ΣΤΑΤΙΣΤΙΚΗ ΙΙΙ

Θεωρία 06

Δήμητρα Κυριακοπούλου

Τμήμα Οικονομικών Επιστημών

Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών

- Στο πλαίσιο της Στατιστικής Επαγωγής, υποθέτουμε κάθε φορά ότι έχουμε ένα παραμετρικό μοντέλο για κάποιο τυχαίο πείραμα που πραγματοποιούμε, όπου το μοντέλο αυτό αναφέρεται σε κάποια (δηλαδή περιέχει κάποια) προς εκτίμηση άγνωστη παράμετρο θ.

- Στη συνέχεια, το πείραμα εκτελείται και παρατηρείται ότι έχει συμβεί κάποιο ενδεχόμενο. Αυτό πρακτικά σημαίνει ότι έχουν συγκεντρωθεί κάποια δεδομένα, δηλαδή ένα τυχαίο δείγμα.

- Αυτό που μας ενδιαφέρει είναι, με τη χρήση των δεδομένων αυτών, να εκτιμηθεί η τιμή της παραμέτρου θ.

- Χρησιμοποιώντας το συγκεκριμένο πιθανοθεωρητικό μοντέλο και με τους νόμους των πιθανοτήτων, είναι δυνατόν να καθορίσουμε την πιθανότητα, με άλλα λόγια την **πιθανοφάνεια**, του ενδεχομένου που έχουμε παρατηρήσει, δηλαδή του τυχαίου δείγματος.

- Η πιθανότητα αυτή θα είναι συνάρτηση της άγνωστης παραμέτρου θ.

Μέγιστη Πιθανοφάνεια (maximum likelihood)

- Η <u>πιθανοφάνεια</u> του τυχαίου δείγματος $(X_1, X_2, ..., X_n)$ ορίζεται ως η συνάρτηση:

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

δηλαδή η από κοινού κατανομή του δείγματος που είναι συνάρτηση της παραμέτρου θ για δοσμένη τιμή του δείγματος, και ισχύει επί της ουσίας ότι:

$$L(\theta) := L(\theta \mid x_1, x_2, ..., x_n).$$

- Η <u>εκτιμήτρια μέγιστης πιθανοφάνειας (maximum likelihood estimator - MLE)</u>, $\hat{\theta}$, προκύπτει από τη λύση του προβλήματος μεγιστοποίησης:

$$\hat{\theta} = \arg\max_{\theta \in \Theta} \mathcal{L}(\theta),$$

όπου

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

- **Παρατηρήσεις:**

1. Στην περίπτωση που η συνάρτηση πιθανοφάνειας $L(\theta)$ είναι διαφορίσιμη, τότε ο MLE $\hat{\theta}$ είναι η λύση της εξίσωσης πιθανοφάνειας:

$$\frac{\partial \log L(\theta)}{\partial \theta} = 0,$$

που ικανοποιεί τη σχέση:

$$\frac{\partial^2 \log L(\theta)}{\partial \theta^2} < 0,$$

δηλαδή λύνουμε το πρόβλημα μεγιστοποίησης στα σημεία που μηδενίζεται η πρώτη παράγωγος της συνάρτησης $\log L(\theta)$ και επιβεβαιώνουμε ότι η ρίζα που βρήκαμε είναι ένα τοπικό μέγιστο.

2. Για την εύρεση του μεγίστου της πιθανοφάνειας υπάρχουν οι ακόλουθες περιπτώσεις:

   - να μην υπάρχει πεπερασμένο μέγιστο,
   - να υπάρχει ακριβώς ένα μέγιστο,
   - να υπάρχουν περισσότερα από ένα μέγιστα.

