# Τρόποι Σύγκλισης Ακολουθιών τ.μ

Έστω $(X_n)_{n\geq 1}$ μία ακολουθία τ.μ.
Θα λέμε ότι η $(X_n)$ συγκλίνει με
πιθανότητα 1 ή σχεδόν βεβαίως ή
ισχυρά σε μία τ.μ. $X$ και θα γράφουμε
(almost surely)

$$\boxed{X_n \xrightarrow{a.s.} X \quad \text{αν} \quad P\left(\lim X_n = X\right) = 1,}$$

όπου $\{\lim X_n = X\} = \{\omega \in \Omega : \exists \lim\limits_{n\to\infty} X_n(\omega) \text{ και } \lim X_n(\omega) = X(\omega)\}$

Πρακτικά: αν πάρουμε
μία ακολουθία $(X_n(\omega))$ τότε $X_n(\omega) \longrightarrow X(\omega)$.

- Θα λέμε ότι μία ακολουθία τ.μ. $X_n$ συγκλίνει **κατά πιθανότητα** ή **στοχαστικά** σε μία τ.μ. $X$, και θα γράφουμε

$$X_n \xrightarrow{P} X \quad \text{(in probability)},$$

αν $\forall \varepsilon > 0$, $P(|X_n - X| > \varepsilon) \xrightarrow[n\to\infty]{} 0$

**Παρατήρηση:**

$$\boxed{\begin{array}{c} X_n \xrightarrow{a.s.} X \implies X_n \xrightarrow{P} X \\ \xrightarrow{a.s} \implies \xrightarrow{P} \end{array}}$$

Απο το παραπάνω έπεται ότι η σύγκλιση με πιθανότητα 1 (σχεδόν βέβαια) είναι ισχυρότερη της κατά πιθανότητας σύγκλισης.

## Παράδειγμα



Έστω $(X_n)$ μία ακολ. τ.μ. με

$X_n \sim Be\left(\frac{1}{n}\right)$, δηλ. $P(X_n = 1) = \frac{1}{n}$.

Τότε $X_n \xrightarrow{P} 0$.

Πράγματι, έστω $\varepsilon > 0$ $(0 < \varepsilon < 1)$

$$P(|X_n - 0| > \varepsilon) = P(X_n > \varepsilon) = P(X_n = 1) = \frac{1}{n} \to 0.$$

- Αν πάρουμε την $(X_n)$ να είναι ανεξάρτητη ακολ. τ.μ. (οι $X_n$, $n \geq 1$ είναι ανεξάρτητες), τότε μπορεί να δειχθεί ότι

$X_n \overset{a.s.}{\not\to} 0$ $\left[ \xrightarrow{a.s.} \text{ είναι γνήσια ισχυρότερη από την } \xrightarrow{P} \right]$.

[όταν είναι ανεξάρτητες: $\forall \varepsilon > 0$.

$$\sum_{n=1}^{+\infty} P(X_n > \varepsilon) = \sum_{n=1}^{+\infty} \frac{1}{n} = +\infty,$$

τότε $X_n \overset{a.s.}{\not\to} 0$ ]

• Αν $(X_n)$ είναι ακολ. τ.μ., τότε λέμε ότι η $(X_n)$ συγκλίνει **κατά κατανομή** στη $X$ και γράφουμε $\boxed{X_n \xrightarrow{d} X}$ (in distribution), αν:

$$\boxed{\lim_{n\to\infty} F_n(x) = F(x), \ \forall x : \text{η } F \text{ να είναι συνεχής στο } x.}$$

όπου $F_n$ η σ.κ. της $X_n$ και $F$ η σ.κ. της $X$.

$$\boxed{X_n \xrightarrow{a.s.} X \underset{\not\Leftarrow}{\Rightarrow} X_n \xrightarrow{P} X \underset{\not\Leftarrow}{\Rightarrow} X_n \xrightarrow{d} X \qquad \xrightarrow{a.s} \Rightarrow \xrightarrow{P} \Rightarrow \xrightarrow{d}}$$

