Κεφάλαιο 4

Διαστηματική Εκτίμηση

4.1 Διαστήματα Εμπιστοσύνης

Έστω $\underset{\sim}{X}=(X_1,X_2,\dots,X_n)$ τ.δ. από την κατανομή $f(x;\theta)$ ή $p(x;\theta)$. Το διάστημα εμπιστοσύνης (δ.ε.) ορίζεται με τη βοήθεια δύο στατιστικών συναρτήσεων του τ.δ. $\underset{\sim}{X}$: $L(\underset{\sim}{X})$, $U(\underset{\sim}{X})$ οι οποίες ορίζουν διάστημα $\left(L(\underset{\sim}{X}),U(\underset{\sim}{X})\right)$. Οι σ.σ. $L(\underset{\sim}{X})$, $U(\underset{\sim}{X})$ είναι το κατώτερο και το ανώτερο όριο του δ.ε. αντίστοιχα, και ικανοποιούν τη σχέση:

$$P_\theta\{L(\underset{\sim}{X})<\theta<U(\underset{\sim}{X})\}=1-\alpha$$

Παράδειγμα Έστω $X_1,\dots,X_n$ τ.δ. από την Κανονική Κατανομή με μ άγνωστη παράμετρο και σ^2 γνωστό. Γνωρίζουμε ότι:

$$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\sim N(0,1).$$

Άρα

$$P\left(-z_{\alpha/2} < \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} < z_{\alpha/2}\right) = 1-\alpha.$$

Ισοδύναμα,

$$P\left(\bar{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right)$$
$$= 1-\alpha.$$

Είναι: $L(\underset{\sim}{X}) = \bar{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$

$$\mathcal{U}(\underset{\sim}{X}) = \bar{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}. \quad \boxtimes$$

Διαπιστώνουμε ότι:

Τα άκρα του δ.ε. είναι τυχαίες μεταβλητές, και ότι το παραπάνω δ.ε. είναι συμμετρικό.

Γενικά, περισσότερα από ένα δ.ε. πληρούν την ιδιότητα: $P_\theta\{L(\underset{\sim}{X}) < \theta < \mathcal{U}(\underset{\sim}{X})\} = 1-\alpha$. Καλύτερο είναι εκείνο με το μικρότερο μήκος.

<u>Ορισμός (δε)</u> Έστω $\underset{\sim}{X}=(X_1,\dots,X_n)$ τ.δ. από πληθυσμό με σ.π.π. $f(x;\theta)$ ή $p(x;\theta)$. Έστω $L=L(\underset{\sim}{X})$, $U=U(\underset{\sim}{X})$ σ.σ. με $L(\underset{\sim}{X})\le U(\underset{\sim}{X})$ $\forall\, \underset{\sim}{X}$ και

$$P_\theta\{L(\underset{\sim}{X})<g(\theta)<U(\underset{\sim}{X})\}\ge 1-\alpha$$

$\forall \theta\in \Theta$. Τότε το τυχαίο διάστημα $(L(\underset{\sim}{X}), U(\underset{\sim}{X}))$ είναι ένα $100(1-\alpha)\%$ δ.ε. για την <u>παραμετρική</u> συνάρτηση $g(\theta)$. ⊠

<u>Παρατηρήσεις</u> (i) Μία πραγματοποίηση του τυχαίου διαστήματος $(\ell(\underset{\sim}{x}), u(\underset{\sim}{x}))$ είναι επίσης ένα $100(1-\alpha)\%$ δ.ε της $g(\theta)$.

(ii) Η πιθανότητα $P_\theta\{L(\underset{\sim}{X})<g(\theta)<U(\underset{\sim}{X})\}$ ονομάζεται <u>πιθανότητα</u> <u>κάλυψης</u> της $g(\theta)$.

