ΣΤΑΤΙΣΤΙΚΗ ΙΙΙ

Θεωρία 03

Δήμητρα Κυριακοπούλου

Τμήμα Οικονομικών Επιστημών

Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών

- Εκτός από τις ιδιότητες των εκτιμητριών που έχουμε ήδη αναφέρει υπάρχουν και οι ασυμπτωτικές ιδιότητες των εκτιμητών.

- Αυτές είναι:

  - η συνέπεια ή σύγκλιση
  - η ασυμπτωτική αποτελεσματικότητα
  - η ασυμπτωτική κανονικότητα.

- Να σημειωθεί ότι οι παραπάνω ιδιότητες αφορούν σε μεγάλα δείγματα, δηλαδή όταν το πλήθος των παρατηρήσεων n θεωρητικά τείνει στο άπειρο, για αυτό το λόγο θα αναφερόμαστε σε μια ακολουθία τυχαίων μεταβλητών που θα συμβολίζουμε ως θ_n.

- Οι μορφές σύγκλισης που μας ενδιαφέρουν είναι η σύγκλιση σε πιθανότητα $\xrightarrow{p}$ και η σύγκλιση σε κατανομή $\xrightarrow{D}$.

- Χρειάζεται να γνωρίζουμε δύο βασικά εργαλεία από την ασυμπτωτική θεωρία: τον (ασθενή) Νόμο των Μεγάλων Αριθμών και το Κεντρικό Οριακό Θεώρημα, που δίνονται στο τέλος αυτής της διάλεξης.

Συνέπεια (Consistency)

- Αν ισχύει ότι:

$$\lim_{n\to\infty} P\left(\left|\hat{\theta}_n - \theta\right| < \varepsilon\right) = 1, \quad \forall \varepsilon > 0$$

δηλαδή με πολύ υψηλή πιθανότητα οι τιμές της εκτιμήτριας $\hat{\theta}_n$ είναι επαρκώς πολύ κοντά στην πραγματική τιμή της παραμέτρου θ,

τότε:

$$\hat{\theta}_n \xrightarrow{p} \theta, \quad \forall \theta$$

- Αυτός είναι ο ορισμός της <u>ασθενούς συνέπειας</u> (weak consistency).
- Επίσης, θα λέμε ότι το όριο πιθανότητας είναι μοναδικό (unique), δηλαδή εάν $Y_n \to \alpha$ τότε δεν μπορεί να συγκλίνει σε καμία άλλη σταθερά β.

- Ο ορισμός της ισχυρής συνέπειας (strong consistency) είναι ο εξής:

$$\hat{\theta_n} \stackrel{a.s.}{\to} \theta, \quad \forall \theta$$

όπου $a.s.$ σημαίνει «σχεδόν βέβαια» (almost surely), με πιθανότητα 1:

$$P\left[\lim_{n\to\infty} \hat{\theta_n} = \theta\right] = 1.$$

- Κάθε ισχυρά συνεπής εκτιμήτρια είναι και ασθενώς συνεπής.

Ιδιότητες της Συνέπειας

- Γραμμικότητα:

$$\theta_n^{(1)} \xrightarrow{p} \theta_1$$

$$\theta_n^{(2)} \xrightarrow{p} \theta_2$$

$\Rightarrow$

$$\alpha\theta_n^{(1)} + \beta\theta_n^{(2)} \xrightarrow{p} \alpha\theta_1 + \beta\theta_2.$$

- <u>Συνέχεια:</u>

  Αν υπάρχει συνάρτηση $g(\cdot)$ συνεχής και ισχύει ότι:

$$\hat{\theta_n} \xrightarrow{p} \theta,$$

τότε:

$$g\left(\hat{\theta_n}\right) \xrightarrow{p} g(\theta).$$

Αυτή είναι μια πολύ σημαντική ιδιότητα μιας συνεπούς εκτιμήτριας, δηλαδή ότι <u>παραμένει αναλλοίωτη (invariant)</u> κάτω από συνεχείς μετασχηματισμούς.

Σε αντίθεση με μια αμερόληπτη εκτιμήτρια, για την οποία ισχύει ότι $E\left(\hat{\theta}\right) = \theta,\ \forall\theta$, όπου <u>γενικά δεν συνεπάγεται</u> ότι $E\left[g\left(\hat{\theta}\right)\right] = g(\theta)$, <u>εκτός κι αν</u> η συνάρτηση $g(\cdot)$ είναι γραμμική σε σχέση με την παράμετρο θ.