• <u>Παράδειγμα:</u>

Έστω τυχαίο δείγμα (X_1, X_2) με συνάρτηση κατανομής

$$f(x;\theta) = \frac{1}{\pi\left[1+(x-\theta)^2\right]}.$$

Τότε:

$$\begin{aligned}\prod_{\iota=1}^{2} f(x_i;\theta) &= \frac{1}{\pi\left[1+(X_1-\theta)^2\right]}\frac{1}{\pi\left[1+(X_2-\theta)^2\right]}\\ &= \frac{1}{\pi^2}\frac{1}{\left[1+(X_1-\theta)^2\right]}\frac{1}{\left[1+(X_2-\theta)^2\right]}\end{aligned}$$

και

$$\mathcal{L} = -\log\pi^2 - \log\left[1+(X_1-\theta)^2\right] - \log\left[1+(X_2-\theta)^2\right].$$

Η λύση θα προκύπτει από:

$$
\begin{aligned}
\frac{\partial \mathcal{L}}{\partial \theta} = 0 &\iff \\
\frac{X_1-\theta}{1+(X_1-\theta)^2} + \frac{X_2-\theta}{1+(X_1-\theta)^2} = 0 &\iff \\
(X_1+X_2-2\theta)\left[1+(X_1-\theta)(X_2-\theta)\right] = 0 &\iff
\end{aligned}
$$

$$\hat{\theta}_1 = \frac{X_1+X_2}{2},$$

όπου έχουμε την εξίσωση:

$$\theta^2 - (X_1+X_2)\,\theta + (1+X_1X_2) = 0$$

της οποίας η Διακρίνουσα δίνεται ίση με:

$$
\begin{aligned}
\Delta &= (X_1+X_2)^2 - 4\,(1+X_1X_2) \\
&= (X_1-X_2)^2 - 4
\end{aligned}
$$

Και

$$\hat{\theta}_{2,3} = \frac{(X_1 + X_2) \pm \sqrt{\Delta}}{2}.$$

- Αν $\Delta < 0 \iff |X_1 - X_2| < 2$, τότε μία ρίζα, $\hat{\theta}_1$
- Αν $\Delta = 0 \iff |X_1 - X_2| = 2$, τότε $\hat{\theta}_1 = \hat{\theta}_{2,3}$
- Αν $\Delta > 0 \iff |X_1 - X_2| > 2$, τότε δύο ρίζες:
  - $\hat{\theta}_2 = \hat{\theta}_1 - \frac{\sqrt{\Delta}}{2}$
  - $\hat{\theta}_3 = \hat{\theta}_1 + \frac{\sqrt{\Delta}}{2}$

Παραδείγματα Εκτιμήτριας Μεγίστης Πιθανοφάνειας (MLE)

- Παράδειγμα 1:

  Έστω ότι η τυχαία μεταβλητή X ακολουθεί την κατανομή Bernoulli $B(p)$ για την οποία ξέρουμε ότι:

$$P(X=1)=p, \quad P(X=0)=1-p, \qquad p\in[0,1].$$

  Η συνάρτηση πιθανοφάνειας δίνεται από τη σχέση:

$$\begin{aligned} L(p \mid x_1, x_2, ..., x_n) &= \prod_{i=1}^{n} f(x;p) \\ &= \prod_{i=1}^{n} p^{x_i}(1-p)^{1-x_i} \\ &= p^{\sum x_i}(1-p)^{n-\sum x_i}. \end{aligned}$$

Ο λογάριθμος της συνάρτησης πιθανοφάνειας είναι:

$$\begin{aligned}\mathcal{L}(p \mid x_1, x_2, ..., x_n) &= \log L(p \mid x_1, x_2, ..., x_n) \\ &= \sum x_i \log p + \left(n - \sum x_i\right) \log(1-p).\end{aligned}$$

Η συνάρτηση $\mathcal{L}(p)$ μεγιστοποιείται στα σημεία που μηδενίζεται η πρώτη παράγωγός της ως προς την παράμετρο p. Άρα:

$$\frac{\partial \mathcal{L}(p)}{\partial p} = \frac{\sum x_i}{p} - \frac{n - \sum x_i}{1-p}$$

$$\begin{aligned}\frac{\partial \mathcal{L}(p)}{\partial p} = 0 \iff (1-p)\sum x_i - p\left(n - \sum x_i\right) = 0 \Rightarrow \\ \hat{p} = \frac{\sum x_i}{n} = \bar{X}.\end{aligned}$$

Επιπλέον:

$$\left.\frac{\partial^2 \mathcal{L}(p)}{\partial p^2}\right|_{np=\sum x_i} = -\frac{\sum x_i}{p^2} - \frac{n-\sum x_i}{(1-p)^2} = -\frac{np}{p^2} - \frac{n-np}{(1-p)^2} = -\frac{n}{p} - \frac{n}{1-p} < 0,$$

διότι $0 \leq p \leq 1$.