Μια ειδική περίπτωση που αν $X = c$ με πιθανότητα 1 (εκφυλισμένη τ.μ.) τότε

$$\boxed{X_n \xrightarrow{P} c \Leftrightarrow X_n \xrightarrow{d} c}$$

6

- Έστω $(X_n)$ μια ακολ. τ.μ. και $X$ άλλη μια τ.μ. Λέμε ότι $(E|X_n|, E|X| < +\infty)$

$X_n \xrightarrow{\mathcal{L}^1} X$ αν $E|X_n - X| \xrightarrow[n\to\infty]{} 0$

(συγκλίνει κατά μέσο) $(EX_n^2, EX^2 < +\infty)$

$X_n \xrightarrow{\mathcal{L}^2} X$ αν $E(X_n - X)^2 \longrightarrow 0$

(συγκλίνει κατά μέσο τετράγωνο).

<u>Σύνδεση:</u>

$X_n \xrightarrow{\mathcal{L}^2} X \Rightarrow X_n \xrightarrow{\mathcal{L}^1} X$, $X_n \xrightarrow{\mathcal{L}^1, \mathcal{L}^2} X \Rightarrow X_n \xrightarrow{P} X$

$$\boxed{\left.\begin{array}{l} \xrightarrow{a.s.} \\ \xrightarrow{L^2} \Rightarrow \xrightarrow{L^1} \end{array}\right\} \Rightarrow \xrightarrow{P} \Rightarrow \xrightarrow{d}}$$

## Σημαντικά Οριακά Θεωρήματα

- **Ισχυρός Νόμος των Μεγάλων Αριθμών.**

Έστω $(X_n)$ μια ακολ. ανεξ + ισον. τ.μ. με $E|X_1|<+\infty$, και θέτουμε $\mu := E(X_1)$.
Τότε αν $\overline{X}_n = \sum_{i=1}^{n} X_i/n$ ο δειγματικός μέσος, τότε

$$\boxed{\overline{X}_n \xrightarrow{a.s.} \mu}$$

# Ασθενής Νόμος των Μεγάλων Αριθμών

Έστω $(X_n)$ μια ακολ. ανεξ. + ισον. τ.μ. με $E|X_1| < +\infty$, αν θέσουμε $\mu \equiv E(X_1)$

τότε

$$\boxed{X_n \xrightarrow{P} \mu}$$

(έπεται άμεσα, από τον ισχυρό νόμο, αφού $\xrightarrow{a.s.} \Rightarrow \xrightarrow{P}$)

# Κεντρικό Οριακό Θεώρημα

Έστω $(X_n)$ μια ακολουθία ανεξ. + ισόνομων τμ. με $0 < V(X_1) < +\infty$. Τότε αν

$$\sigma^2 \equiv V(X_1) \quad \text{και} \quad \mu \equiv E(X_1)$$

αν $S_n = \sum_{i=1}^{n} X_i$ και $\overline{X_n} = \dfrac{\sum_{i=1}^{n} X_i}{n}$ )

(9)

τότε

$$\frac{S_n - E(S_n)}{\sqrt{V(S_n)}} \xrightarrow{d} Z \sim N(0,1)$$

ή

$$\frac{S_n - n\mu}{\sqrt{n\sigma^2}} \xrightarrow{d} Z_1 \sim N(0,1)$$

ή

$$\frac{S_n - n\mu}{\sqrt{n}} \xrightarrow{d} N(0,\sigma^2)$$

ń

$$\frac{\overline{X}_n - E(\overline{X}_n)}{\sqrt{V(\overline{X}_n)}} \xrightarrow{d} Z \sim N(0,1) \quad ń$$

