0$. Ο τρίτος όρος είναι ακριβώς $\text{bias}(\hat{\theta})^2$. $\square$

Συνέπεια του θεωρήματος

Για αμερόληπτο εκτιμητή: $\text{MSE}(\hat{\theta}) = \text{Var}(\hat{\theta})$.

Το MSE επιτρέπει να συγκρίνουμε ακόμη και μεροληπτικούς εκτιμητές με αμερόληπτους — ένας μεροληπτικός εκτιμητής με μικρή διασπορά μπορεί να έχει μικρότερο MSE από έναν αμερόληπτο με μεγάλη διασπορά (κλασικό «bias-variance tradeoff» της Στατιστικής Μάθησης).

Παράδειγμα 1.6.1 — τιμωρία για bias. Έστω ότι έχουμε δύο εκτιμητές της παραμέτρου $\theta = 10$:

- Ο $\hat{\theta}_A$ είναι αμερόληπτος με $\text{Var}(\hat{\theta}_A) = 4$. Άρα $\text{MSE}(\hat{\theta}_A) = 4 + 0 = 4$.
- Ο $\hat{\theta}_B$ έχει $\mathbb{E}(\hat{\theta}_B) = 9$ ($\text{bias} = -1$) και $\text{Var}(\hat{\theta}_B) = 1$. Άρα $\text{MSE}(\hat{\theta}_B) = 1 + (-1)^2 = 2$.

Παρ' όλο που ο $\hat{\theta}_B$ είναι μεροληπτικός, έχει το *μισό* MSE από τον αμερόληπτο. Σε επαναλαμβανόμενα δείγματα, οι εκτιμήσεις του $\hat{\theta}_B$ θα είναι κοντύτερα στο 10 σε ευκλείδεια έννοια.

Σύνδεση με την Οικονομετρία

Το bias-variance tradeoff είναι το θεμελιώδες εργαλείο στην επιλογή μεταξύ OLS και κανονικοποιημένων εκτιμητών (Ridge, Lasso, Elastic Net). Ο OLS είναι αμερόληπτος αλλά με μεγάλη διασπορά όταν οι προβλέπουσες μεταβλητές είναι συγγραμμικές. Ο Ridge εκτιμητής εισάγει bias αλλά μειώνει δραστικά τη διασπορά — και συχνά έχει χαμηλότερο MSE.

1.6 Συνέπεια

Η συνέπεια είναι μια ασυμπτωτική (μεγάλου δείγματος) ιδιότητα: λέει ότι όσο μεγαλώνει το δείγμα, ο εκτιμητής «πέφτει πάνω» στην αληθινή παράμετρο.

Ορισμός 1.7 (Συνέπεια)

Ο εκτιμητής $\hat{\theta}_n$ (όπου ο δείκτης n τονίζει την εξάρτηση από το μέγεθος δείγματος) ονομάζεται **συνεπής** (consistent) για την παράμετρο θ αν

$$\hat{\theta}_n \xrightarrow{\mathbb{P}} \theta \quad \text{καθώς } n \rightarrow \infty,$$

δηλαδή για κάθε $\varepsilon > 0$:

$$\lim_{n \rightarrow \infty} \mathbb{P}(|\hat{\theta}_n - \theta| \geq \varepsilon) = 0.$$

Συμβολικά: $\text{plim}_{n \rightarrow \infty} \hat{\theta}_n = \theta$.

Θεώρημα 1.2 (Επαρκές κριτήριο συνέπειας)

Αν ο εκτιμητής $\hat{\theta}_n$ ικανοποιεί

$$\mathbb{E}(\hat{\theta}_n) \xrightarrow{n \rightarrow \infty} \theta \quad \text{και} \quad \text{Var}(\hat{\theta}_n) \xrightarrow{n \rightarrow \infty} 0,$$

τότε είναι συνεπής για το θ . Ισοδύναμα: αν $\text{MSE}(\hat{\theta}_n) \rightarrow 0$, τότε ο $\hat{\theta}_n$ είναι συνεπής.

Απόδειξη. Από την ανισότητα του Markov εφαρμοσμένη στην $(\hat{\theta}_n - \theta)^2$:

$$\mathbb{P}(|\hat{\theta}_n - \theta| \geq \varepsilon) = \mathbb{P}((\hat{\theta}_n - \theta)^2 \geq \varepsilon^2) \leq \frac{\mathbb{E}[(\hat{\theta}_n - \theta)^2]}{\varepsilon^2} = \frac{\text{MSE}(\hat{\theta}_n)}{\varepsilon^2}.$$

Αφού $\text{MSE}(\hat{\theta}_n) = \text{Var}(\hat{\theta}_n) + \text{bias}(\hat{\theta}_n)^2 \rightarrow 0$, το δεξί μέλος συγκλίνει στο 0. $\square$

Χρήσιμο πρακτικά

Στις περισσότερες ασκήσεις, για να δείξουμε συνέπεια δεν προσπαθούμε να αποδείξουμε σύγκλιση κατά πιθανότητα από τον ορισμό. Αντίθετα:

1. Δείχνουμε ότι ο εκτιμητής είναι (ασυμπτωτικά) αμερόληπτος: $\mathbb{E}(\hat{\theta}_n) \rightarrow \theta$.
2. Δείχνουμε ότι $\text{Var}(\hat{\theta}_n) \rightarrow 0$ όταν $n \rightarrow \infty$.

Το Θεώρημα 1.2 κλείνει τότε το επιχείρημα.

Παράδειγμα 1.7.1 — διαίσθηση μέσω LLN. Για τον δειγματικό μέσο $\bar{X}$:

$$\mathbb{E}(\bar{X}) = \mu, \quad \text{Var}(\bar{X}) = \frac{\sigma^2}{n}.$$

Η μέση τιμή είναι σταθερά ίση με μ (όχι απλώς στο όριο), και η διασπορά συγκλίνει στο 0 με ρυθμό $1/n$. Άρα ο $\bar{X}$ είναι συνεπής για το μ . Αυτό είναι ο **Ασθενής Νόμος των Μεγάλων Αριθμών** (WLLN): ο δειγματικός μέσος συγκλίνει κατά πιθανότητα στη μέση τιμή του πληθυσμού.

Παράδειγμα 1.7.2 — αντιπαράδειγμα. Δεν είναι όλοι οι εκτιμητές συνεπείς. Έστω

$$\hat{\mu}^{\text{bad}} = X_1$$

(δηλαδή χρησιμοποιούμε μόνο την πρώτη παρατήρηση και αγνοούμε τις υπόλοιπες). Τότε $\mathbb{E}(\hat{\mu}^{\text{bad}}) = \mu$ (αμερόληπτος!) αλλά $\text{Var}(\hat{\mu}^{\text{bad}}) = \sigma^2$ που δεν εξαρτάται από το n . Άρα δεν συγκλίνει στο 0, και ο εκτιμητής δεν είναι συνεπής — όσο μεγάλο δείγμα και να έχουμε, η εκτίμηση μένει εξίσου θορυβώδης.

Αμεροληψία και συνέπεια — δύο διαφορετικές έννοιες

- Ένας εκτιμητής μπορεί να είναι αμερόληπτος αλλά όχι συνεπής (παράδειγμα 1.7.2).
- Ένας εκτιμητής μπορεί να είναι συνεπής αλλά όχι αμερόληπτος (π.χ. ο MLE της Εκθετικής, βλ. Άσκηση 9).
- Ένας «καλός» εκτιμητής συνήθως ικανοποιεί και τα δύο (ή τουλάχιστον είναι ασυμπτωτικά αμερόληπτος και συνεπής).

1.7 Σχετική Αποτελεσματικότητα και UMVUE

Ορισμός 1.8 (Σχετική αποτελεσματικότητα)

Έστω $\hat{\theta}_1, \hat{\theta}_2$ δύο αμερόληπτοι εκτιμητές της ίδιας παραμέτρου θ . Η **σχετική αποτελεσματικότητα** (relative efficiency) του $\hat{\theta}_1$ ως προς τον $\hat{\theta}_2$ ορίζεται ως

$$RE(\hat{\theta}_1, \hat{\theta}_2) = \frac{\text{Var}(\hat{\theta}_2)}{\text{Var}(\hat{\theta}_1)}.$$

- Αν $RE(\hat{\theta}_1, \hat{\theta}_2) > 1$, ο $\hat{\theta}_1$ είναι *πιο αποτελεσματικός* (έχει μικρότερη διασπορά).
- Αν $RE(\hat{\theta}_1, \hat{\theta}_2) < 1$, ο $\hat{\theta}_2$ είναι *πιο αποτελεσματικός*.
- Αν $RE(\hat{\theta}_1, \hat{\theta}_2) = 1$, οι δύο εκτιμητές είναι εξίσου αποδοτικοί.

Διαισθητικά

Η RE μάς λέει «πόσες φορές μεγαλύτερο δείγμα θα χρειαστεί ο $\hat{\theta}_2$ για να φτάσει την ακρίβεια του $\hat{\theta}_1$ ». Αν $RE = 2$, χρειάζεται διπλάσιο μέγεθος δείγματος.

Από τη σχετική στην απόλυτη αποτελεσματικότητα: UMVUE

Η σχετική αποτελεσματικότητα συγκρίνει δύο εκτιμητές. Αλλά μπορούμε να ρωτήσουμε κάτι ισχυρότερο: *ποιος είναι ο εκτιμητής με τη μικρότερη δυνατή διασπορά μεταξύ όλων των αμερόληπτων*;

Ορισμός 1.9 (UMVUE)

Ένας εκτιμητής $\hat{\theta}^*$ της παραμέτρου θ ονομάζεται **Ομοιόμορφα Ελάχιστης Διασποράς Αμερόληπτος Εκτιμητής** (Uniformly Minimum Variance Unbiased Estimator — UMVUE) αν:

1. Είναι αμερόληπτος: $\mathbb{E}(\hat{\theta}^*) = \theta$, για κάθε θ .
2. Για κάθε άλλον αμερόληπτο εκτιμητή $\hat{\theta}$: $\text{Var}(\hat{\theta}^*) \leq \text{Var}(\hat{\theta})$, για κάθε θ .