Συνεπώς, η στατιστική συνάρτηση $\hat{p} = \bar{X}$ είναι ο εκτιμητής μεγίστης πιθανοφάνειας της παραμέτρου p.

- <u>Παράδειγμα 2:</u>

  Έστω τυχαίο δείγμα από πληθυσμό με κανονική κατανομή του οποίου η μέση τιμή είναι θ και η διακύμανση 1, $N(\theta, 1)$.

  Η συνάρτηση πυκνότητας πιθανότητας δίνεται από:

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

  Η συνάρτηση πιθανοφάνειας του δείγματος είναι:

$$L(\theta) = (2\pi)^{-n/2} \prod_{i=1}^{n} \exp\left\{-\frac{1}{2}(x_i-\theta)^2\right\}$$

  και λογαριθμίζοντας προκύπτει ότι:

$$\mathcal{L}(\theta) = -\frac{n}{2}\log(2\pi) - \frac{1}{2}\sum_{i=1}^{n}(x_i-\theta)^2.$$

Παραγωγίζοντας ως προς θ και θέτοντας την πρώτη παράγωγο ίση με το μηδέν, έχουμε ότι:

$$\frac{\partial \mathcal{L}(\theta)}{\partial \theta} = 0 \Longleftrightarrow \sum_{i=1}^{n} (x_i - \theta) = 0 \Rightarrow$$

$$\sum_{i=1}^{n} x_i - n\theta = 0 \Rightarrow$$

$$\hat{\theta} = \frac{1}{n} \sum_{i=1}^{n} x_i = \bar{X}.$$

Επιπλέον:

$$\left. \frac{\partial^2 \mathcal{L}(\theta)}{\partial \theta^2} \right|_{n\theta = \sum x_i} = -n < 0.$$

- <u>Παράδειγμα 3:</u>

$$f\left(x;\theta\right)=\theta x^{\theta-1},\quad 0<x<1,\quad \theta>0$$

$$L\left(\theta\right)=\theta^{n}\prod_{i=1}^{n}X_{i}^{\theta-1}=\theta^{n}X_{1}^{\theta-1}X_{2}^{\theta-1}...X_{n}^{\theta-1}\Rightarrow$$

$$\log L=n\log\theta+\left(\theta-1\right)\sum_{i=1}^{n}\log X_{i}.$$

$$\frac{\partial\log L}{\partial\theta}=\frac{n}{\theta}+\sum_{i=1}^{n}\log X_{i}.$$

$$\frac{\partial\log L}{\partial\theta}=0\Rightarrow$$

$$\begin{aligned}\hat{\theta}&=-\frac{n}{\sum_{i=1}^{n}\log X_{i}}\\&=-\left(\overline{\log X}\right)^{-1}.\end{aligned}$$

• Παράδειγμα 4:

Η ομοιόμορφη κατανομή $U(0,\theta)$ με συνάρτηση πυκνότητας πιθανότητας:

$$f(x;\theta) = \frac{1}{\theta}, \quad 0 \leq x \leq \theta$$

και συνάρτηση πιθανοφάνειας:

$$L(\theta) = \prod_{i=1}^{n} f(x_i;\theta) = \frac{1}{\theta^n},$$

εάν $\theta \geq \max(X_1, ..., X_n)$.