(iii) Στον ορισμό και ειδικότερα στην πιθανότητα $P_\theta\{L(\underset{\sim}{X})<g(\theta)<U(\underset{\sim}{X})\}$

οι ποσότητες $L(\underset{\sim}{X}), U(\underset{\sim}{X})$ είναι οι τ.μ. και η ποσότητα $g(\theta)$ είναι σταθερή (άγνωστη, αλλά σταθερή). Ισοδύναμα, μπορούμε να γράψουμε: -127-

$$P_\theta\{g(\theta) \in (L(\underset{\sim}{X}), U(\underset{\sim}{X}))\} \geq 1-\alpha$$

$\forall \theta \in \Theta$. ⊠

<u>Ερμηνεία του δ.ε</u>: Εάν επαναληφθεί η δειγματοληψία αρκετές φορές και καταγράφουμε κάθε φορά την πραγματοποίηση του τυχαίου διαστήματος $(\ell(\underset{\sim}{x}), u(\underset{\sim}{x}))$, τότε αναμένουμε βάση της πιθανότητας, το διάστημα αυτό να περιέχει την παράμετρο θ, με ποσοστό $(1-\alpha)100\%$. <u>Π.Χ.</u> Αν επαναλάβουμε το πείραμα 100 φορές και $\alpha=0.05$, τότε αναμένουμε από τα 100 δ.ε. $(\ell(\underset{\sim}{x}), u(\underset{\sim}{x}))$ που θα προκύψουν, τα 95 από αυτά να περιέχουν την άγνωστη παράμετρο, και στα 5 από αυτά να μην περιέχεται η παράμετρος στο εσωτερικό τους. ⊠

4.2 Αντιστρεπτή Ποσότητα

<u>Ορισμός</u> Έστω $X_1, \ldots, X_n$ τ.δ. από πληθυσμό με σ.π.π. $f(x;\theta)$ ή $p(x;\theta)$. Έστω ακόμα $Q = Q(\underset{\sim}{X};\theta)$ <u>τυχαία συνάρτηση</u> του δείγματος $\underset{\sim}{X} = (X_1, \ldots, X_n)$ η οποία περιλαμβάνει το θ. Εάν η κατανομή της Q <u>δεν</u> εξαρτάται από το θ, τότε η συνάρτηση Q καλείται <u>αντιστρεπτή</u> ποσότητα (pivotal quantity). $\boxtimes$

π.χ. Η ποσότητα $\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}}$ είναι μια αντιστρεπτή ποσότητα, διότι (i) είναι συνάρτηση των τ.μ. του δείγματος και της άγνωστης παραμέτρου μ και (ii) η κατανομή της, $N(0,1)$, είναι ανεξάρτητη από την παράμετρο μ. Η ποσότητα $Q = \bar{X} - \mu$ είναι επίσης αντιστρεπτή ποσότητα διότι η κατανομή της Q είναι η $N(0,\sigma^2)$ και είναι ανεξάρτητη της $\theta = \mu$.

Πώς εφαρμόζουμε την αντιστρεπτή ποσότητα για την κατασκευή δ.ε;

Βήματα 1. Βρίσκουμε μία αντιστρεπτή ποσότητα.

2. Θεωρούμε την πιθανότητα:

$$P_\theta\{q_1 < Q(\underset{\sim}{X};\theta) < q_2\} = 1-\alpha$$

για q_1, q_2 αριθμούς στο στήριγμα της Q με $q_1 < q_2$.

3. Στην παραπάνω πιθανότητα, αντικαθιστούμε όπου $Q(\underset{\sim}{X};\theta)$ με την αναλυτική συναρτησιακή σχέση $Q = Q(\underset{\sim}{X};\theta)$ και λύνουμε ως προς την παράμετρο θ (ή $q(\theta)$ την παραμετρική συνάρτηση) που ενδιαφερόμαστε να κατασκευάσουμε το δ.ε. στην ανισότητα της πιθανότητας. Στόχος είναι το θ ή το $q(\theta)$ να είναι στο κεντρικό σκέλος της ανισότητας στο τελευταίο στάδιο επίλυσης.