Παραδείγματα UMVUE:

- Για κανονικό πληθυσμό $N(\mu, \sigma^2)$ με γνωστή σ^2 : ο $\bar{X}$ είναι UMVUE για το μ .
- Για πληθυσμό Bernoulli(p): ο $\hat{p} = \bar{X}$ είναι UMVUE για το p .
- Για πληθυσμό Poisson(λ): ο $\bar{X}$ είναι UMVUE για το λ .

Η ύπαρξη και κατασκευή του UMVUE συνδέεται στενά με την έννοια της **επάρκειας** (sufficiency) και την **Ανισότητα Cramer-Rao**, τις οποίες θα δούμε αναλυτικά στην Εβδομάδα 2. Σε αυτή την εβδομάδα μάς αρκεί να γνωρίζουμε ότι:

1. Δεν υπάρχει πάντοτε UMVUE — σε ορισμένες κατανομές το παίγνιο «αμερόληπτος με ελάχιστη διασπορά» δεν έχει νικητή.
2. Όταν υπάρχει, ο UMVUE συνήθως φτάνει το *Cramer-Rao κάτω φράγμα* για τη διασπορά κάθε αμερόληπτου εκτιμητή.
3. Στα παραδείγματα που θα δούμε, ο δειγματικός μέσος είναι σχεδόν πάντοτε UMVUE για τη μέση τιμή.

Σύνδεση με την Οικονομετρία

Το πιο διάσημο αποτέλεσμα «UMVUE» στην οικονομετρία είναι το **Θεώρημα Gauss-Markov**: κάτω από τις κλασικές υποθέσεις (γραμμικότητα, αυστηρή εξωγένεια, ομοσκεδαστικότητα), ο OLS εκτιμητής είναι *BLUE* (Best Linear Unbiased Estimator) — δηλαδή έχει την ελάχιστη διασπορά μεταξύ όλων των γραμμικών αμερόληπτων εκτιμητών. Είναι το ανάλογο του UMVUE περιορισμένο στην κατηγορία των γραμμικών εκτιμητών.

1.8 Μέθοδος των Ροπών

Η **Μέθοδος των Ροπών** (Method of Moments — MoM) είναι η παλαιότερη και διαισθητικά απλούστερη τεχνική για την κατασκευή εκτιμητών. Διατυπώθηκε από τον Karl Pearson το 1894 και ήταν η κυρίαρχη τεχνική μέχρι την επικράτηση της Μεθόδου Μέγιστης Πιθανοφάνειας από τον R.A. Fisher στις αρχές του 20ού αιώνα.

Ορισμός 1.10 (Πληθυσμιακές και δειγματικές ροπές)

Έστω X τυχαία μεταβλητή με κατανομή που εξαρτάται από παραμέτρους $\theta_1, \dots, \theta_k$. Η **r -οστή πληθυσμιακή ροπή** (περί την αρχή) ορίζεται ως

$$\mu'_r = \mathbb{E}(X^r), \quad r = 1, 2, 3, \dots,$$

και είναι, γενικά, συνάρτηση των $\theta_1, \dots, \theta_k$.

Για τυχαίο δείγμα $X_1, \dots, X_n$, η **r -οστή δειγματική ροπή** ορίζεται ως

$$m_r = \frac{1}{n} \sum_{i=1}^n X_i^r.$$

Παραδείγματα πληθυσμιακών ροπών:

Κατανομή	$\mu'_1 = \mathbb{E}(X)$	$\mu'_2 = \mathbb{E}(X^2)$
Bernoulli(p)	p	p
Binomial(n, p)	np	$np(1-p) + (np)^2$
Poisson(λ)	λ	$\lambda + \lambda^2$
Geometric(p) ($X \in \{1, 2, \dots\}$)	$1/p$	$(2-p)/p^2$
Exp(rate λ)	$1/\lambda$	$2/\lambda^2$
$U(0, \theta)$	$\theta/2$	$\theta^2/3$
$N(\mu, \sigma^2)$	μ	$\mu^2 + \sigma^2$
Gamma(α, β)	$\alpha\beta$	$\alpha(\alpha+1)\beta^2$

Αρχή της Μεθόδου των Ροπών

Για να εκτιμήσουμε τις k άγνωστες παραμέτρους $\theta_1, \dots, \theta_k$, εξισώνουμε τις k πρώτες πληθυσμιακές ροπές με τις αντίστοιχες δειγματικές:

$$\mu'_r(\theta_1, \dots, \theta_k) = m_r, \quad r = 1, 2, \dots, k.$$

Λύνοντας το σύστημα ως προς $(\theta_1, \dots, \theta_k)$ προκύπτουν οι εκτιμητές της μεθόδου των ροπών $\hat{\theta}_1^{\text{MoM}}, \dots, \hat{\theta}_k^{\text{MoM}}$.

Διαδικασία βήμα-προς-βήμα:

1. Καταγράφουμε τις k πρώτες πληθυσμιακές ροπές ως συναρτήσεις των παραμέτρων.

2. Υπολογίζουμε τις k πρώτες δειγματικές ροπές.
3. Σχηματίζουμε ένα σύστημα k εξισώσεων με k αγνώστους.
4. Λύνουμε το σύστημα.

Γιατί λειτουργεί (διαισθητική θεμελίωση)

Από τον Νόμο των Μεγάλων Αριθμών, $m_r \xrightarrow{\mathbb{P}} \mu'_r$ καθώς $n \rightarrow \infty$. Δηλαδή, σε μεγάλα δείγματα, η δειγματική ροπή είναι ουσιαστικά η πληθυσμιακή ροπή. Άρα αν εξισώσουμε τις (γνωστές πλέον) δειγματικές ροπές με τις (συναρτήσεις των παραμέτρων) πληθυσμιακές ροπές, έχουμε ένα σύστημα που στο όριο μάς δίνει την αληθινή τιμή των παραμέτρων.

Αυτή είναι η *αρχή της αντιστοιχίας* (analogy principle) και αποτελεί τη φιλοσοφική βάση πολλών εκτιμητικών τεχνικών — συμπεριλαμβανομένης της Γενικευμένης Μεθόδου των Ροπών (GMM) που είναι θεμελιώδης στη σύγχρονη οικονομετρία.

Πλεονεκτήματα και μειονεκτήματα

Πλεονεκτήματα:

- Απλή στη διατύπωση και τον υπολογισμό.
- Δεν απαιτεί γνώση της συνάρτησης πυκνότητας — μόνο γνώση των ροπών.
- Δίνει συνήθως συνεπείς εκτιμητές.

Μειονεκτήματα:

- Δεν είναι πάντα αποτελεσματική (συνήθως έχει μεγαλύτερη διασπορά από τη Μέθοδο Μέγιστης Πιθανοφάνειας — βλ. Εβδομάδα 2).
- Σε ορισμένες κατανομές μπορεί να δώσει αδύνατες τιμές (π.χ. αρνητικές για παράμετρο που πρέπει να είναι θετική).
- Δεν χρησιμοποιεί όλη την πληροφορία του δείγματος.

Σύνδεση με την Οικονομετρία

Η **Γενικευμένη Μέθοδος Ροπών (GMM)**, που θα δείτε στην Οικονομετρία II, βασίζεται στην ίδια αρχή αντιστοιχίας αλλά με συνθήκες ροπών που προέρχονται από οικονομικά μοντέλα (π.χ. Euler equations, instrumental variables). Ο Lars Peter Hansen τιμήθηκε με το Νόμπελ Οικονομίας το 2013 ακριβώς για την ανάπτυξη της GMM.

2 Λυμένες Ασκήσεις

2.1 Ομάδα Α: Ιδιότητες Εκτιμητών

Άσκηση 1: Αμεροληψία του δειγματικού μέσου

Έστω $X_1, X_2, \dots, X_n$ τυχαίο δείγμα μηνιαίων μισθών εργαζομένων σε έναν κλάδο, από κατανομή με μέση τιμή $\mathbb{E}(X_i) = \mu$ (άγνωστος μέσος μισθός) και

διασπορά $\text{Var}(X_i) = \sigma^2$. Ορίζουμε τον εκτιμητή

$$\hat{\mu} = \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i.$$

(α') Δείξτε ότι ο $\bar{X}$ είναι αμερόληπτος εκτιμητής του μ .

(β') Υπολογίστε τη διασπορά $\text{Var}(\bar{X})$ και το τυπικό σφάλμα $\text{SE}(\bar{X})$.

(γ') Σε ένα δείγμα $n = 30$ εργαζομένων με $\bar{x} = 1.450 \in$ και $\sigma = 240 \in$ (έστω γνωστή), υπολογίστε αριθμητικά την εκτίμηση και το τυπικό σφάλμα της.

(δ') Πόσο πρέπει να αυξηθεί το δείγμα ώστε το τυπικό σφάλμα να πέσει στο μισό;

Λύση

(α) Από τη γραμμικότητα της μέσης τιμής:

$$\mathbb{E}(\bar{X}) = \mathbb{E}\left[\frac{1}{n} \sum_{i=1}^n X_i\right] = \frac{1}{n} \sum_{i=1}^n \mathbb{E}(X_i) = \frac{1}{n} \cdot n \cdot \mu = \mu.$$

Άρα $\text{bias}(\bar{X}) = \mathbb{E}(\bar{X}) - \mu = 0$, οπότε ο $\bar{X}$ είναι αμερόληπτος.