• Πόρισμα:

Αν $\hat{\theta}_n$ είναι μια εκτιμήτρια της παραμέτρου θ για την οποία ισχύει:

$$\lim_{n\to\infty} E\left(\hat{\theta}_n\right) = \theta,$$

δηλαδή το $\hat{\theta}_n$ είναι ασυμπτωτικά αμερόληπτη εκτιμήτρια του θ, και

$$\lim_{n\to\infty} Var\left(\hat{\theta}_n\right) = 0,$$

τότε το $\hat{\theta}_n$ είναι συνεπής εκτιμήτρια της παραμέτρου θ.

- <u>Παράδειγμα:</u>

  Έστω $(X_1, X_2, ..., X_n)$ ένα τυχαίο δείγμα από κατανομή $f(x;\theta)$ με $E(X) = \theta$. Τότε ο δειγματικός μέσος:

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

  είναι συνεπής εκτιμήτρια της παραμετρικής συνάρτησης θ αφού είναι αμερόληπτη και επί πλέον ισχύει:

$$Var\left(\hat{\theta_n}\right) = \frac{Var(X)}{n} \rightarrow_{n\rightarrow\infty} 0.$$

Ανισότητα Markov

Αν X είναι μια μη αρνητική τυχαία μεταβλητή, δηλαδή $P(X > 0) = 1$, με κατανομή $f(x;\theta)$ τότε ισχύει:

$$P(X \geq \varepsilon) \leq \frac{E(X)}{\varepsilon}, \quad \forall \varepsilon > 0.$$

- Εάν υποθέσουμε ότι η X είναι συνεχής, τότε για κάθε $\varepsilon > 0$ ισχύει ότι:

$$\begin{aligned} E(X) &= \int_0^\infty x f(x)\, dx \geq \int_\varepsilon^\infty x f(x)\, dx \\ &\geq \varepsilon \int_\varepsilon^\infty f(x)\, dx \\ &= \varepsilon P(X \geq \varepsilon), \end{aligned}$$

κι έτσι προκύπτει η παραπάνω ανισότητα Markov.

- Παρατηρείστε ότι εάν αντικαταστήσουμε την τυχαία μεταβλητή X με την τυχαία μεταβλητή $Y = g(X)$, τέτοια ώστε $P[g(X) > 0] = 1$ και η συνάρτηση g είναι αύξουσα στο στήριγμα της X, θα έχουμε:

$$E[g(X)] = \int_0^\infty g(x) f(x) dx \geq \int_\varepsilon^\infty g(x) f(x) dx$$

$$\geq g(\varepsilon) \int_\varepsilon^\infty f(x) dx$$

$$= g(\varepsilon) P(X \geq \varepsilon),$$

όπου χρησιμοποιήσαμε ότι

$$x \geq \varepsilon \Rightarrow g(x) \geq g(\varepsilon)$$

κι έτσι:

$$P(X \geq \varepsilon) = P[g(X) \geq g(\varepsilon)]$$
$$\leq \frac{E[g(X)]}{g(\varepsilon)}.$$

- Παράδειγμα:

Εάν

$$g(x) = e^{tx},\ t > 0$$

τότε για $x \geq 0$:

$$\begin{aligned} P(X \geq \varepsilon) &= P\left[e^{tX} \geq e^{t\varepsilon}\right] \\ &\leq \frac{E\left[e^{tX}\right]}{e^{t\varepsilon}} \\ &= e^{-t\varepsilon} M_X(t). \end{aligned}$$

Συνέπεια μέσω ΜΤΣ

- <u>Πρόταση:</u>

  Έστω η εκτιμήτρια $\hat{\theta_n}$ με
  $$MSE\left(\hat{\theta_n}\right) = \left[Bias\left(\hat{\theta_n}\right)\right]^2 + Var\left(\hat{\theta_n}\right).$$
  Τότε η εκτιμήτρια $\hat{\theta_n}$ καλείται συνεπής αν για το μέσο τετραγωνικό σφάλμα (ΜΤΣ) ισχύει:
  $$\lim_{n\to\infty} MSE\left(\hat{\theta_n}\right) = 0.$$

- <u>Απόδειξη:</u>

$$P\left(\left|\hat{\theta_n}-\theta\right|\geq\varepsilon\right)=P\left(\left|\hat{\theta_n}-\theta\right|^2\geq\varepsilon^2\right)$$
$$\leq\frac{E\left|\hat{\theta_n}-\theta\right|^2}{\varepsilon^2},$$

από την ανισότητα Markov, και ισχύει ότι:

$$\frac{E\left|\hat{\theta_n}-\theta\right|^2}{\varepsilon^2}=\frac{MSE\left(\hat{\theta_n}\right)}{\varepsilon^2}.$$

Αλλά

$$\lim_{n\to\infty}MSE\left(\hat{\theta_n}\right)=0,$$

άρα:

$$\lim_{n\to\infty}P\left(\left|\hat{\theta_n}-\theta\right|\geq\varepsilon\right)=0\Rightarrow$$

η εκτιμήτρια $\hat{\theta_n}$ είναι συνεπής για το θ.