4. Η παραπάνω επίλυση δίνει τη μορφή του δ.ε. Είναι το (ℓ, u), όπου ℓ, u είναι το αριστερό και δεξί σκέλος της ανισότητας, αντίστοιχα. Θα είναι $\ell = \ell(q_1, q_2)$, $u = u(q_1, q_2)$.

5. Υπολογισμός q_1, q_2: Επιστρέφουμε στην πιθανότητα του βήματος 2. Υπολογίζουμε την πιθανότητα, λαμβάνοντας υπόψη την κατανομή της Q και εξισώνουμε με την ποσότητα $1-\alpha$, όπου α δοθέν.

<u>Παράδειγμα 1</u> Για κανονικό πληθυσμό με μ άγνωστο, σ^2 γνωστό, θεωρούμε την ποσότητα $Q = \dfrac{\overline{X} - \mu}{\sigma/\sqrt{n}}$ ως αντιστρεπτή για την οποία ισχύει ότι $Q \sim N(0,1)$. Για το συγκεκριμένο παράδειγμα, τα βήματα της μεθόδου είναι:

1. Η αντιστρεπτή ποσότητα είναι η $Q = \overline{X} - \mu \big/ \sigma/\sqrt{n}$.

2. Έστω $q_1, q_2 \in \mathbb{R}$ τ.ω. $q_1 < q_2$ και
$P_\theta\{q_1 < Q < q_2\} = 1-\alpha$.

3. $P_\theta(q_1 < Q < q_2) = 1-\alpha$

$$\Leftrightarrow P_\theta\left(q_1 < \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} < q_2\right) = 1-\alpha$$

$$\Rightarrow P_\theta\left(q_1 \frac{\sigma}{\sqrt{n}} < \bar{X}-\mu < \frac{\sigma}{\sqrt{n}} q_2\right) = 1-\alpha$$

$$\Rightarrow P_\theta\left(-q_2 \frac{\sigma}{\sqrt{n}} < \mu-\bar{X} < -q_1 \frac{\sigma}{\sqrt{n}}\right) = 1-\alpha$$

$$\Rightarrow P_\theta\left(\bar{X}-q_2 \frac{\sigma}{\sqrt{n}} < \mu < \bar{X}-q_1 \frac{\sigma}{\sqrt{n}}\right) = 1-\alpha$$

4. Το δ.ε. είναι:

$$(\ell, u) = \left(\bar{X}-q_2 \frac{\sigma}{\sqrt{n}}, \bar{X}-q_1 \frac{\sigma}{\sqrt{n}}\right).$$

