

Μια ακολουθία τυχαίων μεταβλητών

$$X_1, X_2, \dots, X_n$$

αποτελεί τυχαίο δείγμα αν είναι ανεξάρτητες

και ισόνομες, δηλ. ακολουθούν την ίδια κατανομή.

Επομένως θα έχουν και ίδια μέση τιμή και διασπορά

Μια συνάρτηση των ζ.μ.  $X_1, \dots, X_n$  που δεν εξαρτάται από αγνώστες παραμέτρους ονομάζεται στατιστική συνάρτηση.

Στατιστική συνάρτηση που μας ενδιαφέρει  
πολύ είναι το  $\sum_{i=1}^n X_i = X_1 + X_2 + \dots + X_n$ ,

ο δειγματικός μέσος  $\bar{X} : \frac{X_1 + \dots + X_n}{n}$

και η δειγματική διασπορά :

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

Μας ενδιαφέρει πολύ αν η στατιστική τι κατανομή  
ακολουθούν αυτές οι στατιστικές συναρτήσεις,  
και ιδιαίτερα το  $\boxed{\bar{X}}$ .

# Τυχαία μεταβλητή $X$ (π.χ. εισόδημα)

Υποτιθέμενη κατανομή

$F$  ή  $f$

Μέση τιμή

$\mu$

Διασπορά

$\sigma^2$

Πληθυσμός

Τυχαίο δείγμα από  
'παραμοιόζουσες'  $X$  :  
 $X_1, \dots, X_n$

Μέση τιμή

$\bar{X}$

Διασπορά

$S^2$

Δείγμα

# Στατιστική

Παραμετρική

Μη παραμετρική ή  
Απαραμετρική

Έστω τ.μ.  $X$  που περιγράφει κάποιο χαρακτηριστικό του πληθυσμού που μας ενδιαφέρει. Η τ.μ.  $X$  ή ένα τυχαίο δείγμα από πανομοιότυπες τ.μ.  $X'$  που το συμβολίζουμε  $X_1, X_2, \dots, X_n$  ακολουθεί κάποια κατανομή  $F$ . Μπορεί να ξέρουμε ή απλά να θέλουμε να κάνουμε κάποιες υποθέσεις για την  $F$  ή να μην θέλουμε να κάνουμε καμία υπόθεση.

Στην παραμετρική στατιστική υποθέτουμε ότι η  $F$  είναι γνωστή αναλυτικά χωρίς εκτός από κάποιες παραμέτρους.

Π.χ. Γάμμα με παραμέτρους  $a$  και  $b$  οι οποίες είναι αγνώστες

ή

Poisson με παράμετρο

$\lambda$  που είναι αγνώστ.

Θέλουμε να μάθουμε τις αγνώστες παραμέτρους

Στην απαραμετρική στατιστική δεν ξέρουμε τίποτα για την  $F$  ή από θέλουμε να μάθουμε τις ιδιότητες της  $F$ .

Οι ημερομηνίες έρχονται από το 2.δ.  
 $X_1, \dots, X_n$ .

## Παραμετρική Στατιστική

Έστω ημερομηνίες (χαρακτηριστικό του ημερομηνίου) που περιγράφεται από την τ.μ.  $X$  η οποία ακολουθεί κατανομή ή πιο συγκεκριμένα σ.η.η. ή σ.μ.η.

Για γνωστές συναρτησιακές μορφές εκτός της παραμέτρου  $\theta$  που είναι αγνώστη.

Το  $\theta$  παίρνει τιμές σε ένα χώρο  $\Theta$  ο οποίος ονομάζεται παραμετρικός χώρος.

Συμβολισμός  $X_1, \dots, X_n$  τ-δείγμα  
 $x_1, \dots, x_n$  δεδομένα δηλαδή  
συγκεκριμένες τιμές των  
τ.μ.  $X_1, \dots, X_n$ .

Θέλουμε να εκτιμήσουμε το  $\theta$  με βάση το 2.δ.