(β) Επειδή τα X_i είναι ανεξάρτητα (τυχαίο δείγμα), από την ιδιότητα της διασποράς αθροίσματος ανεξάρτητων τ.μ.:

$$\text{Var}(\bar{X}) = \text{Var}\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n^2} \sum_{i=1}^n \text{Var}(X_i) = \frac{1}{n^2} \cdot n\sigma^2 = \frac{\sigma^2}{n}.$$

Άρα

$$\text{SE}(\bar{X}) = \sqrt{\text{Var}(\bar{X})} = \frac{\sigma}{\sqrt{n}}.$$

(γ) Με $n = 30$, $\bar{x} = 1.450$, $\sigma = 240$:

$$\hat{\mu}_{\text{obs}} = 1.450 \in, \quad \text{SE}(\bar{X}) = \frac{240}{\sqrt{30}} \approx \frac{240}{5,477} \approx 43,82 \in.$$

(δ) Θέλουμε νέο τυπικό σφάλμα ίσο με $43,82/2 \approx 21,91$:

$$\frac{240}{\sqrt{n'}} = 21,91 \iff \sqrt{n'} = \frac{240}{21,91} \approx 10,95 \iff n' \approx 120.$$

Άρα χρειάζεται τετραπλάσιο δείγμα — αυτή είναι η συνέπεια της παρουσίας του $\sqrt{n}$ στον παρονομαστή.

Παρατήρηση

Το γεγονός ότι $\text{SE}(\bar{X}) = \sigma/\sqrt{n}$ έχει βαθιά οικονομετρική σημασία: για να μειώσουμε στο μισό το τυπικό σφάλμα χρειαζόμαστε τετραπλάσιο δείγμα. Αυτός είναι ο «κανόνας $\sqrt{n}$ » — και ισχύει ασυμπτωτικά για όλους τους ομαλούς εκτιμητές που θα συναντήσετε. Σε σχέση με κόστος δειγματοληψίας, υπάρχουν φθίνουσες αποδόσεις: το πέμπτο εκατοντάδα παρατηρήσεις βελτιώνει την ακρί-

βεια πολύ λιγότερο από το πρώτο.

Άσκηση 2: S^2 vs $\hat{\sigma}^2$ — αμεροληψία και MSE

Έστω $X_1, \dots, X_n$ τυχαίο δείγμα ημερήσιων αποδόσεων μιας μετοχής από κανονικό πληθυσμό $N(\mu, \sigma^2)$. Θεωρούμε δύο εκτιμητές της διασποράς σ^2 :

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2 \quad \text{και} \quad \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2.$$

(α') Δείξτε ότι $\mathbb{E}(S^2) = \sigma^2$ (ο S^2 είναι αμερόληπτος).

(β') Υπολογίστε το $\mathbb{E}(\hat{\sigma}^2)$ και το σφάλμα μεροληψίας $\text{bias}(\hat{\sigma}^2)$.

(γ') Δεδομένου ότι $\text{Var}(S^2) = \frac{2\sigma^4}{n-1}$ για το κανονικό μοντέλο, υπολογίστε $\text{MSE}(S^2)$ και $\text{MSE}(\hat{\sigma}^2)$. Ποιος εκτιμητής έχει μικρότερο MSE;

(δ') Σχεδιάστε νοητά το γράφημα του λόγου $\text{MSE}(\hat{\sigma}^2)/\text{MSE}(S^2)$ ως προς n . Πού πηγαίνει για μεγάλα n ;

Λύση

(α) Χρησιμοποιούμε τη γνωστή ταυτότητα

$$\sum_{i=1}^n (X_i - \bar{X})^2 = \sum_{i=1}^n (X_i - \mu)^2 - n(\bar{X} - \mu)^2.$$

Παίρνοντας μέσες τιμές:

$$\mathbb{E} \left[\sum_{i=1}^n (X_i - \bar{X})^2 \right] = \sum_{i=1}^n \mathbb{E}[(X_i - \mu)^2] - n \mathbb{E}[(\bar{X} - \mu)^2] = n\sigma^2 - n \cdot \frac{\sigma^2}{n} = (n-1)\sigma^2.$$

Συνεπώς

$$\mathbb{E}(S^2) = \frac{1}{n-1} \cdot (n-1)\sigma^2 = \sigma^2. \quad \checkmark$$

(β) Από το προηγούμενο αποτέλεσμα,

$$\mathbb{E}(\hat{\sigma}^2) = \frac{1}{n} \cdot (n-1)\sigma^2 = \frac{n-1}{n}\sigma^2 < \sigma^2.$$

Άρα ο $\hat{\sigma}^2$ υποεκτιμά τη διασπορά, με σφάλμα μεροληψίας

$$\text{bias}(\hat{\sigma}^2) = \mathbb{E}(\hat{\sigma}^2) - \sigma^2 = -\frac{\sigma^2}{n}.$$

(γ) Η σχέση μεταξύ των δύο εκτιμητών είναι $\hat{\sigma}^2 = \frac{n-1}{n}S^2$, οπότε

$$\text{Var}(\hat{\sigma}^2) = \left(\frac{n-1}{n}\right)^2 \text{Var}(S^2) = \left(\frac{n-1}{n}\right)^2 \cdot \frac{2\sigma^4}{n-1} = \frac{2(n-1)\sigma^4}{n^2}.$$

Άρα:

$$\text{MSE}(S^2) = \text{Var}(S^2) + 0 = \frac{2\sigma^4}{n-1},$$

$$\text{MSE}(\hat{\sigma}^2) = \text{Var}(\hat{\sigma}^2) + \text{bias}(\hat{\sigma}^2)^2 = \frac{2(n-1)\sigma^4}{n^2} + \frac{\sigma^4}{n^2} = \frac{(2n-1)\sigma^4}{n^2}.$$

Σύγκριση: Υπολογίζουμε τον λόγο

$$\frac{\text{MSE}(\hat{\sigma}^2)}{\text{MSE}(S^2)} = \frac{(2n-1)\sigma^4/n^2}{2\sigma^4/(n-1)} = \frac{(2n-1)(n-1)}{2n^2}.$$

Πίνακας τιμών:

n	$\text{MSE}(\hat{\sigma}^2)/\sigma^4$	$\text{MSE}(S^2)/\sigma^4$	λόγος
2	0,750	2,000	0,375
5	0,360	0,500	0,720
10	0,190	0,222	0,855
30	0,0656	0,069	0,9483
100	0,0199	0,0202	0,985

(6) Για μεγάλα n :

$$\frac{(2n-1)(n-1)}{2n^2} = \frac{2n^2 - 3n + 1}{2n^2} = 1 - \frac{3}{2n} + \frac{1}{2n^2} \xrightarrow{n \rightarrow \infty} 1.$$

Δηλαδή για μεγάλα δείγματα οι δύο εκτιμητές είναι πρακτικά ισοδύναμοι. Για μικρά δείγματα, ο $\hat{\sigma}^2$ έχει σημαντικά μικρότερο MSE.

Έλεγχος

Παρόλο που ο $\hat{\sigma}^2$ είναι μεροληπτικός, έχει *μικρότερο* MSE από τον αμερόληπτο S^2 — για κάθε $n \geq 2$. Αυτή είναι μια κλασική εκδοχή του «bias-variance tradeoff»: ο μεροληπτικός εκτιμητής «πληρώνει» ένα μικρό bias για να κερδίσει σε διασπορά.

Παρ' όλα αυτά, στην οικονομετρική πρακτική χρησιμοποιείται κατά κανόνα ο S^2 (διορθωμένος), γιατί η αμεροληψία είναι σημαντική για τη σωστή κατασκευή στατιστικών ελέγχου (t-statistic κ.λπ.). Σε οικονομετρικά λογισμικά (Stata, R, Python statsmodels), η εκτίμηση της διασποράς γίνεται με τον S^2 .

Άσκηση 3: Αμεροληψία της δειγματικής αναλογίας

Μια εμπορική τράπεζα μελετά την αναλογία p των στεγαστικών δανείων που πέφτουν σε αθέτηση πληρωμών μέσα στην πρώτη τριετία. Για κάθε δάνειο i από τυχαίο δείγμα μεγέθους n , ορίζουμε

$$X_i = \begin{cases} 1, & \text{αν το δάνειο } i \text{ πέφτει σε αθέτηση,} \\ 0, & \text{διαφορετικά,} \end{cases}$$

οπότε $X_i \sim \text{Bernoulli}(p)$, ανεξάρτητα.

(α') Δείξτε ότι ο εκτιμητής $\hat{p} = \bar{X} = \frac{1}{n} \sum X_i$ είναι αμερόληπτος για το p .

(β') Υπολογίστε τη $\text{Var}(\hat{p})$ και το $\text{SE}(\hat{p})$.

(γ') Σε δείγμα $n = 500$ δανείων παρατηρήθηκαν 28 αθετήσεις. Υπολογίστε την εκτίμηση $\hat{p}_{\text{obs}}$ και το *εκτιμώμενο* τυπικό σφάλμα $\widehat{\text{SE}}(\hat{p}) = \sqrt{\hat{p}(1-\hat{p})/n}$.

(δ') Για ποιο p είναι μέγιστο το τυπικό σφάλμα της $\hat{p}$;

Λύση

(α) Επειδή $X_i \sim \text{Bernoulli}(p)$, έχουμε $\mathbb{E}(X_i) = p$ και $\text{Var}(X_i) = p(1 - p)$. Από την Άσκηση 1:

$$\mathbb{E}(\hat{p}) = \mathbb{E}(\bar{X}) = p. \quad \checkmark$$

(β)

$$\text{Var}(\hat{p}) = \frac{\text{Var}(X_i)}{n} = \frac{p(1-p)}{n}, \quad \text{SE}(\hat{p}) = \sqrt{\frac{p(1-p)}{n}}.$$

(γ) Με $n = 500$ και 28 αθετήσεις:

$$\hat{p}_{\text{obs}} = \frac{28}{500} = 0,056 = 5,6\%.$$

$$\widehat{\text{SE}}(\hat{p}) = \sqrt{\frac{0,056 \cdot 0,944}{500}} = \sqrt{\frac{0,05286}{500}} \approx \sqrt{0,0001057} \approx 0,01028.$$

Άρα η εκτίμηση είναι 5,6% με τυπικό σφάλμα περίπου 1,03 ποσοστιαίες μονάδες.

