

## Θεωρία Αποφάσεων Bayes.

ΠΑΡΑΤΗΡΗΣΗ: Σε πολλές περιπτώσεις οι ερευνητές δεν έχουν προτερες πληροφορίες για το  $\theta$ , για το λόγο αυτό πρέπει να χρησιμοποιήσουν τις λεγόμενες μη πληροφοριακές εκ των προτέρων κατανομές, δηλ εκ των προτέρων κατανομές που δε δίνουν πληροφορίες για την παράμετρο  $\theta$

Μια τέτοια μη πληροφοριακή εκ των προτέρων κατανομή είναι αυτή που έχει προταθεί στη βιβλιογραφία από τον Jeffrey και δίνεται από τη σχέση:

Jeffrey's prior

$$J(\theta) \propto \left[ I_F^*(\theta) \right]^{1/2}, \text{ όπου } I_F^*(\theta) = -E_{X|\theta} \left[ \frac{\partial^2}{\partial \theta^2} \log f(X|\theta) \right]$$

↓  
πληροφορία του Fisher.

Παράδειγμα Έστω ότι  $X|\theta \sim B(n, \theta)$

Να βρεθεί η εκ των προτέρων μη πληροφοριακή κατανομή του Jeffrey και ακολούθως η  $J(\theta)$ .

Γνωρίζουμε ότι:

$$f(x|\theta) = \binom{n}{x} \theta^x (1-\theta)^{n-x} \quad x=0, \dots, n$$

Επομένως θα έχουμε:

$$\log f(x|\theta) = \log \binom{n}{x} + x \log \theta + (n-x) \log (1-\theta) \Rightarrow$$

$$\frac{\partial \log f(x|\theta)}{\partial \theta} = \frac{x}{\theta} + \frac{n-x}{1-\theta} (-1) \Rightarrow$$

$$\frac{\partial \log f(x|\theta)}{\partial \theta} = \frac{x}{\theta} - \frac{n-x}{1-\theta} \Rightarrow \frac{\partial^2 \log f(x|\theta)}{\partial^2 \theta} = - \left[ \frac{x}{\theta^2} + \frac{n-x}{(1-\theta)^2} \right]$$

$$\text{ΟΠΩΤΕ } E_{X|\theta} \left[ \frac{x}{\theta^2} + \frac{n-x}{(1-\theta)^2} \right] = \frac{E(x)}{\theta^2} + \frac{n-E(x)}{(1-\theta)^2} = \frac{n\theta}{\theta^2} + \frac{n-n\theta}{(1-\theta)^2} = \frac{n}{\theta} + \frac{n}{(1-\theta)} =$$

$$= \frac{n(1-\theta) + n\theta}{\theta(1-\theta)} = \frac{n}{\theta(1-\theta)} \Rightarrow E_{X|\theta} \left[ \frac{x}{\theta^2} + \frac{n-x}{(1-\theta)^2} \right] = \frac{n}{\theta(1-\theta)}$$

Άρα θα έχουμε ότι  $f(\theta) \propto \left[ \frac{n}{\theta(1-\theta)} \right]^{1/2} = n^{1/2} \theta^{-1/2} (1-\theta)^{-1/2}$

↑  
αναίλοχη

Μπορώ να θεωρήσω  
το  $n^{1/2}$  ως σταθερά

Μπορούμε να παραλείψουμε τον όρο  $n^{1/2}$ , διότι έχουμε ποσότητα  
αναίλοχη (της  $f(\theta)$ ) και ενδιαφερόμαστε να κρατήσουμε μόνο  
όρους που περιέχουν το  $\theta$ .

ΟΠΩΤΕ  $f(\theta) \propto \theta^{-1/2} (1-\theta)^{-1/2}, 0 < \theta < 1$  (\*)

Σημείωση: Για να εντοπίσω την εκ των προτέρων κατανομή  
θα βασισθώ πάντα στο π.ο της  $\theta$ .