$$\sqrt{n}\,\frac{\overline{X}_n - \mu}{\sigma} \xrightarrow{d} Z \sim N(0,1)$$

$$\left(V(\overline{X}_n) = \frac{\sigma^2}{n}\right).$$

ń

$$\sqrt{n}\,(\overline{X}_n - \mu) \xrightarrow{d} N(0, \sigma^2)$$

Πρακτικά, αν $n$ είναι επαρκώς μεγάλο, τότε (αποδεικνύεται προσεγγιστικά) (11)

$$\boxed{\begin{array}{l} S_n \approx N(n\mu, n\sigma^2) \\ \overline{X}_n \approx N\left(\mu, \frac{\sigma^2}{n}\right) \end{array}}$$

$S_n = \sum_{i=1}^{n} X_i$

Παράδειγμα.

αν $S_n \sim G(n, \theta)$, τότε

$$S_n = \sum_{i=1}^{n} X_i \quad , \quad X_i \sim \mathcal{E}xp(\theta) \, .$$

για μεγάλο $n$, $G(n, \theta) \cong N\left(\frac{n}{\theta}, \frac{n}{\theta^2}\right)$

# Στοιχεία Στατιστικής

Στις Πιθανότητες θεωρούμε ότι γνωρίζουμε τις κατανομές τυχαίων μεταβλητών και βάσει αυτών προσπαθούμε να υπολογίσουμε διάφορες πιθανότητες ενδεχομένων που μας ενδιαφέρουν.

Στη Στατιστική έχουμε παρατηρήσεις, δεδομένα από κάποια χαρακτηριστικά που συνδέονται με κάποια εφαρμογή ενδιαφέροντος και προσπαθούμε να εκτιμήσουμε τις κατανομές των χαρακτηρισικών.

— Ποιά κατανομή/κατανομές βρίσκονται πίσω απο την παραγωγή των δεδομένων?

— Τί συμπεράσμοιτα μπορούμε να εξάγουμε για τον πληθυσμό απο τον οποίο έχουμε πάρει κάποιο δείγμα?

— Θα μπορούσαμε να προβλέψουμε μελλοντικές παρατηρήσεις με μικρό σφάλμα?

Υπάρχουν διάφορες υποθέσεις που κάνουμε για τη φύση της άγνωστης κατανομής.

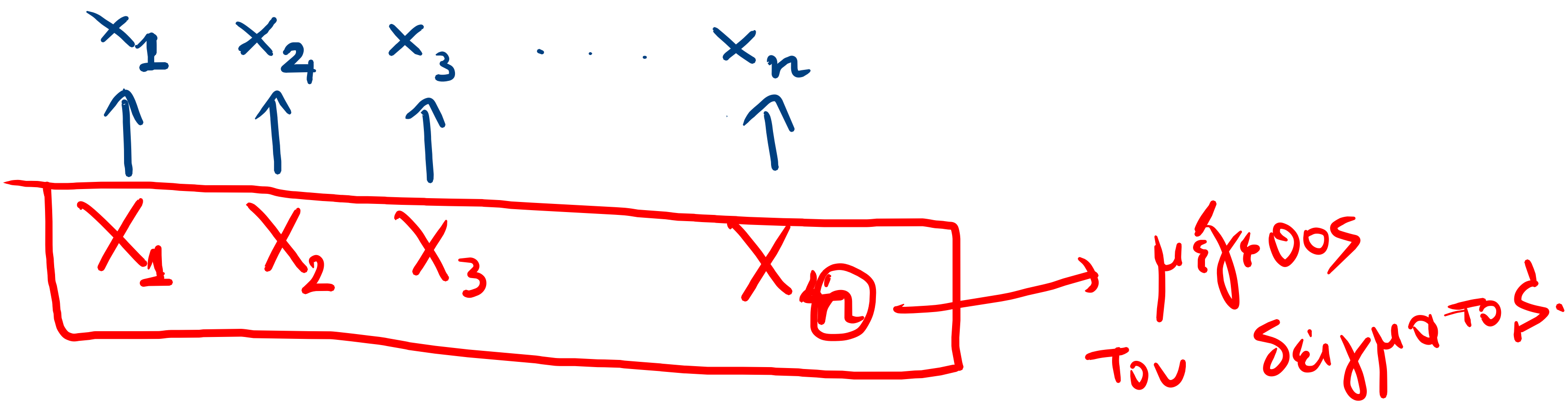
Παραμετρική Στατιστική → παραμετρικές υποθέσεις (πεπερασμ. διάστασης)