(δ) Παραγωγίζουμε ως προς p :

$$\frac{d}{dp}[p(1-p)] = 1 - 2p = 0 \iff p = \frac{1}{2}.$$

Άρα η διασπορά (και το τυπικό σφάλμα) είναι μέγιστα όταν $p = 0,5$. Σε εκείνο το σημείο $p(1-p) = 0,25$.

Παρατήρηση

Παρατηρήστε ότι το τυπικό σφάλμα μεγιστοποιείται για $p = 0,5$ (όπου $p(1-p) = 0,25$) και μικραίνει καθώς το p απομακρύνεται από το 0,5. Γι' αυτό σπάνια συναντάμε αναλογίες αθέτησης κοντά στο 50% με μικρό τυπικό σφάλμα.

Πρακτική συνέπεια: σε σχεδιασμό δημοσκοπήσης ή στατιστικής μελέτης όπου εκτιμάμε αναλογίες, η συντηρητική προσέγγιση για τον υπολογισμό μεγέθους δείγματος είναι να υποθέσουμε $p = 0,5$ — αυτό δίνει το μέγιστο αναγκαίο μέγεθος δείγματος.

Άσκηση 4: Τρεις εκτιμητές της μέσης τιμής — σύγκριση μέσω MSE

Από κανονικό πληθυσμό $N(\mu, \sigma^2)$ παίρνουμε τυχαίο δείγμα X_1, X_2, X_3 μεγέθους $n = 3$. Θεωρούμε τους εξής τρεις εκτιμητές για το μ :

$$\hat{\mu}_A = \frac{X_1 + X_2 + X_3}{3}, \quad \hat{\mu}_B = \frac{X_1 + X_3}{2}, \quad \hat{\mu}_C = \frac{X_1 + 2X_2 + X_3}{4}.$$

(Ο $\hat{\mu}_C$ είναι ένας σταθμισμένος μέσος που δίνει βάρος 2 στη μεσαία παρατήρηση.)

(α') Δείξτε ότι και οι τρεις εκτιμητές είναι αμερόληπτοι.

(β') Υπολογίστε τη διασπορά καθενός.

(γ') Ταξινομήστε τους από τον αποδοτικότερο στον λιγότερο αποδοτικό. Ποιος έχει το μικρότερο MSE;

(δ') Υπολογίστε τη σχετική αποτελεσματικότητα $\text{RE}(\hat{\mu}_A, \hat{\mu}_B)$ και $\text{RE}(\hat{\mu}_A, \hat{\mu}_C)$.

Λύση**(α)** Από γραμμικότητα της μέσης τιμής:

$$\mathbb{E}(\hat{\mu}_A) = \frac{1}{3}(\mu + \mu + \mu) = \mu,$$

$$\mathbb{E}(\hat{\mu}_B) = \frac{1}{2}(\mu + \mu) = \mu,$$

$$\mathbb{E}(\hat{\mu}_C) = \frac{1}{4}(\mu + 2\mu + \mu) = \frac{4\mu}{4} = \mu.$$

Όλοι αμερόληπτοι. ✓

(β) Επειδή τα X_i είναι ανεξάρτητα με $\text{Var}(X_i) = \sigma^2$:

$$\text{Var}(\hat{\mu}_A) = \frac{1}{9}(3\sigma^2) = \frac{\sigma^2}{3} \approx 0,333\sigma^2,$$

$$\text{Var}(\hat{\mu}_B) = \frac{1}{4}(2\sigma^2) = \frac{\sigma^2}{2} = 0,5\sigma^2,$$

$$\text{Var}(\hat{\mu}_C) = \frac{1}{16}(\sigma^2 + 4\sigma^2 + \sigma^2) = \frac{6\sigma^2}{16} = \frac{3\sigma^2}{8} = 0,375\sigma^2.$$

(γ) Αφού όλοι είναι αμερόληπτοι, $\text{MSE} = \text{Var}$. Ταξινόμηση από μικρότερη σε μεγαλύτερη διασπορά:

$$\hat{\mu}_A (0,333\sigma^2) < \hat{\mu}_C (0,375\sigma^2) < \hat{\mu}_B (0,500\sigma^2).$$

Νικητής: ο $\hat{\mu}_A$ (δειγματικός μέσος).**(δ)**

$$\text{RE}(\hat{\mu}_A, \hat{\mu}_B) = \frac{\text{Var}(\hat{\mu}_B)}{\text{Var}(\hat{\mu}_A)} = \frac{\sigma^2/2}{\sigma^2/3} = \frac{3}{2} = 1,5.$$

Δηλαδή ο $\hat{\mu}_A$ είναι 50% πιο αποδοτικός από τον $\hat{\mu}_B$.

$$\text{RE}(\hat{\mu}_A, \hat{\mu}_C) = \frac{\text{Var}(\hat{\mu}_C)}{\text{Var}(\hat{\mu}_A)} = \frac{3\sigma^2/8}{\sigma^2/3} = \frac{9}{8} = 1,125.$$

Δηλαδή ο $\hat{\mu}_A$ είναι 12,5% πιο αποδοτικός από τον $\hat{\mu}_C$.**Παρατήρηση**

Δύο σημαντικές παρατηρήσεις:

1. Ο $\hat{\mu}_B$ είναι χειρότερος γιατί αγνοεί την παρατήρηση X_2 .
2. Ο $\hat{\mu}_C$ είναι χειρότερος από τον $\hat{\mu}_A$ γιατί δίνει ασύμμετρα βάρη. **Αποδεικνύεται γενικότερα:** μεταξύ όλων των γραμμικών αμερόληπτων εκτιμητών της μορφής $\sum w_i X_i$ με $\sum w_i = 1$, αυτός με τη μικρότερη διασπορά είναι ο $w_i = 1/n$ — δηλαδή ο δειγματικός μέσος. Αυτό είναι ένα κλασικό αποτέλεσμα BLUE (Best Linear Unbiased Estimator) — προπομπός του θεωρήματος Gauss–Markov στην Οικονομετρία.

Άσκηση 5: Συνέπεια του δειγματικού μέσου μέσω Chebyshev

Έστω ότι μελετάμε την τιμή πώλησης διαμερισμάτων στο κέντρο της Αθήνας. Παίρνουμε τυχαίο δείγμα $X_1, \dots, X_n$ τιμών (σε χιλ. €) από πληθυσμό με $\mathbb{E}(X_i) = \mu$ και $\text{Var}(X_i) = \sigma^2 < \infty$ (παράμετροι άγνωστοι αλλά πεπερασμένοι).

(α') Χρησιμοποιώντας την ανισότητα Chebyshev, δείξτε ότι ο $\bar{X}$ είναι συνεπής για το μ .

(β') Υποθέστε $\sigma = 60$ χιλ. €. Πόσο μεγάλο πρέπει να είναι το δείγμα ώστε $\mathbb{P}(|\bar{X} - \mu| \geq 10) \leq 0,05$;

(γ') Πόσο μεγάλο πρέπει να είναι το δείγμα ώστε $\mathbb{P}(|\bar{X} - \mu| \geq 5) \leq 0,01$;

Λύση

(α) Η ανισότητα Chebyshev για τη μεταβλητή $\bar{X}$, που έχει μέση τιμή μ και διασπορά σ^2/n , λέει για κάθε $\varepsilon > 0$:

$$\mathbb{P}(|\bar{X} - \mu| \geq \varepsilon) \leq \frac{\text{Var}(\bar{X})}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2}.$$

Καθώς $n \rightarrow \infty$, το δεξί μέλος συγκλίνει στο 0 (αφού σ^2 και ε^2 είναι σταθερά). Άρα

$$\lim_{n \rightarrow \infty} \mathbb{P}(|\bar{X} - \mu| \geq \varepsilon) = 0,$$

δηλαδή $\bar{X} \xrightarrow{\mathbb{P}} \mu$, που είναι ο ορισμός της συνέπειας. ✓

(β) Θέλουμε $\sigma^2/(n\varepsilon^2) \leq 0,05$ με $\varepsilon = 10$, $\sigma^2 = 3.600$:

$$\frac{3.600}{n \cdot 100} \leq 0,05 \iff \frac{36}{n} \leq 0,05 \iff n \geq \frac{36}{0,05} = 720.$$

Άρα χρειάζονται **τουλάχιστον 720 διαμερίσματα** στο δείγμα.

(γ) Με $\varepsilon = 5$ και απαίτηση πιθανότητας 0,01:

$$\frac{3.600}{n \cdot 25} \leq 0,01 \iff \frac{144}{n} \leq 0,01 \iff n \geq 14.400.$$

Πολύ μεγαλύτερο δείγμα — γιατί διπλασιάσαμε την ακρίβεια (μειώσαμε το ε στο μισό, δηλαδή τετραπλασιάσαμε το $1/\varepsilon^2$) και πενταπλασιάσαμε την απαίτηση πιθανότητας.

Παρατήρηση

Το Chebyshev είναι «χαλαρό» όριο — στην πράξη με κανονικό πληθυσμό αρκούν πολύ λιγότερες παρατηρήσεις. Όμως η δύναμη του Chebyshev είναι ότι ισχύει **χωρίς υποθέσεις** για τη μορφή της κατανομής — μόνο πεπερασμένη διασπορά. Στην Οικονομετρία θα δείτε ότι αυτή είναι η τυπική στρατηγική για την απόδειξη συνέπειας του OLS εκτιμητή.