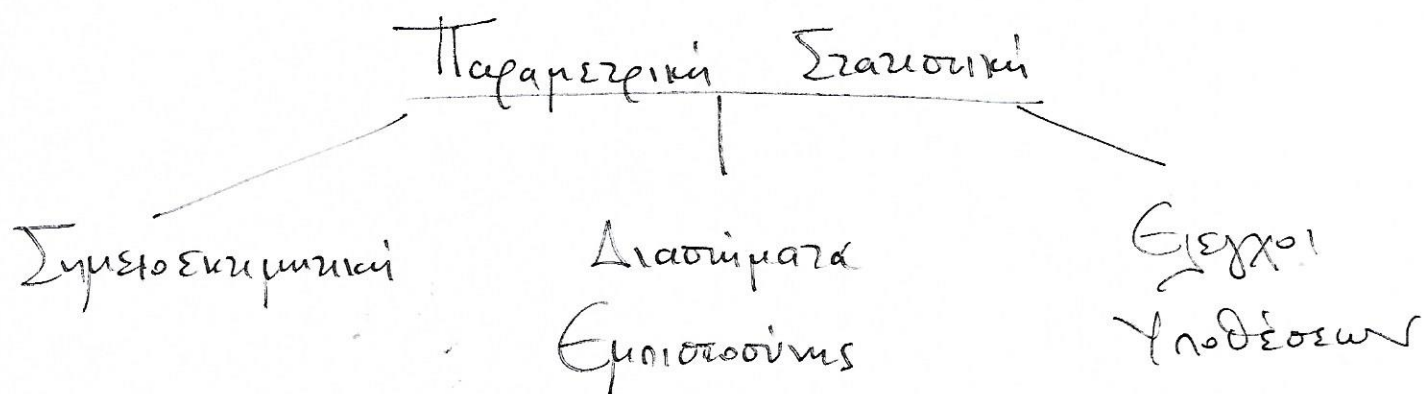
$X_1, \dots, X_n$ . Πιο συγκεκριμένα θέλουμε να βρούμε στατιστική συνάρτηση  $u(X_1, \dots, X_n)$  (εξαρτάται από τις  $X_1, \dots, X_n$  αλλά όχι από αγνώστες παραμέτρους)

τότε ώστε αν  $x_1, \dots, x_n$  δεδομένα που έχουμε στα χέρια μας,



η τιμή  $u(x_1, \dots, x_n)$  να είναι μια αριθμική  
εκτίμηση της παραμέτρου  $\theta$ .

Η στατιστική συνάρτηση  $u(X_1, \dots, X_n)$  η  
οποία είναι τ.μ. ονομάζεται σημειοεκτίμηση  
του  $\theta$ . (point estimate).



Τη συνάρτηση μ.η ή η.η ναί πάλιν δείχνουν ότι την συμβολίζουμε και  
 δώ και τώρα με  $f(x, \theta)$  για να δηλώνουμε την εξάρτηση ως προς  
 των αγνώστων παραμέτρων  $\theta$ . Ας δούμε μερικά παραδείγματα:

a) Έστω  $X \sim N(\mu, \sigma^2)$ ,  $\sigma^2$  γνωστό,  $\mu$  αγνώστο. τότε  

$$f(x; \theta) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\theta)^2}{2\sigma^2}}, \quad x \in \mathbb{R}$$

όπου  $\theta = \mu$  και  $\Theta = \mathbb{R}$

β) Έστω  $X \sim N(\mu, \sigma^2)$ ,  $\mu$  γνωστό και  $\sigma^2$  αγνώστο. τότε  

$$f(x; \theta) = \frac{1}{\sqrt{2\pi\theta}} e^{-\frac{(x-\mu)^2}{2\theta}}, \quad x \in \mathbb{R}$$

όπου  $\theta = \sigma^2$  και  $\Theta = (0, \infty)$ .

γ) Έστω  $X \sim P(\lambda)$ ,  $\lambda$  αγνώστο. τότε  

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

όπου  $\theta = \lambda$  και  $\Theta = (0, \infty)$

(δ)  $X \sim N(\mu, \sigma^2)$ ,  $\mu, \sigma^2$  αγνώστα. τότε

$$f(x, \theta) = \frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{(x-\theta_1)^2}{2\sigma_2^2}}, \quad x \in \mathbb{R}$$

όπου  $\theta = (\theta_1, \theta_2) = (\mu, \sigma^2)$  και  $\Theta = \{\theta = (\theta_1, \theta_2) : \theta_1 \in \mathbb{R}, \theta_2 \in (0, \infty)\}$ .