(\*) Θυμίζει Βήτα κατανομή  $\frac{x^{\alpha-1} (1-x)^{\beta-1}}{B(\alpha, \beta)}$

όπου  $x=\theta$   
 $\alpha-1 = -\frac{1}{2} \Rightarrow \alpha = \frac{1}{2}$

$\beta-1 = -\frac{1}{2} \Rightarrow \beta = \frac{1}{2}$

Επομένως η  
μη πληροφοριακή εκ των προτέρων  
κατανομή του  $\theta$ , θα είναι

η  $f(\theta) = \frac{\theta^{-1/2} (1-\theta)^{-1/2}}{B(\frac{1}{2}, \frac{1}{2})}$ , δηλ.  $\theta \sim \text{Beta}(\frac{1}{2}, \frac{1}{2})$



### Ορισμός 1: [Εκ των υστέρων Εκτιμητής Bayes]

Έστω τ.δ.  $X_1, \dots, X_n$  από ένα πληθυσμό με σ.π. ή σ.π.  $f(x|\theta)$ , όπου  $\theta$  η τιμή της τ.μ.  $\theta$  με εκ των προτέρων κατανομή  $f(\theta)$ .

Ο εκ των υστέρων Εκτιμητής Bayes της  $g(\theta)$  θα ορίζεται από τη σχέση  $T^* = E_{\theta|X}(g(\theta)) = \begin{cases} \sum_{\theta} g(\theta) \cdot (\text{σ.π. της } \theta|X) \\ \int_{\theta} g(\theta) \cdot f(\theta|X) d\theta. \end{cases}$

### Ορισμός 2: [Συνάρτηση Διακινδυνεύσεως κατά Bayes]

Ορίζουμε ως συνολ. Διακινδυνεύσεως κατά Bayes  $R(\mathcal{F}, d)$  για τον κανόνα απόφασης  $d$  έχοντας εκ των προτέρων κατανομή  $f$ , να είναι η συνολ. για την οποία ισχύει:

$$R(\mathcal{F}, d) = E_{\theta}[R(\theta, d)], \text{ όπου } R(\theta, d) = E_{X|\theta}[L(\theta, d)]$$

### Ορισμός 3: [Εκτιμητής Bayes]

Ένας εκτιμητής  $d^*$  ορίζεται να είναι εκτιμητής Bayes

$$\text{αν } R(\mathcal{F}, d^*) = \min_{d \in D} R(\mathcal{F}, d).$$

Αν δηλ. ελαχιστοποιεί την συνολ. Διακινδυνεύσεως κατά Bayes.

$f = \mathcal{F} = \pi$  το ίδιο

**ΘΕΩΡΗΜΑ:** Έστω τ.δ.  $X_1, \dots, X_n$  από ένα πληθυσμό με σ.π. ή σ.π.  $f(x|\theta)$ , όπου  $\theta$  η τιμή της τ.μ.  $\theta$  με εκ των προτέρων  $f(\theta)$

Αν  $L(\theta, d)$  είναι μια συνολ. απώλειας, τότε ο εκτιμητής Bayes της παραμετρικής συνολ. του  $\theta$ ,  $g(\theta)$  θα είναι το στατιστικό  $d^*$  που ελαχιστοποιεί την ποσότητα  $E_{\theta|X}[L(g(\theta), d)]$ .

Αποδ.

## Αποδ

Γνωρίζουμε ότι ο  $d^*$  εφ' ορισμού ελαχιστοποιεί την ποσότητα  $R(f, d)$   
δηλ. εφ' ορισμού ελαχιστοποιεί τη  $E_\theta[R(\theta, d)]$ .

Επομένως (στη συνεχή περίπτωση) θα έχουμε ότι ελαχιστοποίηση  
της  $E_\theta[R(\theta, d)]$ , σημαίνει ελαχιστοποίηση της ποσότητας

$$\int_\theta R(\theta, d) f(\theta) d\theta, \text{ όμως εφ' ορισμού } R(\theta, d) = E_{x|\theta}[L(\theta, d)] = \int_{\mathcal{X}} L(\theta, d) f(x|\theta) dx$$

$$\text{οπότε } \int_\theta E_{x|\theta} L(g(\theta, d)) f(\theta) d\theta = \int_\theta \int_{\mathcal{X}} L(g(\theta, d)) f(x|\theta) f(\theta) d\theta dx =$$

$$= \int_\theta \int_{\mathcal{X}} L(g(\theta, d)) \underbrace{h(x, \theta)}_{\substack{\text{η από κοινού} \\ \text{των } x, \theta.}} d\theta dx$$