Έχουμε ότι:

$$\max_{\theta} L(\theta) = \frac{1}{\left(\max_{1\leq i\leq n} X_i\right)^n},$$

διότι

$$\theta \geq X_i \Rightarrow \hat{\theta} \geq \max_{1\leq i\leq n} X_i,$$

και άρα

$$\max_{\theta} L(\theta) = \frac{1}{\left(\min \hat{\theta}\right)^n} = \frac{1}{\left(\max_{1\leq i\leq n} X_i\right)^n}.$$

Επομένως, ο MLE είναι:

$$\hat{\theta} = \max_{1\leq i\leq n} X_i.$$

• Παράδειγμα 5:

Έστω τυχαίο δείγμα από πληθυσμό με κανονική κατανομή του οποίου η μέση τιμή είναι μ και είναι γνωστή και η άγνωστη παράμετρος είναι η διακύμανση την οποία συμβολίζουμε ως θ, δηλαδή έχουμε την κατανομή: $N(\mu, \theta)$.

Η συνάρτηση πυκνότητας πιθανότητας δίνεται από:

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

Η συνάρτηση πιθανοφάνειας του δείγματος είναι:

$$L(\theta) = (2\pi\theta)^{-n/2} \prod_{i=1}^{n} \exp\left\{-\frac{1}{2\theta}(x_i-\mu)^2\right\}$$

και λογαριθμίζοντας προκύπτει ότι:

$$\mathcal{L}(\theta) = -\frac{n}{2}\log\theta - \frac{n}{2}\log(2\pi) - \frac{1}{2\theta}\sum_{i=1}^{n}(x_i-\mu)^2.$$

Παραγωγίζοντας ως προς θ και θέτοντας την πρώτη παράγωγο ίση με το μηδέν, έχουμε ότι:

$$\frac{\partial \mathcal{L}(\theta)}{\partial \theta} = 0 \Longleftrightarrow -\frac{n}{2\theta} + \frac{1}{2\theta^2} \sum_{i=1}^{n} (x_i - \mu)^2 = 0 \Rightarrow$$

$$\frac{1}{\theta} \sum_{i=1}^{n} (x_i - \mu)^2 - n = 0 \Rightarrow$$

$$\hat{\theta} = \frac{1}{n} \sum_{i=1}^{n} (x_i - \mu)^2 .$$

Επιπλέον, η λύση αυτή είναι μοναδική, αντιστοιχεί σε μέγιστο επειδή:

$$\begin{aligned} \frac{\partial^2 \mathcal{L}(\theta)}{\partial \theta^2} &= \frac{n}{\theta^2} - \frac{1}{\theta^3} \sum_{i=1}^{n} (x_i - \mu)^2 \\ &= -\frac{n}{2\theta^2} < 0, \end{aligned}$$

το οποίο είναι ολικό μέγιστο αφού $\lim_{\theta \to 0} \mathcal{L}(\theta) = \lim_{\theta \to \infty} \mathcal{L}(\theta) = -\infty$.

- Παράδειγμα 6:

  Έστω ότι η τυχαία μεταβλητή X ακολουθεί την Γεωμετρική κατανομή με άγνωστη παράμετρο θ για την οποία ξέρουμε ότι:

$$P(x;\theta) = (1-\theta)^{x-1}\theta.$$

  Η συνάρτηση πιθανοφάνειας δίνεται από τη σχέση:

$$\begin{aligned} L(\theta \mid x_1, x_2, ..., x_n) &= \prod_{i=1}^{n} P(x_i;\theta) \\ &= P(x_1;\theta)\, P(x_2;\theta) \dots P(x_n;\theta) \\ &= (1-\theta)^{\left[\sum x_i - n\right]}\, \theta^n. \end{aligned}$$

Ο λογάριθμος της συνάρτησης πιθανοφάνειας είναι:

$$\begin{aligned}\mathcal{L}\left(\theta \mid x_1, x_2, ..., x_n\right) &= \log L\left(\theta \mid x_1, x_2, ..., x_n\right) \\ &= \left(\sum x_i - n\right)\log\left(1-\theta\right) + n\log\left(\theta\right).\end{aligned}$$