Άσκηση 6: Σχετική αποτελεσματικότητα δύο εκτιμητών

Έστω X_1, X_2, X_3, X_4 τυχαίο δείγμα από κατανομή με $\mathbb{E}(X_i) = \mu$ και $\text{Var}(X_i) = \sigma^2$. Ένας αναλυτής προτείνει δύο εκτιμητές του μ :

$$\hat{\mu}_1 = \frac{X_1 + X_2 + X_3 + X_4}{4} = \bar{X}, \quad \hat{\mu}_2 = \frac{X_1 + 2X_2 + 2X_3 + X_4}{6}.$$

(α') Επαληθεύστε ότι και οι δύο εκτιμητές είναι αμερόληπτοι.

(β') Υπολογίστε τη διασπορά καθενός.

(γ') Υπολογίστε τη σχετική αποτελεσματικότητα $\text{RE}(\hat{\mu}_1, \hat{\mu}_2)$ και ερμηνεύστε.

(δ') Πόσο μεγαλύτερο δείγμα θα χρειαζόταν ο $\hat{\mu}_2$ για να έχει την ίδια διασπορά με τον $\hat{\mu}_1$;

Λύση**(α)**

$$\mathbb{E}(\hat{\mu}_1) = \frac{4\mu}{4} = \mu, \quad \mathbb{E}(\hat{\mu}_2) = \frac{\mu+2\mu+2\mu+\mu}{6} = \frac{6\mu}{6} = \mu. \quad \checkmark$$

(β)

$$\text{Var}(\hat{\mu}_1) = \frac{1}{16}(4\sigma^2) = \frac{\sigma^2}{4}.$$

$$\text{Var}(\hat{\mu}_2) = \frac{1}{36}(\sigma^2 + 4\sigma^2 + 4\sigma^2 + \sigma^2) = \frac{10\sigma^2}{36} = \frac{5\sigma^2}{18} \approx 0,278\sigma^2.$$

Πράγματι $\sigma^2/4 = 0,25\sigma^2 < 0,278\sigma^2$, οπότε ο $\hat{\mu}_1$ έχει μικρότερη διασπορά.

(γ)

$$\text{RE}(\hat{\mu}_1, \hat{\mu}_2) = \frac{\text{Var}(\hat{\mu}_2)}{\text{Var}(\hat{\mu}_1)} = \frac{5\sigma^2/18}{\sigma^2/4} = \frac{20}{18} = \frac{10}{9} \approx 1,111.$$

Ο $\hat{\mu}_1$ είναι περίπου 11,1% πιο αποδοτικός από τον $\hat{\mu}_2$.

(δ) Αν θέλαμε τον $\hat{\mu}_2$ με διασπορά $\sigma^2/4$ από δείγμα μεγέθους m , θα έπρεπε

$$\frac{5\sigma^2}{18} \cdot \frac{4}{m} = \frac{\sigma^2}{4} \iff \frac{20}{18m} = \frac{1}{4} \iff m = \frac{80}{18} \approx 4,44.$$

Δηλαδή για κάθε 4 παρατηρήσεις που χρησιμοποιεί ο $\hat{\mu}_1$, ο $\hat{\mu}_2$ θα χρειαζόταν περίπου 4,44 — ένα κέρδος αποδοτικότητας περίπου 11% υπέρ του $\hat{\mu}_1$.

Άσκηση 7: Όταν ένας μεροληπτικός εκτιμητής νικάει — εκτίμηση διασποράς μισθών

Από κανονικό πληθυσμό $N(\mu, \sigma^2)$ που περιγράφει τους μηνιαίους μισθούς εργαζομένων, παίρνουμε τυχαίο δείγμα $X_1, \dots, X_n$. Θεωρούμε τον μεροληπτικό εκτιμητή της διασποράς

$$\hat{\sigma}_c^2 = c \cdot \sum_{i=1}^n (X_i - \bar{X})^2,$$

όπου $c > 0$ είναι σταθερά. (Π.χ. $c = 1/(n-1)$ δίνει τον αμερόληπτο S^2 , $c = 1/n$ δίνει τον MLE $\hat{\sigma}^2$.)

(α') Δείξτε ότι $\mathbb{E}(\hat{\sigma}_c^2) = c(n-1)\sigma^2$ και άρα $\text{bias}(\hat{\sigma}_c^2) = (c(n-1) - 1)\sigma^2$.

(β') Δεδομένου ότι $\text{Var}(\sum (X_i - \bar{X})^2) = 2(n-1)\sigma^4$, υπολογίστε το $\text{MSE}(\hat{\sigma}_c^2)$ ως συνάρτηση του c .

(γ') Βρείτε τη βέλτιστη τιμή c^* που ελαχιστοποιεί το MSE.

(δ') Συγκρίνετε με $c = 1/(n-1)$ (αμερόληπτος) και $c = 1/n$ (MLE).

Λύση

(α) Από την Άσκηση 2 ξέρουμε ότι $\mathbb{E}[\sum (X_i - \bar{X})^2] = (n-1)\sigma^2$. Άρα

$$\mathbb{E}(\hat{\sigma}_c^2) = c(n-1)\sigma^2.$$

Το σφάλμα μεροληψίας:

$$\text{bias}(\hat{\sigma}_c^2) = c(n-1)\sigma^2 - \sigma^2 = (c(n-1) - 1)\sigma^2.$$

(β)

$$\text{Var}(\hat{\sigma}_c^2) = c^2 \cdot 2(n-1)\sigma^4.$$

$$\text{MSE}(\hat{\sigma}_c^2) = \text{Var} + \text{bias}^2 = 2c^2(n-1)\sigma^4 + (c(n-1) - 1)^2\sigma^4.$$

Αναπτύσσοντας:

$$\begin{aligned} \text{MSE}(\hat{\sigma}_c^2)/\sigma^4 &= 2c^2(n-1) + c^2(n-1)^2 - 2c(n-1) + 1. \\ &= c^2[(n-1)(2 + (n-1))] - 2c(n-1) + 1 = c^2(n-1)(n+1) - 2c(n-1) + 1. \end{aligned}$$

(γ) Παραγωγίζουμε ως προς c και θέτουμε ίσο με 0:

$$\begin{aligned} \frac{d}{dc} [c^2(n-1)(n+1) - 2c(n-1) + 1] &= 2c(n-1)(n+1) - 2(n-1) = 0 \\ \Leftrightarrow c^* &= \frac{1}{n+1}. \end{aligned}$$

Άρα η βέλτιστη επιλογή ως προς MSE είναι $c^* = 1/(n+1)$.

(δ) Συνοπτικός πίνακας για $n = 10$ (όπου $c^* = 1/11 \approx 0,0909$):

Εκτιμητής	c	bias / σ^2	MSE / σ^4
S^2 (αμερόληπτος)	$1/9 \approx 0,1111$	0	$2/9 \approx 0,222$
$\hat{\sigma}^2$ (MLE)	$1/10 = 0,100$	$-1/10 = -0,10$	$19/100 = 0,190$
$\hat{\sigma}_{c^*}^2$ (βέλτ. MSE)	$1/11 \approx 0,0909$	$-2/11 \approx -0,182$	$9/55 \approx 0,164$

Το παράδοξο της αμεροληψίας

Με βάση το MSE κριτήριο, ο εκτιμητής $c^* = 1/(n+1)$ νικάει και τον αμερόληπτο και τον MLE. Όμως δεν χρησιμοποιείται ποτέ στην πράξη. Γιατί;

Λόγοι:

1. Η μεροληψία του είναι αρνητική και αυξάνεται (σε απόλυτη τιμή) σε σχέση με τον MLE.
2. Όλοι οι κλασικοί έλεγχοι υποθέσεων (Εβδομάδα 3) και τα διαστήματα εμπιστοσύνης βασίζονται στον αμερόληπτο S^2 — αν χρησιμοποιήσουμε άλλο εκτιμητή, καταρρέει ολόκληρη η θεωρία ελέγχων.
3. Η βελτίωση στο MSE είναι μικρή, ενώ η απώλεια διαφάνειας μεγάλη.

Μάθημα: η επιλογή εκτιμητή δεν γίνεται μόνο με βάση το MSE. Άλλα κριτήρια (αμεροληψία, αναλυτική απλότητα, συμβατότητα με μετέπειτα ελέγχους) έχουν σημασία.

2.2 Ομάδα Β: Μέθοδος των Ροπών — Μονοπαραμετρικές Κατανομές

Άσκηση 8: MoM για Bernoulli

Ένα marketing department θέλει να εκτιμήσει την πιθανότητα θ ότι ένας τυχαίος πελάτης θα αγοράσει ένα νέο προϊόν. Σε τυχαίο δείγμα n πελατών, καταγράφεται $X_i \sim \text{Bernoulli}(\theta)$, ανεξάρτητα.

(α') Υπολογίστε τον εκτιμητή της Μεθόδου των Ροπών για το θ .

(β') Ελέγξτε αν είναι αμερόληπτος και συνεπής.

(γ') Σε δείγμα 200 πελατών, 46 αγόρασαν. Δώστε την εκτίμηση και το εκτιμώμενο τυπικό σφάλμα.

Λύση

(α) Η μοναδική παράμετρος είναι το θ , οπότε χρειαζόμαστε μόνο την πρώτη ροπή.

Πληθυσμιακή ροπή: $\mu'_1 = \mathbb{E}(X) = \theta$.

Δειγματική ροπή: $m_1 = \frac{1}{n} \sum_{i=1}^n X_i = \bar{X}$.