Μη Παραμετρική Στατιστική → απειροδιάστατους χώρους (συναρτησιακούς)

θα ασχοληθούμε πιο πολύ

Εκτιμητική / Διαστήματα Εμπιστοσύνης / Έλεγχοι Υποθέσεων

# Εκτιμητική




Τα δεδομένα $x_1, x_2, \dots, x_n$ είναι πραγματοποιήσεις τυχαίων μεταβλητών.

<u>απλούστερη περίπτωση</u>: ανεξάρτητες + ισόνομες τ.μ.

$\equiv$ τυχαίο δείγμα.

# Παραμετρική Στατιστική



~~$X$~~$_i$ προέρχονται από κάποια παραμετρική οικογένεια κατανομών.

π.χ. $\left\{ f_\theta \,:\, \theta \in \Theta \right\}$

→ υποψήφιες συναρτήσεις πιθανότητας (αν είναι διακριτή η $X_i$)

ή συναρτ. πυκνότητας πιθανότητας (αν είναι συνεχής η $X_i$).

$\Theta \subseteq \mathbb{R}^d$ → πεπερ. διάστασης χώρος.

$\theta = (\theta_1, \theta_2, \dots, \theta_d) \to$ παράμετρος.

θα περιγράψει πλήρως την άγνωστη κατανομή.

π.χ. $X_i \sim N(\underbrace{\mu, \sigma^2}_{\text{άγνωστα}})$. $\theta = (\mu, \sigma^2)$

μέση τιμή, διασπορά.

$\theta \in \Theta \to$ παραμετρικός χώρος (φιλοξενεί τις δυνατές τιμές του θ).

π.χ. $\theta = (\mu, \sigma^2) \in \mathbb{R} \times (0, +\infty)$

# Μέθοδοι Εκτίμησης της παραμέτρου θ.

Για να εκτιμήσουμε το θ, χρησιμοποιούμε κατάλληλες τυχαίες μεταβλητές, που πρέπει να μας δίνουν τιμές κοντά στο άγνωστο θ.

**Ορσ**: Κάθε συνάρτηση $T = T(X)$ του τυχαίου δείγματος $X = (X_1, X_2, \dots, X_n)$ λέγεται στατιστική συνάρτηση (σ.σ.) και δεν θα πρέπει να περιέχει άγνωστες παραμέτρους.

π.χ. η $T = X_1 + X_2 + \dots + X_n$

είναι σ.σ., όπως επίσης και τα παρακάτω

$\overline{X_n}$, $X_1^2 + X_2^2 + \dots + X_n^2$, $\max\{X_1, \dots, X_n\} \equiv X_{(n)}$,
$\min\{X_1, \dots, X_n\} \equiv X_{(1)}$, $X_{(n)} - X_{(1)}$

όπως σε ένα τ.δ. από $N(\mu, \sigma^2)$

άγνωστο

Το $\overline{X} - \mu$ δεν είναι σ.σ.

άγνωστη παράμετρος.

**Ορσ**: Μια σ.σ. $T = T(X)$ λέγεται εκτιμήτρια του $\theta$, αν $T(X) \in \Theta$.

Όμοια αν μας ενδιαφέρει κάποια $g(\theta)$ ορίζουμε την εκτιμήτρια του $g(\theta)$.

π.χ. $\theta = (\mu, \sigma^2)$, $g(\theta) = \dfrac{\mu}{\sigma^2}$

ή $g(\theta) = \mu^2$

Σημαντικές σ.σ. που συνδέονται με ένα Τ.Δ.