Οι μέθοδοι <sup>σημαντικές</sup> εκτίμηση <sup>και</sup>  $\theta$  που θα αναπτύξουμε εδώ είναι δύο: η μέθοδος μέγιστης πιθανοφάνειας (maximum likelihood method) και η μέθοδος των ροών (method of moments). Ξεκινάμε με τη μέθοδο μέγιστης πιθανοφάνειας. Η από κοινού σ.π.η. ή σ.π.η των  $X_1, X_2, \dots, X_n$  μπορούμε σαν συνάρτηση να  $\theta$  ονομάζεται συνάρτηση πιθανοφάνειας (likelihood function) και δίνεται από τον τύπο

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

$\uparrow$   
γινόμενο  
λόγω ανεξαρτησίας

$\uparrow$   
ομοσ  $f$   
διότι έχουν  
την ίδια  
κατανομή

Έστω  $x_1, x_2, \dots, x_n$  σταθερά (fixed)

και θεωρούμε την  $L(\theta, x_1, \dots, x_n)$  σαν συνάρτηση του  $\theta$



## Παράδειγμα - Διαισθητική ερμηνεία της μεθόδου μέγιστης πιθανοφάνειας

Εστω διωνυμικά περάματα  $X_1, \dots, X_n$  οδού

$$X_i \sim b(1, p) \text{ ή } X_i \sim b(1, \theta)$$

Η πιθανότητα επιτυχίας  $p$  είναι αγνώστη και  
θελουμε να εκτιμηθεί.

Αφού γίνουν τα περάματα εστω ότι έχουμε δεδομένα

0, 0, 1, 0, 0, 1, 1, 0, 0, 0, 0, ...

$$H \quad P(X_1=0, X_2=0, X_3=1, X_4=0, X_5=0, X_6=1, \dots, \theta)$$

δηλαδή η πιθανότητα να παρατηρήσουμε αυτά τα νούμερα  
εξαρτάται από την πιθανότητα επιτυχίας  $\theta$ .

Μπορούμε να θέσουμε τα εξής ερωτήματα:

Ποια είναι η πιο πιθανή τιμή για το  $\theta$ , η τιμή  
0.1 ή 0.9? Εδώ θα σκεφτείμε με βάση τα νούμερα  
να αναπλάσουμε 0.1 αφού βλέπουμε πολλά μηδενικά  
ή απλώς αγγίζει αποτυχίες.

Οι ερωτήσεις συνεχίζονται.

Ποια είναι η πιο πιθανή τιμή για το  $\theta$ , η τιμή  
0.2 ή 0.8? Κ.λ.π.

Ποια είναι η πιο πιθανή τιμή από όλες τις τιμές ( $\theta \in (0,1)$ ) να παράγασε αυτά τα νούμερα;

Ή αλλιώς, ποια τιμή του  $\theta$  μεγιστοποιεί την πιθανότητα να παρατηρήσουμε αυτά τα νούμερα;

Θέλουμε λοιπόν να μεγιστοποιήσουμε ως προς  $\theta$

$$\text{των } P(X_1=0, X_2=0, X_3=1, \dots, \theta)$$

Αυτή όμως είναι η συνάρτηση μέγας πιθανότητας, η απο κοινού των  $X_1, \dots, X_n$  ή όπως των ορίσαμε πριν η συνάρτηση πιθανοφάνειας.

Θέλουμε να βρούμε το

$$\max_{\theta \in \Theta} L(\theta, x_1, \dots, x_n) \quad \text{για fixed } x_1, \dots, x_n$$

$$= L(u(x_1, \dots, x_n), x_1, \dots, x_n)$$

Η τιμή  $u(x_1, \dots, x_n)$  είναι μια εκτίμηση του  $\theta$  με βάση τα fixed νούμερα  $x_1, \dots, x_n$ .

Η στατιστική συνάρτηση  $u(X_1, \dots, X_n) \equiv \hat{\theta}$  είναι

η εκτιμήτρια συνάρτηση μέγιστης πιθανοφάνειας του  $\theta$

Τις βρίσκουμε μερικο μας συνάρσεις; παίρνουμε παράγωγο.