Η συνάρτηση $\mathcal{L}\left(\theta\right)$ μεγιστοποιείται στα σημεία που μηδενίζεται η πρώτη παράγωγός της ως προς την παράμετρο θ. Άρα:

$$\frac{\partial\mathcal{L}\left(\theta\right)}{\partial\theta} = \left(\sum x_i - n\right)\frac{-1}{1-\theta} + \frac{n}{\theta}$$

$$\frac{\partial\mathcal{L}\left(\theta\right)}{\partial\theta} = 0 \Longleftrightarrow \left(\sum x_i - n\right)\frac{-1}{1-\theta} + \frac{n}{\theta} = 0 \Rightarrow$$

$$\hat{\theta} = \frac{n}{\sum x_i}.$$

(η ίδια εκτιμήτρια με τη μέθοδο των ροπών)

- Παράδειγμα 7:

  Έστω $(X_1, X_2, ..., X_n)$ τυχαίο δείγμα από κατανομή με συνάρτηση πυκνότητας πιθανότητας:

$$f(x;\lambda) = \frac{\lambda^3}{2} x^2 e^{-\lambda x}, \quad x > 0.$$

  Ο λογάριθμος της συνάρτησης πιθανοφάνειας του δείγματος δίνεται από τον τύπο:

$$\begin{aligned}
\mathcal{L}(\lambda) := \log L(\lambda) = \log \prod_{i=1}^{n} f(x_i;\lambda) \\
&= \sum_{i=1}^{n} \log f(x_i;\lambda) \\
&= \sum_{i=1}^{n} \log \left(\frac{\lambda^3}{2} x_i^2 e^{-\lambda x_i} \right) \\
&= \sum_{i=1}^{n} \left(3\log\lambda - \log 2 + \log x_i^2 - \lambda x_i \right) \\
&= 3n\log\lambda - n\log 2 + \sum_{i=1}^{n} \log x_i^2 - \lambda \sum_{i=1}^{n} x_i.
\end{aligned}$$

Παρατηρούμε ότι η συνάρτηση της παραμέτρου λ είναι παραγωγίσιμη σε όλον τον παραμετρικό χώρο $(0,\infty)$. Ειδικότερα έχουμε ότι:

$$\frac{\partial \mathcal{L}(\lambda)}{\partial \lambda} = \frac{3n}{\lambda} - \sum_{i=1}^{n} x_i$$

και θέτοντας την παράγωγο ίση με το μηδέν:

$$\frac{3n}{\lambda} - \sum_{i=1}^{n} x_i = 0 \Rightarrow$$

$$\hat{\lambda} = \frac{3}{\bar{X}}.$$

Απομένει να επαληθεύσουμε ότι η παραπάνω τιμή αποτελεί μέγιστο. Πράγματι, ισχύει ότι:

$$\frac{\partial^2 \mathcal{L}(\lambda)}{\partial \lambda^2} = -\frac{3n}{\lambda^2} < 0.$$

Βασικές Ιδιότητες

- Έστω $(X_1, X_2, ..., X_n)$ τυχαίο δείγμα από τυχαία μεταβλητή X με κατανομή $f(x;\theta)$ και $\hat{\theta}$ η εκτιμήτρια μεγίστης πιθανοφάνειας (MLE) της παραμέτρου θ.

- [Ιδιότητα 1.] Αν $g(\theta)$ είναι αμφιμονοσήμαντη συνάρτηση (δηλαδή 1-1) της παραμέτρου θ, τότε ο MLE της συνάρτησης αυτής είναι ο $g\left(\hat{\theta}\right)$.

  - Παράδειγμα 1:

    Έστω $X_1, X_2, ..., X_n \sim P(\lambda)$. Ο MLE είναι:

    $$\hat{\lambda} = \bar{X}.$$

    Τότε για $P(X = 0) = e^{-\lambda}$, ο MLE σε αυτή την περίπτωση είναι $e^{-\bar{X}}$.