Γνωρίζουμε από τη θεωρία Πιθανοτήτων ότι  $P(A|B)P(B) = P(A \cap B) = P(B|A)P(A)$   
με αντιστοίχη εφαρμογή σε όρους τ.μ. να είναι:  $f(x|\theta)f(\theta) = h(x, \theta) = f(\theta|x)f(x)$

Επομένως θέλουμε να ελαχιστοποιήσουμε ως προς  $d$  την ποσότητα

$$\int_\theta \int_{\mathcal{X}} \underbrace{L(g(\theta, d)) f(\theta|x)}_{\substack{\text{Εξαρτά από} \\ \text{το } d}} \underbrace{f(x)}_{\substack{\text{περιθώρια} \\ \text{του } x.}} d\theta dx$$

αρκεί επομένως να ελαχιστοποιήσουμε την ποσότητα

$$\int_\theta L(g(\theta, d)) f(\theta|x) d\theta = E_{\theta|x}[L(g(\theta, d))] \text{ άρα}$$

αποδείξαμε το ζητούμενο.



Πόρισμα 1: Αν η  $L(\theta, d)$  ~~είναι~~<sup>ως</sup> η συν/ση απώλειας, ορίζεται το τετραγωνικό σφάλμα τότε:

ο  $d^*$  εκτιμητής Bayes θα είναι η αναμενόμενη τιμή της  $g(\theta)$  ως προς  $\theta$  δθέντος  $X$ , δηλ  $d^* = E_{\theta|X}(g(\theta))$

Αποδ

Έχουμε αποδείξει από το προηγ. θεώρημα ότι ο  $d^*$  ελαχιστοποιεί τη  $E_{X|g}[L(g(\theta), d)]$ . Επομένως δεδομένου ότι ως συν/ση απώλειας έχουμε το Τ.Σ. πρέπει να ελαχιστοποιήσουμε ως προς  $d$  την ποσότητα  $E_{\theta|X}[(g(\theta) - d)^2]$

Στη συνεχή περίπτωση θα ισχύει:

$$E_{\theta|X}(g(\theta) - d)^2 = \int (g(\theta) - d)^2 f(\theta|X) d\theta \quad \text{παραγωγίζουμε ως προς } d$$

$$\int -2(g(\theta) - d) f(\theta|X) d\theta = 0 \Rightarrow \int (g(\theta) - d) f(\theta|X) d\theta = 0$$

$$\Rightarrow \int g(\theta) f(\theta|X) d\theta = \underbrace{\int d f(\theta|X) d\theta}_1 = d \quad \text{οπότε } d^* = \int g(\theta) f(\theta|X) d\theta = E_{\theta|X}[g(\theta)]$$

Υπολογίζοντας και τη 2<sup>η</sup> παραγωγή θα έχουμε  $2 \int f(\theta|X) d\theta = 2 \cdot 1 = 2 > 0$

Επομένως για  $d^*$  έχουμε ελαχιστοποίηση και άρα αποδεικνύεται το ζητούμενο.

Πόρισμα 2 Αν η συνιστη αμώλειας είναι το απόλυτο σφάλμα, δηλ.

$L(\theta, d) = |d - \theta|$ , τότε ο εκτιμητής Bayes της  $\theta$  θα είναι

η διαίρεσος της  $f(\theta|x)$ , δηλ.  $d^* = \text{median}(f(\theta|x))$

$d^* = \text{median}(f(\theta|x))$ , <sup>αίρα</sup> το σημείο εκείνο για το οποίο ισχύει ότι

$$\int_{-\infty}^{d^*} f(\theta|x) d\theta = \int_{d^*}^{+\infty} f(\theta|x) d\theta = 0,5 \quad \text{διότι} \quad \int_{-\infty}^{+\infty} f(\theta|x) d\theta = 1$$

και

$d^*$  διαίρεσος.