- <u>Παράδειγμα 1:</u>

  Ο δειγματικός μέσος $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$.

  Γνωρίζουμε ότι είναι αμερόληπτη εκτιμήτρια του μέσου του πληθυσμού μ, οπότε:

$$MSE\left(\hat{\theta_n}\right) = 0 + \frac{\sigma^2}{n}$$
$$\rightarrow_{n\rightarrow\infty} 0$$

  και συνεπώς η εκτιμήτρια $\bar{X}$ είναι συνεπής για την παράμετρο μ.

- Παράδειγμα 2:

$$\hat{\sigma^2} = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2$$

Ισχύει ότι:

$$\begin{aligned} Bias\left(\hat{\sigma^2}\right) &= \frac{n-1}{n}\sigma^2 - \sigma^2 \\ &= -\frac{1}{n}\sigma^2 \end{aligned}$$

και συνεπώς:

$$\lim_{n\to\infty}\left[Bias\left(\hat{\sigma^2}\right)\right]^2 = 0.$$

Επίσης, ισχύει:

$$\begin{aligned} Var\left(\hat{\sigma^2}\right) &= \frac{1}{n}\left[E\left(X-\mu\right)^4 - \frac{n-3}{n-1}\sigma^4\right] \\ &\to_{n\to\infty} 0. \end{aligned}$$

• <u>Παράδειγμα 3:</u>

$$\hat{\mu_n} = \frac{1}{n+1}\sum_{j=1}^{n} X_j$$

Ισχύει ότι:

$$E\left(\hat{\mu_n}\right) = \frac{n}{n+1}\mu$$

και συνεπώς:

$$\begin{aligned} Bias\left(\hat{\mu_n}\right) &= \frac{n}{n+1}\mu - \mu \\ &= -\frac{\mu}{n+1} \end{aligned}$$

και

$$\lim_{n\to\infty}\left[Bias\left(\hat{\mu_n}\right)\right]^2 = 0.$$

Επίσης:

$$\begin{aligned} Var\left(\hat{\mu_n}\right) &= \frac{1}{(n-1)^2}\sigma^2 \\ &\to_{n\to\infty} 0. \end{aligned}$$

- Παράδειγμα:

  Έστω $X_1, X_2, ..., X_n$ τυχαίο δείγμα από την κανονική κατανομή $N\left(\mu, \sigma^2\right)$. Θεωρούμε την κεντρική δειγματική ροπή 2ης τάξης:

$$\hat{\theta_n} = m_2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2.$$

  Τότε, ως γνωστόν, η τυχαία μεταβλητή:

$$\frac{n\hat{\theta_n}}{\sigma^2} = \sum_{i=1}^{n}\left(\frac{X_i - \bar{X}}{\sigma}\right)^2 \sim X_{n-1}^2$$

οπότε:

$$E\left(\frac{n\hat{\theta_n}}{\sigma^2}\right) = n-1, \quad Var\left(\frac{n\hat{\theta_n}}{\sigma^2}\right) = 2\left(n-1\right) \Rightarrow$$

$$E\left(\hat{\theta_n}\right) = \frac{n-1}{n}\sigma^2, \quad Var\left(\hat{\theta_n}\right) = \frac{2\left(n-1\right)}{n^2}\sigma^4$$

και καταλήγουμε στο συμπέρασμα ότι η m_2 είναι συνεπής εκτιμήτρια του σ^2.

- Παράδειγμα:

Έστω $X_1, X_2, ..., X_n$ τυχαίο δείγμα από την κατανομή Cauchy με συνάρτηση πυκνότητας:

$$f(x;\theta) = \frac{1}{\pi} \cdot \frac{1}{1 + (x-\theta)^2}.$$

Τότε ο δειγματικός μέσος $\hat{\theta_n} = \bar{X}$ έχει ακριβώς την ίδια κατανομή με τις παρατηρήσεις X_i και για οποιαδήποτε παραμετρική συνάρτηση $g(\theta)$ η πιθανότητα:

$$P\left[\left|\hat{\theta_n} - g(\theta)\right| > \varepsilon\right]$$

είναι ανεξάρτητη του n και μη μηδενική. Άρα η εκτιμήτρια $\hat{\theta_n}$ δεν είναι συνεπής.