### Βήματα ως μεθόδου

1)  $L(x_1, \dots, x_n, \theta)$

2)  $\ln L(x_1, \dots, x_n, \theta)$

3)  $\frac{d \ln L(x_1, \dots, x_n, \theta)}{d\theta} = 0$

και λύνουμε ως προς  $\theta$  οδου τζιρέ

$$\theta = u(x_1, \dots, x_n)$$

4)  $\hat{\theta} = u(X_1, \dots, X_n)$  η σημειοεκτιμήτρια του  $\theta$   
Ε.Μ.Π. του  $\theta$ .

### Παράδειγμα

Εστω  $X_1, \dots, X_n$  Τ.δ. από την  $b(1, \theta)$  κατανομή.

Να βρεθεί εκτιμήτρια μεγίστης πιθανότητας (ΕΜΠ) του  $\theta$ .

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

$$\theta^{x_1} (1-\theta)^{1-x_1} \theta^{x_2} (1-\theta)^{1-x_2} \dots \theta^{x_n} (1-\theta)^{1-x_n} =$$

$$\theta^{x_1+x_2+\dots+x_n} (1-\theta)^{n-(x_1+x_2+\dots+x_n)} =$$

$$\theta^{\sum x_i} (1-\theta)^{n-\sum x_i}$$

$$2) \ln L = \sum_{i=1}^n x_i \ln \theta + (n - \sum x_i) \ln(1 - \theta)$$

$$3) \frac{d \ln L}{d \theta} = \sum x_i \cdot \frac{1}{\theta} + (n - \sum x_i) \frac{-1}{1 - \theta} = 0 \Rightarrow$$

$$\sum x_i (1 - \theta) - (n - \sum x_i) \theta = 0 \Rightarrow$$

$$\sum x_i - \cancel{\theta \sum x_i} - n\theta + \cancel{\sum x_i \cdot \theta} = 0 \Rightarrow$$

$$\sum x_i - n\theta = 0 \Rightarrow \theta = \frac{\sum x_i}{n} \quad \leftarrow \begin{array}{l} \text{εκτίμηση} \\ \text{για} \\ \text{fixed } x_1, \dots, x_n \end{array}$$

$$4) \boxed{\hat{\theta} = \frac{\sum x_i}{n} = \bar{X}}$$

Αν κάποια διωνυμικά πειράματα δώσαν αποτελέσματα  
( $n=11$ )  $0, 0, 1, 0, 0, 1, 1, 0, 0, 0, 0$   
Ποια η εκτίμηση του  $\theta$ ?

$$\theta = \frac{0+0+1+0+0+1+1+0+0+0+0}{11} = \frac{3}{11} = 0.27$$

Η ΕΜΠ του  $\theta$  είναι  $\bar{X}$ .

(Παρατήρηση: για να εμαστε πιο σωστά πρέπει να ελέγξουμε  
ση  $\frac{d^2 \ln L}{d \theta^2} < 0$  αλλά δεν είναι πάντα εύκολο.)



## Παράδειγμα.

Έστω  $X_1, \dots, X_n$  τ.δ. από την  $N(\theta, 1)$ .

Να βρεθεί ΕΜ.Π. του  $\theta$  του μέσου της κανονικής κατανομής.

$$1) L(x_1, \dots, x_n, \theta) = \prod_{i=1}^n f(x_i, \theta) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x_1-\theta)^2}{2}} \cdot \dots \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{(x_n-\theta)^2}{2}} =$$
$$\left(\frac{1}{\sqrt{2\pi}}\right)^n e^{-\frac{\sum_{i=1}^n (x_i-\theta)^2}{2}}$$

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

$$3) \frac{d \ln L}{d\theta} = -\frac{1}{2} \cdot 2 \sum_{i=1}^n (x_i-\theta) \cdot (-1) = \sum_{i=1}^n (x_i-\theta) = 0 \Rightarrow$$

$$\sum x_i - n\theta = 0 \Rightarrow \theta = \frac{\sum x_i}{n} = \bar{x}$$

$$4) \hat{\theta} = \bar{X} \quad \text{ΕΜΠ του } \theta.$$

## Παράδειγμα

Έστω  $X_1, \dots, X_n$  τ.δ. από την εκθετική κατανομή

με σ.η.η.  $f(x, \theta) = \theta \cdot e^{-\theta x}$ ,  $x > 0, \theta > 0$ .