21

$$\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

$$\overline{X^2} = \frac{\sum_{i=1}^{n} X_i^2}{n}.$$

$$M_2 = \frac{\sum_{i=1}^{n} (X_i - \overline{X})^2}{n} \rightarrow \text{δειγματική διασπορά}$$

$$S^2 = \frac{\sum_{i=1}^{n} (X_i - \overline{X})^2}{n-1} \rightarrow \text{αμερόληπτη δειγματική διασπορά}$$

$X_{(n)}, X_{(1)}, X_{(n)} - X_{(1)}$ : δειγματικό μέγιστο/ελάχιστο/εύρος δειγματική διάμεσος : $\begin{cases} X_{(k)+1}, & n=2k+1 \\ \dfrac{X_{(k)}+X_{(k)+1}}{2}, & n=2k \end{cases}$

π.χ. $X_1, X_2, X_3 \longrightarrow$ δ.δ : $X_{(2)}$

$X_1, X_2, X_3, X_4 \longrightarrow$ δ.δ : $\dfrac{X_{(2)}+X_{(3)}}{2}$.

# Μέθοδος Ροπών (απλός τρόπος εξαγωγής εκτιμητριών).

Παρ. 1.

Έστω $X_1, \ldots, X_n$ τ.δ. $P(\lambda)$, $\lambda > 0$.

Εξισώνουμε τη θεωρητική με τη δειγματική ροπή 1ης τάξης:

$$E_\lambda(X_1) = \overline{X} \iff \lambda = \overline{X}$$

Άρα παίρνουμε ως εκτιμήτρια ροπών τη

$$\boxed{\overline{\lambda} = \overline{X}}$$

## Παρ. 2.

Έστω $X_1, \dots, X_n$ τ.δ. $N(\mu, \sigma^2)$ και τα δύο άγνωστα.

$\theta = (\mu, \sigma^2)$.

$$\left.\begin{aligned} E_\theta(X_1) &= \overline{X_n} \\ E_\theta(X_1^2) &= \overline{X_n^2} \end{aligned}\right\} \Longleftrightarrow \left.\begin{aligned} \mu &= \overline{X_n} \\ \sigma^2 + \mu^2 &= \overline{X_n^2} \end{aligned}\right| \Longleftrightarrow$$

$$\mu = \overline{X_n}$$

$$\sigma^2 = \overline{X_n^2} - \left(\overline{X_n}\right)^2 = M_2 = \frac{\sum_{i=1}^{n} (X_i - \overline{X_n})^2}{n}$$

→ κεντρική δειγμ. ροπή 2ης τάξης ή δειγμ. διασπορά!

Τελικά πήραμε ως λύση

$$\overline{\theta} = \left(\overline{\mu}, \overline{\sigma^2}\right) = \left(\overline{X}_n, M_2\right)$$

- Υπάρχει πιο γρήγορος τρόπος.

$$\left.\begin{aligned} E_\theta(X_1) &= \overline{X}_n \\ V_\theta(X_1) &= M_2 \end{aligned}\right| \Longleftrightarrow \boxed{\begin{aligned} \mu &= \overline{X}_n \\ \sigma^2 &= M_2 \end{aligned}}$$

## Γενική μέθοδος

αν $\theta = (\theta_1, \ldots, \theta_d)$

$E_\theta(X_1) = \overline{X}$

$E_\theta(X_1^2) = \overline{X^2}$

$\vdots$

$E_\theta(X_1^s) = \overline{X^s}$

ή ισοδ.

$$\begin{cases} E_\theta(X_1) = \overline{X} \\ V_\theta(X_1) = M_2 \\ \vdots \\ E_\theta(X_1 - \mu)^{s} = \dfrac{\sum_{i=1}^{n}(X_i - \overline{X}_n)^{s}}{n} \end{cases}$$

[δε θα χρειαστούμε $s \geq 3$]

- Υπάρχει περίπτωση να χρειαστεί να συμπληρώσουμε με $> s$ (π.χ. $U[-\theta,\theta] \to 0 = \overline{X}$)

- Υπάρχουν κάποια μειονεκτήματα.