  - Παράδειγμα 2:

    Στο Παράδειγμα 5 προηγουμένως, έχοντας βρει την εκτιμήτρια μεγίστης πιθανοφάνειας (MLE) της άγνωστης διακύμανσης της Κανονικής κατανομής, $\hat{\theta}$, μπορούμε μέσω αυτής της ιδιότητας να βρούμε την εκτιμήτρια MLE της τυπικής απόκλισης παίρνοντας κατευθείαν την τετραγωνική ρίζα του $\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2$.

- [Ιδιότητα 2.] Επιπλέον, ισχύει η ιδιότητα ότι οι εκτιμήτριες μεγίστης πιθανοφάνειας είναι αναλλοίωτες (invariant) ως προς αντιστρέψιμους μετασχηματισμούς των δεδομένων:

    - εάν X έχει συνάρτηση κατανομής $f_X(x;\theta)$ και $Y = h(X)$,τότε:

$$f_Y(y;\theta) = f_X\left(h^{-1}(y);\theta\right).$$

- [Ιδιότητα 3.] Επίσης, εάν T είναι επαρκής στατιστική συνάρτηση για την παράμετρο θ, τότε ο MLE $\hat{\theta}$ εξαρτάται από το τυχαίο δείγμα μόνο ως συνάρτηση του T:
  - από το Θεώρημα Παραγοντοποίησης, ισχύει ότι η στατιστική T είναι επαρκής έαν και μόνο έαν:

$$f\left(x;\theta\right) = g\left(T;\theta\right)h\left(x\right) \Rightarrow$$
$$\mathcal{L}\left(\theta\right) = \log g\left(T;\theta\right) + \log h\left(x\right)$$

  - ο MLE $\hat{\theta}$ μεγιστοποιεί τη λογαριθμική συνάρτηση πιθανοφάνειας $\mathcal{L}\left(\theta\right)$, με άλλα λόγια (λόγω της παραπάνω σχέσης) μεγιστοποιεί το $\log g\left(T;\theta\right)$
  - από το τελευταίο προκύπτει ότι ο MLE $\hat{\theta}$ είναι συνάρτηση του επαρκούς στατιστικού T.

- <u>Παράδειγμα (της Ιδιότητας 3):</u>

  Έστω ότι $(X_1, X_2, ..., X_n)$τυχαίο δείγμα από πληθυσμό που ακολουθεί την Εκθετική κατανομή με άγνωστη παράμετρο λ.

  Τότε:

$$f\left(x;\lambda\right) = \prod_{i=1}^{n} \lambda e^{-\lambda x_i} = \lambda^n e^{-\lambda \sum x_i}$$
$$= \lambda^n e^{-\lambda n \bar{X}}.$$

  Από το Θεώρημα Παραγοντοποίησης, ο δειγματικός μέσος $\bar{X}$ είναι επαρκής στατιστική για το λ.

  Για τον MLE $\hat{\lambda}$:

$$\frac{\partial}{\partial \lambda}\left(n \log \lambda - n\lambda\bar{X}\right) = \frac{n}{\lambda} - n\bar{X}$$

  και

$$\frac{n}{\lambda} - n\bar{X} = 0 \Rightarrow \hat{\lambda} = \frac{1}{\bar{X}}.$$

  Οπότε, ο MLE $\hat{\lambda}$ είναι συνάρτηση του επαρκούς στατιστικού $\bar{X}$.

  (<u>Να σημειωθεί επίσης ότι</u> συμπίπτει με την εκτιμήτρια ροπών του λ.)