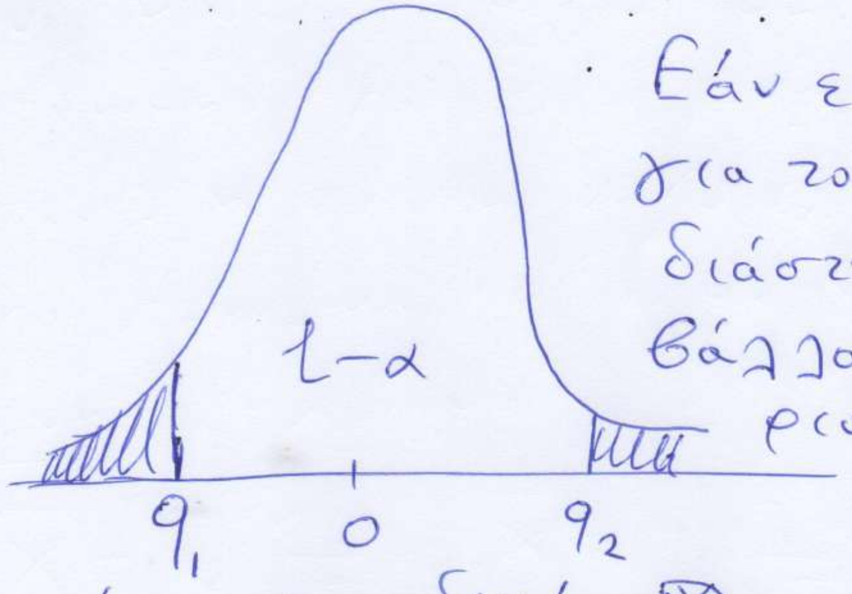
5. $P_\theta(q_1 < Q < q_2) = 1-\alpha$

$$\Leftrightarrow \int_{q_1}^{q_2} f(q)\,dq = 1-\alpha, \text{ όπου } f(q)$$

είναι η σ.π.π. της τυπικής Κανονικής κατανομής. Άρα $\Phi(q_2)-\Phi(q_1)=1-\alpha$ όπου Φ η σ.κ. της $N(0,1)$.

Άρα, συνολικά, το δ.ε. με βαθμό εμπιστοσύνης $1-\alpha$ είναι:

$$\left(\overline{X}-q_2\frac{\sigma}{\sqrt{n}},\ \overline{X}-q_1\frac{\sigma}{\sqrt{n}}\right)$$ με τα q_1, q_2 να πληρούν τη σχέση: $\Phi(q_2)-\Phi(q_1)=1-\alpha$. Το παραπάνω δ.ε. είναι συνάρτηση των q_1, q_2 δηλαδή δύο αγνώστων ενώ η σχέση που γνωρίζουμε για τους αγνώστους q_1, q_2 είναι μία. Άρα, το δ.ε. που προκύπτει δεν είναι μοναδικά ορισμένο. Τα δυνατά δ.ε είναι άπειρα, διότι υπάρχει ένας ελεύθερος άγνωστος. Οποιοδήποτε ζευγάρι τιμών q_1, q_2 για τα οποία το εμβαδόν που περικλείεται μεταξύ των καθέτων στα q_1, q_2 και της καμπύλης της πυκνότητας της Κανονικής κατανομής είναι $1-\alpha$ δίνει ένα $(1-\alpha)100\%$ βαθμό εμπιστοσύνης για το $\theta=\mu$.



Εάν ενδιαφερόμαστε για το βέλτιστο διάστημα, τότε επιβάλλουμε έναν περιορισμό και η λύση ως προς q_1, q_2 θα είναι μοναδική. $\boxtimes$

<u>Παράδειγμα 2</u> Συνεχίζοντας το Παράδειγμα 1, όπου βρέθηκε το $100(1-\alpha)\%$ δ.ε. για το μ στην περίπτωση ενός τ.δ. από την $N(\mu, \sigma^2)$, αναζητούμε το ζευγάρι τιμών q_1, q_2 που περικλείουν πιθανότητα $1-\alpha$ και το δ.ε. είναι βέλτιστο. Το δ.ε. είναι:

$$\left(\bar{X} - q_2 \frac{\sigma}{\sqrt{n}}, \bar{X} - q_1 \frac{\sigma}{\sqrt{n}}\right)$$ με τα q_1, q_2

να τηρούν τη σχέση:

$\Phi(q_2) - \Phi(q_1) = 1-\alpha$. Το μήκος του δ.ε. είναι: $\text{length} = \frac{\sigma}{\sqrt{n}}(q_2 - q_1)$.

Αναζητούμε το ελάχιστο της παραπάνω συνάρτησης ή ισοδύναμα της $(q_2 - q_1)$ ως προς τα q_1, q_2 με συνθήκη: $\Phi(q_2) - \Phi(q_1) = 1-\alpha$. Προσεγγίζουμε το πρόβλημα με πολλαπλασιαστές Lagrange. Έστω:

$$g(q_1, q_2, \lambda) = q_2 - q_1 + \lambda\left(\Phi(q_2) - \Phi(q_1) - 1 + \alpha\right)$$

Αναζητούμε το $\min g(q_1, q_2, \lambda)$.