Εξίσωση:

$$\mu'_1 = m_1 \iff \theta = \bar{X} \implies \boxed{\hat{\theta}^{\text{MoM}} = \bar{X}}.$$

(β) Από την Άσκηση 3:

- Αμεροληψία: $\mathbb{E}(\hat{\theta}^{\text{MoM}}) = \theta$. ✓
- Διασπορά: $\text{Var}(\hat{\theta}^{\text{MoM}}) = \theta(1 - \theta)/n \rightarrow 0$. ✓ (Συνεπής)

(γ) $\hat{\theta}_{\text{obs}}^{\text{MoM}} = 46/200 = 0,23 = 23\%$.

$$\widehat{\text{SE}}(\hat{\theta}) = \sqrt{\frac{0,23 \cdot 0,77}{200}} = \sqrt{\frac{0,1771}{200}} \approx 0,0298 = 2,98\%.$$

Άσκηση 9: MoM για Poisson

Σε μια ασφαλιστική εταιρεία, ο αριθμός αξιώσεων αυτοκινήτου ανά μήνα μοντελοποιείται ως $X_i \sim \text{Poisson}(\lambda)$, ανεξάρτητα. Δίνεται τυχαίο δείγμα 12 μηνών:

3, 5, 4, 6, 2, 7, 4, 5, 3, 6, 5, 4.

(α') Βρείτε τον MoM εκτιμητή του λ .

(β') Υπολογίστε αριθμητικά την εκτίμηση.

(γ') Δείξτε ότι ο εκτιμητής είναι αμερόληπτος και συνεπής.

(δ') Πόσος είναι ο εκτιμώμενος μέσος ετήσιος αριθμός αξιώσεων;

Λύση

(α) Για την Poisson: $\mathbb{E}(X) = \lambda$. Άρα

$$\mu'_1 = m_1 \iff \lambda = \bar{X} \implies \boxed{\hat{\lambda}^{\text{MoM}} = \bar{X}}.$$

(β)

$$\bar{X} = \frac{1}{12}(3 + 5 + 4 + 6 + 2 + 7 + 4 + 5 + 3 + 6 + 5 + 4) = \frac{54}{12} = 4,5.$$

Άρα $\hat{\lambda}_{\text{obs}}^{\text{MoM}} = 4,5$ αξιώσεις/μήνα.

(γ) Για Poisson, $\text{Var}(X) = \lambda$, οπότε

$$\mathbb{E}(\hat{\lambda}^{\text{MoM}}) = \lambda \quad \checkmark, \quad \text{Var}(\hat{\lambda}^{\text{MoM}}) = \frac{\lambda}{n} \rightarrow 0. \quad \checkmark$$

(δ) Επειδή οι μηνιαίες αξιώσεις είναι iid Poisson(λ), ο ετήσιος αριθμός $Y = \sum_{m=1}^{12} X_m$ είναι Poisson(12λ) με $\mathbb{E}(Y) = 12\lambda$. Άρα $\widehat{\mathbb{E}}(Y) = 12 \times 4,5 = 54$ αξιώσεις/έτος.

Παρατήρηση

Στις «καλές» κατανομές που η πρώτη ροπή ταυτίζεται με την παράμετρο (Bernoulli, Poisson, Εκθετική με rate $\lambda = 1/\mu$, κ.λπ.), ο MoM εκτιμητής της παραμέτρου είναι κατά κανόνα ο $\bar{X}$ ή κάποια απλή συνάρτησή του.

Άσκηση 10: MoM για Εκθετική κατανομή

Ο χρόνος (σε λεπτά) που περιμένει ένας πελάτης σε ATM μέχρι να εξυπηρετηθεί μοντελοποιείται ως $X_i \sim \text{Exp}(\lambda)$, ανεξάρτητα, όπου $\lambda > 0$ είναι η παράμετρος ρυθμού (rate). Υπενθυμίζουμε ότι για την εκθετική:

$$f_X(x) = \lambda e^{-\lambda x}, \quad x > 0; \quad \mathbb{E}(X) = \frac{1}{\lambda}; \quad \text{Var}(X) = \frac{1}{\lambda^2}.$$

(α') Βρείτε τον MoM εκτιμητή του λ .

(β') Είναι αμερόληπτος; (Προσοχή — το αντίστροφο της μέσης τιμής δεν έχει μέση τιμή ίση με την αντίστροφη της μέσης τιμής.)

(γ') Είναι συνεπής; Δικαιολογήστε.

(δ') Σε δείγμα 50 πελατών, ο μέσος χρόνος αναμονής βρέθηκε 3 λεπτά. Δώστε την εκτίμηση.

Λύση

(α) $\mu'_1 = \mathbb{E}(X) = 1/\lambda$. Άρα

$$\frac{1}{\lambda} = \bar{X} \iff \hat{\lambda}^{\text{MoM}} = \frac{1}{\bar{X}}.$$

(β) Είναι ο $1/\bar{X}$ αμερόληπτος για το λ ; **Όχι, εν γένει.** Από την ανισότητα Jensen, για κυρτή συνάρτηση $g(x) = 1/x$ ($x > 0$):

$$\mathbb{E}\left(\frac{1}{\bar{X}}\right) \geq \frac{1}{\mathbb{E}(\bar{X})} = \lambda,$$

με γνήσια ανισότητα όταν η $\bar{X}$ δεν είναι σταθερή. Άρα ο $1/\bar{X}$ υπερεκτιμά το λ κατά μέσο όρο. (Πιο συγκεκριμένα, για την εκθετική μπορεί να αποδειχθεί ότι $\mathbb{E}(1/\bar{X}) = \frac{n}{n-1}\lambda$ για $n \geq 2$, οπότε το σφάλμα μεροληψίας είναι $\lambda/(n-1)$.)

(γ) Είναι συνεπής όμως. Από τον Νόμο των Μεγάλων Αριθμών $\bar{X} \xrightarrow{\mathbb{P}} 1/\lambda$, και από το *Continuous Mapping Theorem* (αν $Y_n \xrightarrow{\mathbb{P}} c$ και g συνεχής στο c , τότε $g(Y_n) \xrightarrow{\mathbb{P}} g(c)$):

$$\frac{1}{\bar{X}} \xrightarrow{\mathbb{P}} \frac{1}{1/\lambda} = \lambda.$$

(δ) $\bar{x} = 3$ λεπτά. Άρα

$$\hat{\lambda}_{\text{obs}}^{\text{MoM}} = \frac{1}{3} \approx 0,333 \text{ πελάτες/λεπτό}.$$

Δηλαδή ο ρυθμός εξυπηρέτησης είναι περίπου 1/3 πελάτη το λεπτό (ή 20 πελάτες την ώρα).

Παρατήρηση

Πολύ σημαντικό μάθημα: ο MoM εκτιμητής δεν είναι πάντα αμερόληπτος, ακόμη και όταν προέρχεται από αμερόληπτη πρώτη ροπή. Όταν η παράμετρος είναι μη γραμμική συνάρτηση της ροπής, εισάγεται bias.

Επίσης, η συνέπεια δεν χάνεται — διατηρείται μέσω του Continuous Mapping Theorem. Στην Οικονομετρία θα χρησιμοποιείτε αυτή την τεχνική συνεχώς (π.χ. ο εκτιμητής του β στο OLS είναι γραμμικός, αλλά τα εκτιμητικά μέτρα ελαστικότητας είναι μη γραμμικά).

Άσκηση 11: MoM για Γεωμετρική κατανομή

Σε μια αλυσίδα καταστημάτων μελετάται ο αριθμός X εμφανίσεων ενός υποψήφιου πελάτη στο κατάστημα μέχρι την πρώτη αγορά. Υποθέτουμε $X \sim \text{Geom}(\theta)$ στη γραφή $X \in \{1, 2, 3, \dots\}$, με

$$\mathbb{P}(X = k) = (1 - \theta)^{k-1}\theta, \quad k = 1, 2, 3, \dots$$

και $\mathbb{E}(X) = 1/\theta$, $\text{Var}(X) = (1 - \theta)/\theta^2$.

Σε δείγμα 8 πελατών παρατηρήθηκαν: 2, 1, 3, 4, 1, 2, 5, 2.

(α') Βρείτε τον MoM εκτιμητή του θ .

(β') Υπολογίστε αριθμητικά την εκτίμηση και ερμηνεύστε την.

(γ') Δείξτε ότι ο εκτιμητής είναι συνεπής.

Λύση

(α) $\mu'_1 = \mathbb{E}(X) = 1/\theta$. Άρα

$$\frac{1}{\theta} = \bar{X} \iff \boxed{\hat{\theta}^{\text{MoM}} = \frac{1}{\bar{X}}}.$$

(β)

$$\bar{x} = \frac{2+1+3+4+1+2+5+2}{8} = \frac{20}{8} = 2,5.$$

$$\hat{\theta}_{\text{obs}}^{\text{MoM}} = \frac{1}{2,5} = 0,4 = 40\%.$$

Δηλαδή, σύμφωνα με την εκτίμηση, υπάρχει 40% πιθανότητα ένας πελάτης να αγοράσει σε κάθε επίσκεψη.

(γ) Από τον LLN $\bar{X} \xrightarrow{\mathbb{P}} 1/\theta$, και από το Continuous Mapping Theorem

$$\frac{1}{\bar{X}} \xrightarrow{\mathbb{P}} \theta. \quad \checkmark$$

Όπως και στην εκθετική, ο εκτιμητής δεν είναι αμερόληπτος (Jensen — υπερεκτιμά το θ).

Άσκηση 12: MoM για Ομοιόμορφη κατανομή $U(0$

Ο χρόνος (σε ώρες) μέχρι την ολοκλήρωση μιας συναλλαγής σε ένα ηλεκτρονικό σύστημα μοντελοποιείται ως $X_i \stackrel{\text{iid}}{\sim} U(0, \theta)$ με $\theta > 0$ άγνωστο. Δίνεται τυχαίο

δείγμα $n = 5$:

1,8, 0,7, 3,1, 2,5, 1,4.