Νόμος των Μεγάλων Αριθμών (Law of Large Numbers)

Έστω $(X_1, X_2, X_3, ...)$ ένα τυχαίο δείγμα, μια άπειρη ακολουθία ανεξάρτητων τυχαίων μεταβλητών όπου κάθε μία έχει ίδια κατανομή με μια τυχαία μεταβλητή X, από πληθυσμό με μέση τιμή $E(X) = \mu < \infty$ (πεπερασμένη), τότε:

$$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i \xrightarrow{p} \mu = E(X), \quad n \to \infty$$

δηλαδή ο δειγματικός μέσος συγκλίνει στο μέσο του πληθυσμού, που ισοδυναμεί με την έκφραση:

$$\lim_{n \to \infty} P\left[\left(\bar{X} - \mu\right) > \varepsilon\right] = 0, \quad \forall \varepsilon > 0.$$

- <u>Απόδειξη:</u>

$$
\begin{aligned}
P\left(\left|\bar{X}-\mu\right|>\varepsilon\right) &= P\left(\left|\bar{X}-E\left(\bar{X}\right)\right|>\varepsilon\right)\\
&\leq \frac{Var\left(\bar{X}\right)}{\varepsilon^2}\\
&= \frac{\sigma^2}{n\varepsilon^2} \to_{n\to\infty} 0,
\end{aligned}
$$

κάνοντας χρήση στο πρώτο αποτέλεσμα ότι:

$$
E\left(\bar{X}\right) = E\left(\frac{1}{n}\sum_i X_i\right) = \frac{1}{n}\sum_i E\left(X_i\right) = \mu.
$$

Ο Νόμος των Μεγάλων Αριθμών έρχεται με <u>γενίκευση:</u>

$$\frac{1}{n}\sum_{i=1}^{n} g\left(X_i\right) \xrightarrow{p} E\left[g\left(X\right)\right].$$

Για παράδειγμα:

$$\frac{1}{n}\sum_{i=1}^{n} X_i^2 \xrightarrow{p} E\left(X^2\right).$$

- <u>Παράδειγμα:</u>

  Έστω $(X_1, X_2, ..., X_n)$ ένα τυχαίο δείγμα από την κανονική κατανομή $N\left(\mu, \sigma^2\right)$. Να δείξετε ότι η κεντρική δειγματική ροπή δεύτερης τάξης είναι συνεπής εκτιμήτρια του σ^2:

$$\hat{\sigma^2} := m_2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2 \xrightarrow{p} \sigma^2$$

  Ξεκινάμε με το εξής:

$$\hat{\sigma^2} = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)^2 .$$

Τότε:

$$\theta_n^{(1)} = \frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{p} E\left(X\right) = \theta^{(1)}$$
$$\theta_n^{(2)} = \frac{1}{n}\sum_{i=1}^{n} X_i^2 \xrightarrow{p} E\left(X^2\right) = \theta^{(2)}$$
$$\theta_n^{(3)} = \left(\bar{X}\right)^2 \xrightarrow{p} \left(E\left(X\right)\right)^2 = \theta^{(3)}$$

Και τέλος:

$$\begin{aligned}\theta_n^{(2)} - \theta_n^{(3)} &\xrightarrow{p} \theta^{(2)} - \theta^{(3)}\\ &= E\left(X^2\right) - \left(E\left(X\right)\right)^2\\ &= Var\left(X\right)\\ &= \sigma^2.\end{aligned}$$

Κεντρικό Οριακό Θεώρημα

- Το Κεντρικό Οριακό Θεώρημα (Central Limit Theorem -CLT) εξηγεί το μεγάλο εύρος εφαρμογής της κανονικής κατανομής και το πώς συνδέει οποιαδήποτε κατανομή με την κανονική.
- Η διατύπωσή του είναι η εξής:
    - αν από έναν πληθυσμό που ακολουθεί μια οποιαδήποτε κατανομή, με μέση τιμή μ και διακύμανση σ^2, επιλέξουμε τυχαία δείγματα μεγέθους n και υπολογίσουμε τους μέσους τους, τότε για μεγάλα n, δηλαδή $n \to \infty$, η κατανομή αυτών των μέσων (των δειγματικών) είναι κατά προσέγγιση κανονική κατανομή με μέση τιμή μ και διακύμανση $\frac{\sigma^2}{n}$
    - όσο πιο μεγάλο είναι το μέγεθος n των δειγμάτων, τόσο καλύτερη (ακριβέστερη) είναι η προσέγγιση της κατανομής των δειγματικών μέσων από την κανονική κατανομή.