$$\frac{\partial g}{\partial q_1} = -1 - \lambda \frac{\partial \Phi(q_1)}{\partial q_1} = 0 \Rightarrow 1 = -\lambda f(q_1)$$

$$\frac{\partial g}{\partial q_2} = 1 + \lambda \frac{\partial \Phi(q_2)}{\partial q_2} = 0 \Rightarrow 1 = -\lambda f(q_2)$$

$$\Rightarrow f(q_1) = f(q_2).$$

$$\frac{\partial g}{\partial \lambda} = 0 \Rightarrow \Phi(q_2) - \Phi(q_1) = 1 - \alpha.$$

Η σχέση $f(q_1)=f(q_2)$ συνεπάγεται:
$q_1=q_2$ ή $q_1=-q_2$, λόγω συμμετρίας της πυκνότητας της f. Η πρώτη λύση απορρίπτεται γιατί πρέπει $q_1<q_2$ ώστε:

$$P_\theta(q_1<Q<q_2)=1-\alpha\neq 0.$$

Τελικά,
$$\left.\begin{aligned} &q_1=-q_2\\ &\Phi(q_2)-\Phi(q_1)=1-\alpha \end{aligned}\right\}$$

$\Rightarrow q_2=-q_1=z_{\alpha/2}$. Συνολικά, το ζευγάρι που ελαχιστοποιεί το δ.ε. είναι:
$(q_1,q_2)=(-z_{\alpha/2},z_{\alpha/2})$ επειδή $q_1<q_2$.
Αντικαθιστώντας τα (q_1,q_2) στο δ.ε.:

$$\left(\overline{X}-z_{\alpha/2}\frac{\sigma}{\sqrt{n}},\ \overline{X}+z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right).\ \boxtimes$$

<u>Ασκήσεις</u> <u>#1</u> Έστω $X_1,\dots,X_n$ τ.δ. από πληθυσμό με πυκνότητα:

$f(x;\theta)=\dfrac{\theta}{x^2}$, $\theta\le x$, $\theta>0$. Εάν $Q=\dfrac{\theta}{X_{(1)}}$ όπου $X_{(1)}$ είναι το ελάχιστο διατεταγμένο στατιστικό του δείγματος, (i) επιβεβαιώστε ότι η τ.μ. Q είναι αντιστρεπτή,

(ii) με τη βοήθεια της Q κατασκευάστε ένα $100(1-\alpha)\%$ δ.ε. για το θ,

(iii) Προσδιορίστε το βέλτιστο διάστημα εμπιστοσύνης.

<u>Λύση</u> (i) Για να είναι η Q αντιστρεπτή, αρχικά προσδιορίζουμε την κατανομή του $T = X_{(1)}$ από τον τύπο των διατεταγμένων στατιστικών για $r=1$.

$$f_T(t) = n\left(F_X(t)\right)^0 \left(1-F_X(t)\right)^{n-1} f_X(t)$$

και $F_X(t) = \displaystyle\int_\theta^t \frac{\theta}{x^2}\,dx = \left[-\frac{\theta}{x}\right]_\theta^t$

$= 1 - \theta/t$, $t \geq \theta$. Άρα,

$$f_T(t) = n\left(1 - 1 + \frac{\theta}{t}\right)^{n-1} \frac{\theta}{t^2}$$

$$= n\,\frac{\theta^{n-1}\theta}{t^{n-1}t^2} = \frac{n\theta^n}{t^{n+1}},\ t \geq \theta.$$

Θέτουμε: $q = \dfrac{\theta}{t} \Rightarrow t = \theta/q \geq \theta$

$\Rightarrow \dfrac{1}{q} \geq 1 \Rightarrow q \leq 1$ άρα $0 \leq q \leq 1$, το στήριγμα της Q. Επίσης, $\dfrac{dt}{dq} = -\dfrac{\theta}{q^2}$.