π.χ. δεν μπορούμε να είμαστε σίγουροι ότι οι λύσεις που παίρνουμε είναι εκτιμήτριες (αν υπάρχουν λύσεις), οπότε χρειάζεται διόρθωση παίρνοντας υπόψη μας περιορισμούς παραμέτρων.

- Η πιο δημοφιλής μέθοδος εξαγωγής εκτιμητριών είναι η μέθοδος μέγιστης πιθανοφάνειας.

→ Έστω $x = (x_1, \ldots, x_n)$.

Λέμε συνάρτηση πιθανοφάνειας που αντιστοιχεί στο $x$, τη συνάρτηση

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

$x$ → είναι γνωστό.

όπου $f(x_i; \theta)$ θα είναι η συνάρτηση πιθανότητας (για διακριτές) ή πυκνότητας πιθανότητας (για συνεχείς).

των $X_i$, αν το $\theta$ είναι η συγκεκριμένη τιμή κάτω από την οποία υπολογίζονται (ως συνάρτηση του $\theta$).

π.χ.

Έστω $X_1, \ldots, X_n$ τ.δ. $\sim Be(p)$, $0 < p < 1$.

και $x = (x_1, \ldots, x_n)$ τα δεδομένα μας.

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

$$= p^{\sum_{i=1}^{n} x_i}(1-p)^{n-\sum_{i=1}^{n} x_i}$$

Η εκτίμηση μέγιστης πιθανοφάνειας για την τιμή του $p$ προκύπτει ως η τιμή εκείνη του $p$, που μεγιστοποιεί την $L(p)$.

$\underset{\text{όρισμα}}{\arg}\max L(p)$ ?, δηλ. το $\hat{p}$ : $L(\hat{p}) = \max\limits_{p\in(0,1)} L(p)$

Εδώ $L(p) = p^{\sum_{i=1}^{n} x_i} (1-p)^{n-\sum_{i=1}^{n} x_i}$

Για να μεγιστοποιήσουμε αυτή τη συνάρτηση είναι πιο εύκολο να δουλεύουμε με log και

$$l(p) \equiv \log L(p) = \left(\sum_{i=1}^{n} x_i\right)\cdot \log p + \left(n - \sum_{i=1}^{n} x_i\right)\log(1-p).$$

$$\arg\max_{p\in(0,1)} L(p) = \arg\max_{p\in(0,1)} \ell(p),$$

όπου $\ell(p) \equiv \log L(p)$ (λογαριθμοπιθανοφάνεια).

δηλ. η log ↗

• (i) αν $\sum x_i = 0 \iff x_1 = \dots = x_n = 0$

$$\ell(p) = n\log(1-p) \searrow$$ (Τυπικά ∄ η ε.μ.π.)

• (ii) αν $\sum x_i = n \iff x_1 = \dots = x_n = 1$

$$\ell(p) = n\log p \nearrow$$ (Τυπικά ∄ η ε.μ.π.)

• (iii) αν $0 < \sum_{i=1}^{n} x_i < n$

Βρίσκουμε $\ell'(p)$ και λύνουμε

$\ell'(p)=0$ (εξίσωση πιθανοφάνειας)

$$\ell'(p) = \frac{\sum_{i=1}^{n} x_i}{p} - \frac{n-\sum_{i=1}^{n} x_i}{1-p} = \frac{S_n}{p} - \frac{n-S_n}{1-p}$$

$$\ell'(p)=0 \iff p^* = \frac{S_n}{n} = \overline{X}_n$$

(μοναδικό στάσιμο σημείο)

$$\ell''(p) = -\frac{S_n}{p^2} - \frac{n-S_n}{(1-p)^2} < 0$$

(γνήσια κοίλη)