Να βρεθεί ΕΜΠ του  $\theta$ .

$$1) L(x_1, \dots, x_n, \theta) = \prod_{i=1}^n f(x_i, \theta) = \theta \cdot e^{-\theta x_1} \cdot \dots \cdot \theta e^{-\theta x_n} = \theta^n e^{-\theta \sum x_i}$$

$$2) \ln L = n \ln \theta - \theta \sum x_i$$

$$3) \frac{d \ln L}{d \theta} = \frac{n}{\theta} - \sum x_i = 0 \Rightarrow n - \theta \sum x_i = 0 \Rightarrow$$

$$n = \theta \sum x_i \Rightarrow \theta = \frac{n}{\sum x_i} = \frac{1}{\bar{x}}$$

$$4) \hat{\theta} = \frac{1}{\bar{X}}$$

## Θεώρημα του αναλλοίωτου

Έστω ότι θέλουμε να εκτιμήσουμε μια συνάρτηση του  $\theta$  ως προς την  $h(\theta)$ .

Όταν η συνάρτηση  $h$  είναι 1-1 τότε για τις εκτιμήσεις μεθόδου πιθανοφάνειας ισχύει ότι

$$\widehat{h(\theta)} = h(\widehat{\theta})$$

### Εφαρμογή

Για το προηγούμενο παράδειγμα της εκθετικής κατανομής,

$$\mu = E(X_1) = \dots = E(X_n) = \frac{1}{\theta}.$$

Να βρεθεί ΕΗΠ της μέσης τιμής  $\mu$ .

$$\widehat{\mu} = \frac{\widehat{1}}{\widehat{\theta}} = \frac{1}{\widehat{\theta}} = \frac{1}{\frac{1}{\overline{X}}} = \overline{X}$$

↑  
"ανταποκρίνου"

$$h(\theta) = \frac{1}{\theta} \quad 1-1 \text{ συνάρτηση.}$$

Από παράδειγμα:

$$h(\theta) = \theta^3 \quad 1-1$$

$$\text{Τότε} \quad \widehat{h(\theta)} = \widehat{\theta^3} = \widehat{\theta}^3 = \left(\frac{1}{\overline{X}}\right)^3$$

Συμπέρασμα: η εκτίμηση  $\overline{X}$  πάντα

εκτιμά τη μέση τιμή  $\mu$  της κατανομής των  $X_i, i=1, \dots, n$ .

# Άλλες περιπτώσεις ΕΜΠ.

Πεδίο ορισμού εξαρτάται από το  $\theta$ .

## Παράδειγμα

Έστω  $X_1, \dots, X_n$  τ.δ. από την κατανομή με π.π.  $f(x, \theta) = \frac{1}{\theta}$ ,  $0 < x \leq \theta$ ,  $\theta > 0$ .

Το πεδίο ορισμού εξαρτάται από το  $\theta$  οπότε εδώ δεν μπορεί να χρησιμοποιηθεί η διαδικασία διαφόρισης.

Η συνάρτηση πιθανότητας είναι

$$L(\theta, x_1, \dots, x_n) = \begin{cases} \left(\frac{1}{\theta}\right)^n & 0 < x_i \leq \theta \quad \forall i \\ 0 & \text{αλλιώς} \end{cases}$$

Θα τη γράψουμε και σαν (για δεδομένα (fixed)  $x_1, \dots, x_n$ )

$$\left(\frac{1}{\theta}\right)^n I_{(x_1 \leq \theta)} I_{(x_2 \leq \theta)} \cdots I_{(x_n \leq \theta)} =$$

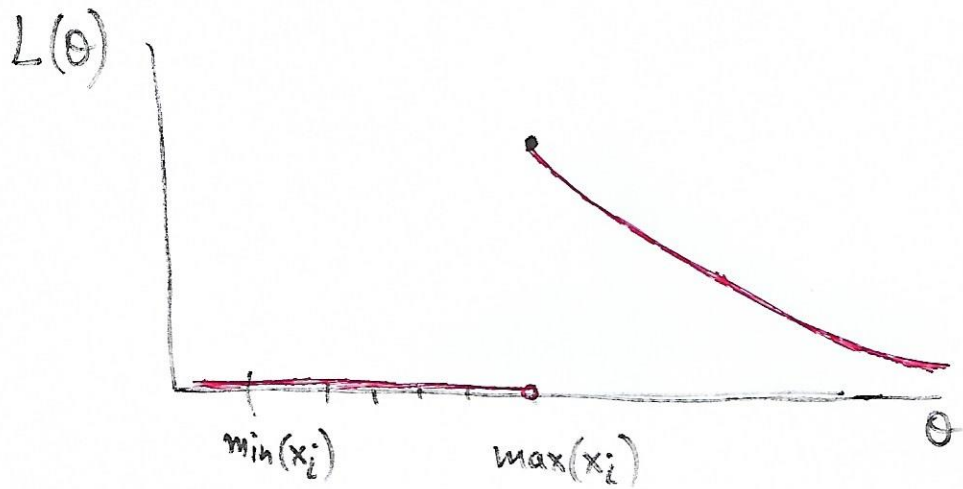
$$\left(\frac{1}{\theta}\right)^n I_{(\theta \geq x_1)} I_{(\theta \geq x_2)} \cdots I_{(\theta \geq x_n)} \stackrel{(*)}{=}$$