Άρα, $f_Q(q) = \frac{n\theta^n}{\left(\frac{\theta}{q}\right)^{n+1}} \cdot \left|-\frac{\theta}{q^2}\right|$

$= \frac{n\theta^{n+1}}{\theta^{n+1}} \frac{q^{n+1}}{q^2} = nq^{n-1}, q \in [0,1],$

άρα Q αντιστρεπτή.

(ii) Δ.ε. $P(q_1 < Q < q_2) = 1-\alpha$

$\Rightarrow P\left(q_1 < \frac{\theta}{T} < q_2\right) = 1-\alpha \xRightarrow{T>0}$

$\Rightarrow P(Tq_1 < \theta < Tq_2) = 1-\alpha$

$\Rightarrow P(X_{(1)}q_1 < \theta < X_{(1)}q_2) = 1-\alpha.$

Άρα, το δ.ε. για το θ είναι της μορφής

$(q_1 X_{(1)}, q_2 X_{(1)})$ με q_1, q_2:

$P(q_1 < Q < q_2) = 1-\alpha \Rightarrow \int_{q_1}^{q_2} f_Q(q)\,dq$

$= 1-\alpha \Rightarrow \int_{q_1}^{q_2} nq^{n-1}\,dq = 1-\alpha$

$\Rightarrow [q^n]_{q_1}^{q_2} = 1-\alpha \Rightarrow q_2^n - q_1^n = 1-\alpha \; (1)$

με $q_1, q_2 \in (0,1)$ με $q_1 < q_2$ που ικανοποιούν την (1). π.χ. αν $q_1 = 1/2$ και $1-\alpha = 0.95$, $q_2^n = \left(\frac{1}{2}\right)^n + 1 - \alpha \Rightarrow$

$\Rightarrow q_2 = \left(0.5^n + 1 - \alpha\right)^{1/n}$ και το δ.ε. είναι: $\left(\frac{1}{2} X_{(1)}, \left(0.5^n + 1 - \alpha\right)^{1/n} X_{(1)}\right)$.

Αν έχουμε $n=40$, τότε το 95% δ.ε. θα είναι το: $(0.5 X_{(1)}, 0.9987 X_{(1)})$ και το $X_{(1)}$ θα είναι η μικρότερη μέτρηση του δείγματος.

(iii) Για το βέλτιστο δ.ε. αναζητούμε το ζεύγος (q_1, q_2) που ικανοποιεί τις παραπάνω συνθήκες και ελαχιστοποιεί το μήκος του διαστήματος. Το μήκος είναι $X_{(1)}(q_2 - q_1)$. Ζητάμε ελαχιστοποίηση του $X_{(1)}(q_2 - q_1)$ ή ισοδύναμα του $\ell(q_1, q_2) = q_2 - q_1$ δηλαδή

$$\min_{q_1, q_2} \{q_2 - q_1\} \text{ με συνθήκη } q_2^n - q_1^n = 1 - \alpha.$$

Είναι: $\dfrac{\partial \ell(q_1, q_2)}{\partial q_1} = \dfrac{\partial q_2}{\partial q_1} - 1$ και παραγωγίζοντας ως προς q_1:

$$\frac{\partial}{\partial q_1}\left[q_2^n - q_1^n\right] = 0 \text{ ή } n q_2^{n-1}\frac{\partial q_2}{\partial q_1} - n q_1^{n-1} = 0$$

$$\Rightarrow \frac{\partial q_2}{\partial q_1} = \frac{q_1^{n-1}}{q_2^{n-1}}.$$ Αντικαθιστώντας

στην $\frac{\partial \ell(q_1,q_2)}{\partial q_1}$ προκύπτει:

$$\frac{\partial \ell(q_1,q_2)}{\partial q_1} = \frac{q_1^{n-1}}{q_2^{n-1}} - 1$$ που μηδενίζεται

για $q_1^{n-1} = q_2^{n-1}$ ή $q_1 = q_2$ άτοπο, γιατί $q_1 < q_2$ ώστε να ορίζεται διάστημα.