- Μια ισοδύναμη διατύπωση του Θεωρήματος είναι η εξής:
    - αν από έναν πληθυσμό που ακολουθεί οποιαδήποτε κατανομή, με μέση τιμή μ και διακύμανση σ^2, επιλέξουμε τυχαία δείγματα μεγέθους n και υπολογίσουμε το άθροισμα των παρατηρήσεων κάθε δείγματος, τότε για μεγάλα n, δηλαδή $n \to \infty$, η κατανομή αυτών των αθροισμάτων είναι κατά προσέγγιση κανονική κατανομή με μέση τιμή $n\mu$ και διακύμανση $n\sigma^2$, δηλαδή αν συμβολίσουμε με S_n την τυχαία μεταβλητή που εκφράζει αυτά τα αθροίσματα, τότε κατά προσέγγιση:

$$S_n \sim N\left(n\mu, n\sigma^2\right)$$

- Η τυπική απόκλιση της κατανομής των δειγματικών μέσων:

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$$

ονομάζεται **τυπικό σφάλμα (standard error)**.

- Η κατανομή των δειγματικών μέσων (και των αθροισμάτων των παρατηρήσεων κάθε δείγματος) είναι κανονική κατανομή ανεξάρτητα από την τιμή του n:
  - αν $(X_1, X_2, ..., X_n)$ είναι ανεξάρτητες και ισόνομες (i.i.d.) τυχαίες μεταβλητές με $Var\,(X) = \sigma^2 < \infty$, τότε:

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n} \sim N\left(\mu, \frac{\sigma^2}{n}\right), \quad n \to \infty$$

και

$$S_n = \sum_{i=1}^{n} X_i \sim N\left(n\mu, n\sigma^2\right), \quad n \to \infty.$$

- Το προηγούμενο αποτέλεσμα του Θεωρήματος ισοδυναμεί με:

$$\sqrt{n}\left(\bar{X}-\mu\right) \xrightarrow{D} N\left(0,\sigma^2\right), \quad n\to\infty$$

ή

$$\sqrt{n}\left(\frac{\bar{X}-\mu}{\sigma}\right) \xrightarrow{D} N\left(0,1\right), \quad n\to\infty,$$

το οποίο σημαίνει ότι:

$$\lim_{n\to\infty} P\left[\sqrt{n}\left(\frac{\bar{X}-\mu}{\sigma}\right) \leq x\right] = \Phi\left(x\right),$$

όπου $\Phi\left(x\right)$ η (αθροιστική) συνάρτηση κατανομής πιθανότητας.

- Καθώς το σ^2 είναι γενικά **άγνωστο**, θα πρέπει να το αντικαταστήσουμε με οποιαδήποτε συνεπή εκτιμήτρια.

- Το συμπέρασμα είναι ότι για μέγαλο δείγμα μπορούμε προσεγγιστικά να βγάζουμε την κατανομή του $\bar{X}$ όπως στην περίπτωση που έχουμε κανονικό πληθυσμό και γνωστή διακύμανση, χρησιμοποιώντας την εκτίμηση της διακύμανσης.

- <u>Λήμμα:</u>

Αν

$$\sqrt{n}\left(\hat{\theta_n}-\theta\right) \xrightarrow{D} N\left(0,\sigma_\theta^2\right),$$

όπου σ_θ^2 είναι η ασυμπτωτική διακύμανση του $\hat{\theta_n}$, και $g(\theta)$ συνεχής με $g'(\theta) \neq 0$, κι έστω ότι:

$$\varphi = g(\theta)$$

και

$$\hat{\varphi_n} = g\left(\hat{\theta_n}\right).$$

- Τότε:

$$\begin{aligned}\sqrt{n}\left(\hat{\varphi_n}-\varphi\right) &= \sqrt{n}\left(g\left(\hat{\theta_n}\right)-g(\theta)\right)\\ &\xrightarrow{D} N\left(0,\sigma_\varphi^2\right),\end{aligned}$$

όπου

$$\sigma_\varphi^2 = \sigma_\theta^2\left(\frac{dg}{d\theta}\right)^2.$$