$$\text{οδον } I_A(x) = \begin{cases} 1 & \text{αν } x \in A \\ 0 & \text{αν } x \notin A \end{cases} \quad \text{δείκτης συνάρτηση}$$

$$\stackrel{(*)}{=} \left(\frac{1}{\theta}\right)^n I_{(\theta \geq \max_{1 \leq i \leq n} (x_i))}$$



Η γραφική παράσταση της  $L(\theta, x_1, \dots, x_n)$  σαν συνάρτηση του  $\theta$  για fixed  $x_1, \dots, x_n$  είναι:



Η  $L(\theta)$  είναι φθίνουσα συνάρτηση, άρα το μέγιστο της επιτυγχάνεται για  $\theta = \max(x_i)$  ή

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

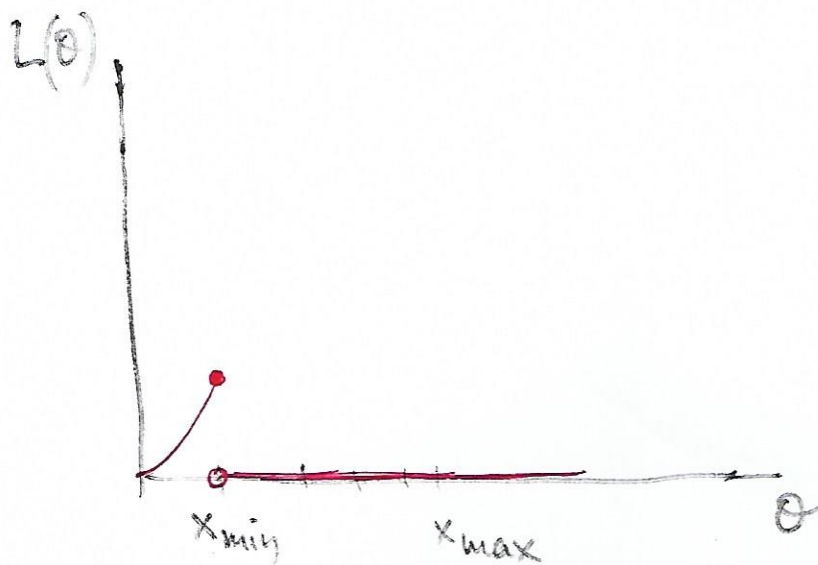
### Παράδειγμα

Έστω  $X_1, \dots, X_n$  ζ.δ. από την κατανομή με π.π.π.

$$f(x) = 2\theta^{-2} x^{-3}, \quad x \geq \theta, \quad \theta > 0$$

Να βρεθεί ΕΜΠ του  $\theta$ .

$$\begin{aligned} L(x_1, \dots, x_n, \theta) &= 2^n \theta^{-2n} \left( \prod_{i=1}^n x_i \right)^{-3} I_{(x_1 \geq \theta)} \cdots I_{(x_n \geq \theta)} = \\ &= 2^n \theta^{-2n} \left( \prod_{i=1}^n x_i \right)^{-3} I_{(\theta \leq x_1)} \cdots I_{(\theta \leq x_n)} = \\ &= 2^n \theta^{-2n} \left( \prod_{i=1}^n x_i \right)^{-3} I_{(\theta \leq \min(x_i))} \end{aligned}$$



Το μέγιστο ενοχληστικό για  $\theta = \min(x_i)$

$$\hat{\theta} = \min_{1 \leq i \leq n} (x_i)$$