Αποδ

Έχουμε αποδείξει ότι ο  $d^*$  ελαχιστοποιεί την ποσότητα  $E_{\theta|x} L(\theta, d)$ , δηλ

την ποσότητα  $\int |d - \theta| f(\theta|x) d\theta$  που είναι ισοδύναμη με την ποσότητα

$$\underbrace{\int_{-\infty}^d (d - \theta) f(\theta|x) d\theta}_{\theta < d} + \underbrace{\int_d^{+\infty} - (d - \theta) f(\theta|x) d\theta}_{\theta > d}$$

άρα  $|d - \theta| = d - \theta$

$|d - \theta| = - (d - \theta)$

ΥΠΕΝΟΜΙΣΗ (κανόνας Leibnitz)

Έστω η συνιστη  $F(x) = \int_{A(x)}^{B(x)} f(x, t) dt$ , όπου οι συνιστες  $A(x)$  και  $B(x)$

είναι παραγωγίσιμες ως προς  $x$  και  $f(x, t)$  και  $\frac{\partial f(x, t)}{\partial x}$

είναι συνεχείς ως προς  $x, t$ . Τότε η παράγωγος

$$\frac{dF}{dx} = \int_{A(x)}^{B(x)} \frac{\partial f(x, t)}{\partial x} dt + f(x, B(x)) \frac{dB}{dx} - f(x, A(x)) \frac{dA}{dx}$$



Επομένως παραχρησιμοποιώντας την ποσότητα  $A$  ως προς  $d$  θα έχουμε:

$$\frac{\partial A}{\partial d} = \int_{-\infty}^d f(\theta|x) d\theta + (d-d) f(d|x) \cdot 1 + \int_d^{+\infty} -f(\theta|x) d\theta + (d-d) f(d|x) \cdot 1 = 0 \Rightarrow$$

$$\boxed{\int_{-\infty}^{d^*} f(\theta|x) d\theta = \int_{d^*}^{+\infty} f(\theta|x) d\theta}$$

Πρέπει επιπλέον να ελέγξουμε ότι η 2<sup>η</sup> παράγωγος της ποσότητας  $A$  είναι θετική για να έχω σημείο ελαχίστου.

Θα είναι:  $\frac{\partial^2 A}{\partial d^2} = f(d|x) + f(d|x) = 2 \underbrace{f(d|x)}_{>0}$ , άρα αποδεικνύεται το ζητούμενο.

Παράδειγμα: Έστω το  $X_1, \dots, X_n$  από  $P(\theta)$ , όπου  $\theta$  η τιμή της τ.μ  $\theta$  με εκ των προτέρων κατανομή τη Γαμμα  $(p, 1/q)$

- Να βρεθεί ο εκτιμητής Bayes της  $\theta$  και της  $e^{-\theta}$ , με συνθήκη απώλειας το  $T.S$
- Να βρεθεί ο εκτιμητής Bayes της  $\theta$ , με συνθήκη απώλειας τη συνθήκη Linex  
δηλ.  $L(\theta, d) = e^{k(d-\theta)} - k(d-\theta) - 1$ .

ΛΥΣΗ

i) Γνωρίζουμε ότι  $f(\theta) = \frac{\theta^{p-1} e^{-\theta q}}{(\frac{1}{q})^p \Gamma(p)}$ ,  $\theta > 0$  και  $f(x|\theta) = \frac{e^{-\theta} \theta^x}{x!}$   $x=0,1,2,\dots$

Το π.ο της  $f(x|\theta)$  είναι ανεξάρτητο της παραμέτρου  $\theta$ , άρα μπορούμε να χρησιμοποιήσουμε την αναλογία  $f(\theta|x) \propto f(\theta) f(x|\theta)$ .

Θα έχουμε ότι  $f(\theta|x) \propto \frac{\theta^{p-1} e^{-\theta q}}{(\frac{1}{q})^p \Gamma(p)} \cdot \frac{e^{-n\theta} \theta^{\sum x_i}}{\prod_{i=1}^n x_i!}$ ,  $\theta > 0$

οπότε  $f(\theta|x) \propto \underbrace{\theta^{p-1} e^{-\theta q} e^{-n\theta} \theta^{\sum x_i}}_A = \theta^{p+\sum x_i-1} \cdot e^{-\theta(q+n)}$ ,  $\theta > 0$





β Τρόπος (Αν δινόταν τυπολόγιο)

Γνωρίζουμε ότι  $\theta/x \sim \text{Γάμμα} (\alpha = p + \sum x_i, \beta = \frac{1}{q+n})$

$$\text{άρα } d_1^* = E(\theta) = \alpha \cdot \beta = (p + \sum x_i) \frac{1}{q+n} = \frac{p + \sum x_i}{q+n}$$