Ασυμπτωτικές Ιδιότητες

- Κάτω από ορισμένες προϋποθέσεις (κατάλληλες συνθήκες ομαλότητας) ισχύει ότι:

$$\hat{\theta} \xrightarrow{p} \theta, \quad n \to \infty$$

και

$$\sqrt{nI(\theta)}\left(\hat{\theta} - \theta\right) \xrightarrow{D} N(0,1), \quad n \to \infty$$

δηλαδή ο MLE είναι συνεπής και ασυμπτωτικά πλήρως αποτελεσματικός, όπου:

$$I(\theta) = -E\left[\frac{\partial^2 \log f(x;\theta)}{\partial \theta^2}\right].$$

- Η ασυμπτωτική διακύμανση του $\hat{\theta}$ δίνεται από:

$$AsVar\left(\hat{\theta}\right) = \frac{1}{nI(\theta)}.$$

- Η πρακτική χρησιμότητα του παραπάνω αποτελέσματος της Ασυμπτωτικής Κανονικότητας είναι ότι για μεγάλα δείγματα n, ισχύει ότι:

$$\hat{\theta} \xrightarrow{D} N\left(\theta, \frac{1}{nI\left(\theta\right)}\right)$$

και άρα ότι:

$$Var\left(\hat{\theta}\right) \to 0, \quad n \to \infty.$$

- Αποδεικνύεται λοιπόν ότι ασυμπτωτικά η εκτιμήτρια $\hat{\theta}$ είναι άριστη εκτιμήτρια του θ.

- Επομένως, η βέλτιστη εκτιμήτρια μιας παραμέτρου θ όταν έχουμε ένα σχετικά μεγάλο δείγμα θα είναι η MLE $\hat{\theta}$.

Παραδείγματα Συνέπειας

- Για τυχαία μεταβλητή X από πληθυσμό που ακολουθεί την κατανομή Bernoulli(p), η εκτιμήτρια μεγίστης πιθανοφάνειας είναι:

$$\hat{p} = \bar{X}$$

  - εφαρμόζοντας το Νόμο των Μεγάλων Αριθμών (LLN), $\hat{p}$ είναι συνεπής εκτιμήτρια της παραμέτρου p.

- Για τυχαία μεταβλητή X από πληθυσμό που ακολουθεί την Ομοιόμορφη κατανομή $U(0,\theta)$, η εκτιμήτρια μεγίστης πιθανοφάνειας είναι:

$$\hat{\theta} = \max X_i \leq \theta$$

  - η κατανομή του $\hat{\theta}$ δίνεται από:

$$P\left(\hat{\theta} \leq x\right) = \theta^{-n} x^{n}, \quad 0 \leq x \leq \theta$$

  και ισχύει ότι:

$$\begin{aligned} P\left(\left|\theta - \hat{\theta}\right| \geq \varepsilon\right) &= P\left(\hat{\theta} \leq \theta - \varepsilon\right) \\ &= \left(1 - \theta^{-1}\varepsilon\right)^{n} \to 0, \quad n \to \infty \end{aligned}$$

Παράδειγμα Ασυμπτωτικής Κανονικότητας

- Έστω $(X_1, X_2, ..., X_n)$ τυχαίο δείγμα από πληθυσμό με κατανομή Bernoulli(p), με συνάρτηση πιθανότητας:

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

Η λογαριθμική συνάρτηση πιθανοφάνειας δίνεται από:

$$\mathcal{L}(p) = x \log p + (1-x) \log (1-p)$$

και η πρώτη και δεύτερη παράγωγος δίνονται ως εξής:

$$\frac{\partial \mathcal{L}(p)}{\partial p} = \frac{x}{p} - \frac{1-x}{1-p},$$
$$\frac{\partial^2 \mathcal{L}(p)}{\partial p^2} = -\frac{x}{p^2} - \frac{1-x}{(1-p)^2}.$$

Η πληροφορία κατά Fisher (Fisher Information) είναι ίση με:

$$I(p) = -E\left(\frac{\partial^2 \mathcal{L}(p)}{\partial p^2}\right) = \frac{p}{p^2} + \frac{1-p}{(1-p)^2}$$
$$= \frac{1}{p(1-p)}.$$

Οπότε έχουμε ότι:

$$\sqrt{n}(\hat{p} - p) \sim N(0, p(1-p)),$$

το οποίο συμβαδίζει απόλυτα με το αποτέλεσμα που αποκτάται μέσω της εφαρμογής του Κεντρικού Οριακού Θεωρήματος (CLT).