(α') Βρείτε τον MoM εκτιμητή του θ .

(β') Υπολογίστε αριθμητικά την εκτίμηση.

(γ') Παρατηρήστε ότι ο μέγιστος του δείγματος είναι 3,1. Είναι λογική η εκτίμηση; Εξετάστε τη γενική περίπτωση: μπορεί ο MoM εκτιμητής να είναι μικρότερος από το $\max(X_i)$; Τι θα σήμαινε αυτό για τη λογική του μοντέλου;

Λύση

(α) Για $X \sim U(0, \theta)$: $E(X) = \theta/2$. Άρα

$$\frac{\theta}{2} = \bar{X} \iff \boxed{\hat{\theta}^{\text{MoM}} = 2\bar{X}}.$$

(β)

$$\bar{x} = \frac{1,8+0,7+3,1+2,5+1,4}{5} = \frac{9,5}{5} = 1,9.$$

$$\hat{\theta}_{\text{obs}}^{\text{MoM}} = 2 \cdot 1,9 = 3,8 \text{ ώρες.}$$

(γ) Στο συγκεκριμένο δείγμα η εκτίμηση $\hat{\theta} = 3,8$ είναι μεγαλύτερη από $\max(X_i) = 3,1$, οπότε όλα είναι συνεπή.

Όμως: Είναι θεωρητικά δυνατόν ο MoM εκτιμητής να βγει μικρότερος από το $\max(X_i)$. Π.χ. αν το δείγμα ήταν 0,1, 0,2, 0,3, 0,1, 4,9:

$$\bar{X} = 1,12, \quad \hat{\theta}^{\text{MoM}} = 2,24 < 4,9 = \max(X_i).$$

Αυτό είναι **παράλογο**: η εκτίμηση του θ πρέπει να είναι τουλάχιστον $\max(X_i)$, αλλιώς η πιθανότητα του παρατηρούμενου δείγματος υπό την εκτίμηση θα ήταν 0 (δηλαδή η εκτίμηση δηλώνει ότι η παρατήρηση 4,9 ήταν αδύνατη).

Παρατήρηση

Αυτό το παράδειγμα δείχνει ένα κλασικό μειονέκτημα του MoM: μπορεί να παράγει εκτιμήσεις στο όριο του παραμετρικού χώρου ή και εκτός του λογικού εύρους. Ο εκτιμητής Μέγιστης Πιθανοφάνειας (Εβδομάδα 2) δίνει εδώ $\hat{\theta}^{\text{MLE}} = \max(X_i)$, το οποίο είναι πάντοτε λογικό.

2.3 Ομάδα Γ: Μέθοδος των Ροπών — Διπααραμετρικές Κατανομές

Άσκηση 13: MoM για Κανονική κατανομή — δύο παράμετροι

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

(α') Βρείτε τους MoM εκτιμητές για τα μ και σ^2 .

(β') Είναι αμερόληπτοι;

(γ') Είναι συνεπείς;

(δ') Σε δείγμα $n = 10$ ημερών δίνονται: $\sum x_i = 8,5$ και $\sum x_i^2 = 14,9$ (αποδόσεις σε

ποσοστιαίες μονάδες). Υπολογίστε τις εκτιμήσεις.

Λύση

(α) Έχουμε δύο άγνωστες παραμέτρους, οπότε χρειαζόμαστε δύο ροπές.

Πληθυσμιακές ροπές:

$$\mu'_1 = \mathbb{E}(X) = \mu, \quad \mu'_2 = \mathbb{E}(X^2) = \text{Var}(X) + [\mathbb{E}(X)]^2 = \sigma^2 + \mu^2.$$

Δειγματικές ροπές: $m_1 = \bar{X}$ και $m_2 = \frac{1}{n} \sum X_i^2$.

Σύστημα:

$$\begin{cases} \mu = \bar{X}, \\ \sigma^2 + \mu^2 = \frac{1}{n} \sum X_i^2. \end{cases}$$

Αντικαθιστούμε στη δεύτερη:

$$\sigma^2 = \frac{1}{n} \sum X_i^2 - \bar{X}^2 = \frac{1}{n} \sum (X_i - \bar{X})^2.$$

Άρα

$$\hat{\mu}^{\text{MoM}} = \bar{X}, \quad \hat{\sigma}^{2, \text{MoM}} = \frac{1}{n} \sum (X_i - \bar{X})^2.$$

(β) Από προηγούμενες ασκήσεις:

- Ο $\bar{X}$ είναι αμερόληπτος για το μ . ✓
- Ο $\hat{\sigma}^{2, \text{MoM}} = \frac{n-1}{n} S^2$ είναι μεροληπτικός για το σ^2 , με $\text{bias} = -\sigma^2/n$ (Άσκηση 2).

(γ) Είναι αμφότεροι συνεπείς:

- $\bar{X} \xrightarrow{\mathbb{P}} \mu$ (LLN).
- Από LLN, $\frac{1}{n} \sum X_i^2 \xrightarrow{\mathbb{P}} \mathbb{E}(X^2) = \sigma^2 + \mu^2$. Άρα $\hat{\sigma}^{2, \text{MoM}} = \frac{1}{n} \sum X_i^2 - \bar{X}^2 \xrightarrow{\mathbb{P}} (\sigma^2 + \mu^2) - \mu^2 = \sigma^2$.

(δ) Με $n = 10$, $\sum x_i = 8,5$, $\sum x_i^2 = 14,9$:

$$\bar{x} = 0,85, \quad \frac{1}{n} \sum x_i^2 = 1,49.$$

$$\hat{\mu}_{\text{obs}}^{\text{MoM}} = 0,85, \quad \hat{\sigma}_{\text{obs}}^{2, \text{MoM}} = 1,49 - 0,85^2 = 1,49 - 0,7225 = 0,7675.$$

Άρα μέση ημερήσια απόδοση 0,85% και διασπορά 0,7675 (τυπική απόκλιση $\approx 0,876\%$).

Άσκηση 14: MoM για Binomial με γνωστό n

Μια εταιρεία διεξάγει $n_0 = 20$ ανεξάρτητες κλήσεις πωλήσεων ανά ημέρα. Για κάθε ημέρα i , ο αριθμός επιτυχών πωλήσεων είναι $Y_i \sim \text{Binomial}(20, p)$, όπου p είναι η άγνωστη πιθανότητα επιτυχίας ανά κλήση. Παρατηρούμε τα ημερήσια αποτελέσματα για 15 ημέρες:

$$4, 6, 5, 3, 7, 5, 4, 6, 5, 4, 6, 5, 4, 5, 7.$$

- (α') Βρείτε τον MoM εκτιμητή του p .
 (β') Υπολογίστε αριθμητικά την εκτίμηση.
 (γ') Είναι αμερόληπτος και συνεπής;

Λύση

(α) Για $\text{Binomial}(n_0, p)$: $\mathbb{E}(Y) = n_0 p$. Άρα

$$n_0 p = \iff \hat{p}^{\text{MoM}} = \frac{\cdot}{n_0}.$$

(β)

$$= \frac{4+6+5+3+7+5+4+6+5+4+6+5+4+5+7}{15} = \frac{76}{15} \approx 5,067.$$

$$\hat{p}_{\text{obs}}^{\text{MoM}} = \frac{5,067}{20} \approx 0,2533 \approx 25,3\%.$$

(γ)

$$\mathbb{E}(\hat{p}^{\text{MoM}}) = \frac{\mathbb{E}(\cdot)}{n_0} = \frac{n_0 p}{n_0} = p. \quad \checkmark$$

$$\text{Var}(\hat{p}^{\text{MoM}}) = \frac{\text{Var}(\cdot)}{n_0^2} = \frac{n_0 p(1-p)/n}{n_0^2} = \frac{p(1-p)}{n \cdot n_0} \rightarrow 0. \quad \checkmark$$

Άρα αμερόληπτος και συνεπής.

Παρατήρηση

Παρατηρήστε ότι ο εκτιμητής $\hat{p} = \cdot / n_0$ είναι ισοδύναμος με την αναλογία επιτυχιών αν θεωρήσουμε όλες τις κλήσεις ($n \cdot n_0 = 300$ συνολικά) ως ένα Bernoulli δείγμα. Πράγματι, $\sum Y_i = 76$ και η αναλογία είναι $76/300 \approx 0,2533$ — ίδια εκτίμηση. Αυτό είναι αναμενόμενο: η Binomial είναι άθροισμα ανεξάρτητων Bernoulli.

Άσκηση 15: MoM για κατανομή Γάμμα — εφαρμογή σε διάρκεια δανείου

Η διάρκεια (σε έτη) μέχρι την πλήρη αποπληρωμή ενός τραπεζικού δανείου μοντελοποιείται ως $X_i \stackrel{\text{iid}}{\sim} \text{Gamma}(\alpha, \beta)$ με σ.π.π.

$$f_X(x) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x/\beta}, \quad x > 0, \alpha, \beta > 0.$$

Είναι γνωστό ότι $\mathbb{E}(X) = \alpha\beta$ και $\text{Var}(X) = \alpha\beta^2$.

Σε τυχαίο δείγμα $n = 8$ δανείων, παρατηρήθηκαν τα παρακάτω (σε έτη):

$$4,2, 6,5, 3,8, 5,1, 7,0, 4,9, 6,2, 5,5.$$

- (α') Βρείτε τις γενικές εκφράσεις των MoM εκτιμητών για α και β .
 (β') Υπολογίστε αριθμητικά τις εκτιμήσεις.