Ειδικά $\ell''(p^*) < 0 \implies p^*$ τοπικό μέγιστο

Επειδή όμως είναι και μοναδικό στάσιμο σημείο, είναι ολικό μέγιστο

$\Rightarrow \quad \hat{p} = \bar{X}_n$. Η ε.μ.π. $\boxed{\hat{p}(X) = \bar{X}_n}$

Παρατήρηση

Αρχικά αν $p=0$, $Be(0) \equiv 0$

$p=1$, $Be(1) \equiv 1$

$p \in [0,1]$

$$L(p) = p^{\sum_{i=1}^{n} x_i}(1-p)^{n-\sum_{i=1}^{n} x_i}$$

Επέκταση.

$0^0 = 1 \Rightarrow \log 0^0 = 0 \Rightarrow$

$0 \cdot \underbrace{\log 0}_{(-\infty)} = 0$

$\hat{p} = 0$, όταν $x_1 = x_2 = \cdots = x_n = 0$.

# Συμπαγοποίηση ενός π.χ.

$p \in (0,1) \longrightarrow \underbrace{[0,1]}$

κλειστό + φραγμένα ≡ συμπαγές

Εδώ $\boxed{\hat{p} = \overline{X}_n}$ (για όλες τις περιπτώσεις)

↳ γενικά στα διακριτά εμφανίζονται τέτοια φαινόμενα

• <u>Ορσ</u>

Λέμε <u>εκτιμήτρια μέγιστης πιθανοφάνειας</u> $\hat{\theta} = \hat{\theta}(X)$, την σ.σ. που προκύπτει μέσα από τη μεγιστοποίηση της πιθανοφάνειας $L(\theta)$, και άρα πρέπει να ικανοποιεί

$$\hat{\theta} = \arg\max L(\theta)$$

ή να έχουμε ότι

$$L(\hat{\theta}) = \max_{\theta \in \Theta} L(\theta).$$

Η ε.μ.π. δεν υπάρχει πάντα, όταν υπάρχει μπορεί να μην είναι μοναδική, μπορεί να υπάρχουν και άπειρες και γενικά χρειάζεται μια διερεύνηση. Σε απλά παραδείγματα μπορούμε να πάρουμε μοναδικές λύσεις και η μέθοδος αυτή επεκτείνεται και σε πολυδιάστατη παράμετρο με τον ίδιο ορισμό.

**Παράδ.**

Έστω $X_1, \ldots, X_n$ τ.δ. $\underline{\mathcal{E}xp(\vartheta)}$, $\vartheta > 0$.
Να βρεθεί η ε.μ.π.

**Λύση.**

Έστω $x = (x_1, \ldots, x_n)$, με $x_i > 0$, $1 \le i \le n$.

$$L(\theta) = \prod_{i=1}^{n} f(x_i;\vartheta) = \prod_{i=1}^{n} \vartheta e^{-\vartheta x_i} = \theta^n e^{-\theta \sum_{i=1}^{n} x_i}$$

$\Longrightarrow$

$$\ell(\theta) = \log L(\theta) = n \log \vartheta - \vartheta \cdot \sum_{i=1}^{n} x_i, \quad \theta > 0.$$

$$\ell'(\theta) = \frac{n}{\theta} - \sum_{i=1}^{n} x_i, \quad \ell'(\vartheta) = 0 \Leftrightarrow \theta^* = \frac{1}{\overline{x}_n}$$

μοναδικό στάσιμο σημείο.

$$\ell''(\theta) = -\frac{n}{\theta^2} < 0 \quad \forall \theta > 0.$$

(γνήσια κοίλη η $\ell(\theta)$)

(Δ) $\boxed{\hat{\theta} = \frac{1}{\overline{X}_n}}$ είναι η ε.μ.π. του $\theta$.