Λύση**(α)** Έχουμε:

$$\mu'_1 = \mathbb{E}(X) = \alpha\beta, \quad \mu'_2 = \mathbb{E}(X^2) = \text{Var}(X) + [\mathbb{E}(X)]^2 = \alpha\beta^2 + (\alpha\beta)^2 = \alpha\beta^2(1 + \alpha).$$

Σύστημα MoM:

$$\begin{cases} \alpha\beta = \bar{X} = m_1, \\ \alpha\beta^2(1 + \alpha) = m_2. \end{cases}$$

Από την πρώτη: $\beta = m_1/\alpha$. Αντικαθιστούμε στη δεύτερη:

$$\alpha \cdot \frac{m_1^2}{\alpha^2}(1 + \alpha) = m_2 \iff \frac{m_1^2(1 + \alpha)}{\alpha} = m_2.$$

Λύνουμε ως προς α :

$$m_1^2 + m_1^2\alpha = m_2\alpha \iff m_1^2 = \alpha(m_2 - m_1^2) \iff \alpha = \frac{m_1^2}{m_2 - m_1^2}.$$

Παρατηρούμε ότι $m_2 - m_1^2 = \frac{1}{n} \sum X_i^2 - \bar{X}^2 = \hat{\sigma}^2, \text{MoM}$ (η δειγματική διασπορά χωρίς διόρθωση).**Τελικοί τύποι:**

$$\hat{\alpha}^{\text{MoM}} = \frac{\bar{X}^2}{\hat{\sigma}^2, \text{MoM}}, \quad \hat{\beta}^{\text{MoM}} = \frac{\hat{\sigma}^2, \text{MoM}}{\bar{X}}.$$

(Λογική: για την Gamma, ο λόγος $\mathbb{E}(X)^2 / \text{Var}(X) = \alpha$ — άρα οι δειγματικές αντίστοιχες δίνουν τον εκτιμητή του α .)**(β)** Από τα δεδομένα:

$$\sum x_i = 4,2 + 6,5 + 3,8 + 5,1 + 7,0 + 4,9 + 6,2 + 5,5 = 43,2.$$

$$\begin{aligned} \sum x_i^2 &= 4,2^2 + 6,5^2 + 3,8^2 + 5,1^2 + 7,0^2 + 4,9^2 + 6,2^2 + 5,5^2 \\ &= 17,64 + 42,25 + 14,44 + 26,01 + 49,00 + 24,01 + 38,44 + 30,25 = 242,04. \end{aligned}$$

$$\bar{x} = \frac{43,2}{8} = 5,4, \quad m_2 = \frac{242,04}{8} = 30,255.$$

$$\hat{\sigma}_{\text{obs}}^2, \text{MoM} = m_2 - \bar{x}^2 = 30,255 - 29,16 = 1,095.$$

$$\hat{\alpha}_{\text{obs}}^{\text{MoM}} = \frac{5,4^2}{1,095} = \frac{29,16}{1,095} \approx 26,63.$$

$$\hat{\beta}_{\text{obs}}^{\text{MoM}} = \frac{1,095}{5,4} \approx 0,2028.$$

Έλεγχος**Επαλήθευση:** Από τις εκτιμήσεις θα πρέπει $\hat{\alpha}\hat{\beta} \approx \bar{x}$.

$$26,63 \times 0,2028 \approx 5,401 \approx 5,4 = \bar{x}. \quad \checkmark$$

Επίσης $\hat{\alpha}\hat{\beta}^2 \approx \hat{\sigma}^2, \text{MoM}$:

$$26,63 \times 0,2028^2 \approx 26,63 \times 0,04113 \approx 1,095. \quad \checkmark$$

Οι εκτιμήσεις είναι εσωτερικά συνεπείς.

Παρατήρηση

Ερμηνεία: Με $\hat{\alpha} \approx 26,6$, το shape parameter είναι αρκετά μεγάλο, δείχνοντας ότι η κατανομή των διαρκείων δανείων είναι σχετικά «συγκεντρωμένη» γύρω από τη μέση τιμή (όσο μεγαλώνει το α , η Gamma πλησιάζει κανονική). Το $\hat{\beta} \approx 0,2$ είναι το scale parameter.

Πρακτική παρατήρηση: για μικρά δείγματα (όπως $n = 8$), οι MoM εκτιμητές μπορεί να έχουν μεγάλη μεταβλητότητα. Στην Εβδομάδα 2 θα δούμε ότι ο MLE εκτιμητής του α είναι πιο σταθερός αλλά απαιτεί αριθμητική λύση.

3 Σύνοψη Τύπων

Σύνοψη Εβδομάδας 1

Βασικοί ορισμοί:

- Εκτιμητής: $\hat{\theta} = T(X_1, \dots, X_n)$ — τυχαία μεταβλητή.
- Εκτίμηση: αριθμός $\hat{\theta}_{\text{obs}}$ για συγκεκριμένο δείγμα.
- Δειγματοληπτική κατανομή: η κατανομή του $\hat{\theta}$.

Ιδιότητες εκτιμητών:

$$\text{bias}(\hat{\theta}) = \mathbb{E}(\hat{\theta}) - \theta$$

$$\text{Var}(\hat{\theta}) = \mathbb{E}[(\hat{\theta} - \mathbb{E}(\hat{\theta}))^2]$$

$$\text{SE}(\hat{\theta}) = \sqrt{\text{Var}(\hat{\theta})}$$

$$\text{MSE}(\hat{\theta}) = \mathbb{E}[(\hat{\theta} - \theta)^2] = \text{Var}(\hat{\theta}) + \text{bias}(\hat{\theta})^2$$

$$\text{RE}(\hat{\theta}_1, \hat{\theta}_2) = \text{Var}(\hat{\theta}_2) / \text{Var}(\hat{\theta}_1) \quad (\text{και οι δύο αμερόληπτοι})$$

Συνέπεια: $\hat{\theta}_n \xrightarrow{\mathbb{P}} \theta$. Επαρκές κριτήριο: $\mathbb{E}(\hat{\theta}_n) \rightarrow \theta$ και $\text{Var}(\hat{\theta}_n) \rightarrow 0$.

Chebyshev: $\mathbb{P}(|\hat{\theta} - \mathbb{E}(\hat{\theta})| \geq \varepsilon) \leq \text{Var}(\hat{\theta}) / \varepsilon^2$.

Δειγματικός μέσος (για iid):

$$\mathbb{E}(\bar{X}) = \mu, \quad \text{Var}(\bar{X}) = \sigma^2/n, \quad \text{SE}(\bar{X}) = \sigma/\sqrt{n}.$$

Δειγματική διασπορά:

$$S^2 = \frac{1}{n-1} \sum (X_i - \bar{X})^2 \quad (\text{αμερόληπτη}), \quad \hat{\sigma}^2 = \frac{1}{n} \sum (X_i - \bar{X})^2 \quad (\text{μεροληπτική, minMSE}).$$

Δειγματική αναλογία:

$$\hat{p} = \bar{X}, \quad \mathbb{E}(\hat{p}) = p, \quad \text{Var}(\hat{p}) = p(1-p)/n.$$

Μέθοδος Ροπών: εξισώνουμε $\mu'_r(\theta_1, \dots, \theta_k) = m_r$ για $r = 1, \dots, k$ και λύνουμε.
Συνηθισμένοι MoM εκτιμητές:

Κατανομή	Παράμετρος(οι)	MoM εκτιμητής
Bernoulli(p)	p	$\hat{p} = \bar{X}$
Binomial(n_0, p), n_0 γνωστό	p	$\hat{p} = \bar{X}/n_0$
Poisson(λ)	λ	$\hat{\lambda} = \bar{X}$
Geometric(θ)	θ	$\hat{\theta} = 1/\bar{X}$
Exp(rate λ)	λ	$\hat{\lambda} = 1/\bar{X}$
$U(0, \theta)$	θ	$\hat{\theta} = 2\bar{X}$
$N(\mu, \sigma^2)$	μ, σ^2	$\hat{\mu} = \bar{X}, \hat{\sigma}^2 = \frac{1}{n} \sum (X_i - \bar{X})^2$
Gamma(α, β)	α, β	$\hat{\alpha} = \bar{X}^2/\hat{\sigma}^2, \hat{\beta} = \hat{\sigma}^2/\bar{X}$

Τι ακολουθεί στην Εβδομάδα 2

Στην επόμενη εβδομάδα θα προχωρήσουμε σε:

- **Πληροφορία Fisher και Ανισότητα Cramer-Rao** — το κάτω φράγμα για τη διασπορά κάθε αμερόληπτου εκτιμητή. Αυτό θα μάς δώσει μια *απόλυτη* αναφορά για το πόσο καλός μπορεί να είναι ένας αμερόληπτος εκτιμητής.
- **Επάρκεια** και θεώρημα παραγοντοποίησης Fisher-Neyman — πότε ένα στατιστικό «συμπυκνώνει» όλη την πληροφορία του δείγματος για μια παράμετρο.
- **Μέθοδος Μέγιστης Πιθανοφάνειας (MLE)** — η κυρίαρχη τεχνική εκτίμησης σε όλη την οικονομετρία και τη στατιστική μάθηση. Σχεδόν κάθε σύγχρονο μοντέλο (logit, probit, GLM, ARIMA, GARCH) εκτιμάται με MLE.
- **Ασυμπτωτικές ιδιότητες MLE**: συνέπεια, ασυμπτωτική κανονικότητα $\sqrt{n}(\hat{\theta}_{\text{MLE}} - \theta) \xrightarrow{d} N(0, I(\theta)^{-1})$, ασυμπτωτική αποτελεσματικότητα.
- **Delta method** — πώς να βρίσκουμε την κατανομή μη γραμμικών συναρτήσεων εκτιμητών. Κρίσιμο για ελαστικότητες, παράγωγα δοκιμών, οικονομικούς δείκτες.