Άρα η συνάρτηση $\ell(q_1,q_2)$ είναι γνήσια μονότονη ως προς q_1. Επειδή

$$0 < q_1 < q_2 \Rightarrow q_1^{n-1} < q_2^{n-1} \Rightarrow \frac{q_1^{n-1}}{q_2^{n-1}} < 1$$

άρα η $\ell(q_1,q_2)$ είναι γνήσια φθίνουσα ως προς q_1 και λαμβάνει ελάχιστο στην μέγιστη τιμή της q_1, δηλ. $q_1 = 1$.

Για $q_1 = 1$ η (1) δίνει $q_2 = (2-\alpha)^{\frac{1}{n}}$ και άρα το ελάχιστο μήκος του δ.ε. επιτυγχάνεται για $(q_1, q_2) = \left(1, (2-\alpha)^{\frac{1}{n}}\right)$.

Το ελάχιστο δ.ε. με εμπιστοσύνη $100(1-\alpha)\%$ είναι:

$$\left(q_1 X_{(1)}, q_2 X_{(1)}\right) = \left(X_{(1)}, X_{(1)}(2-\alpha)^{1/n}\right).$$

#2 Δίνεται τ.δ. από πληθυσμό με κατανομή Gamma$(\alpha=1, \beta=1/\theta)$.

(i) Να δειχθεί ότι $Q = \frac{2}{\theta}\sum_{i=1}^{n} X_i$ είναι μία αντιστρεπτή ποσότητα.

(ii) Να βρεθεί ένα $100(1-\alpha)\%$ δ.ε. για το θ.

Λύση (i) $Y = \sum_{i=1}^{n} X_i \sim \text{Gamma}\left(n, \frac{1}{\theta}\right)$

$\Rightarrow Q = \frac{2}{\theta} Y$. Θέτουμε $q = \frac{2}{\theta} y$

$\Rightarrow y = \frac{\theta}{2} q \Rightarrow \frac{\partial y}{\partial q} = \frac{\theta}{2}, y > 0, q > 0.$

$$f_Q(q) = \frac{\left(\frac{\theta q}{2}\right)^{n-1} e^{-\frac{\theta q}{2\theta}}}{\Gamma(n)\,\theta^n} \cdot \frac{\theta}{2}$$

$$= \frac{\theta^{n-1} q^{n-1} e^{-q/2}}{\Gamma(n)\,\theta^n} \, \frac{\theta}{2^n} = \frac{q^{n-1} e^{-q/2}}{\Gamma(n)\, 2^n},$$

$q>0$. Η παραπάνω κατανομή είναι ανεξάρτητη της παραμέτρου, άρα είναι αντιστρεπτή. Επιπλέον, $Q \sim Gamma(n, \frac{1}{2}) \equiv \chi^2_{2n}$.

Κατασκευή του δ.ε.

$P(q_1 < Q < q_2) = 1-\alpha \Rightarrow P\left(q_1 < \frac{2n}{\theta}\bar{X} < q_2\right)$

$= 1-\alpha \Rightarrow P\left(\frac{1}{q_2} < \frac{\theta}{2n\bar{X}} < \frac{1}{q_1}\right) = 1-\alpha$

$\Rightarrow P\left(\frac{2n\bar{X}}{q_2} < \theta < \frac{2n\bar{X}}{q_1}\right) = 1-\alpha.$

Άρα, το δ.ε. είναι το: $\left(\frac{2n\bar{X}}{q_2}, \frac{2n\bar{X}}{q_1}\right)$.

Προσδιορισμός q_1, q_2:

$$\int_{q_1}^{q_2} f_Q(q)\,dq = 1-\alpha$$

$$\Rightarrow F_{\chi^2_{2n}}(q_2) - F_{\chi^2_{2n}}(q_1) = 1-\alpha \quad (*)$$

Άρα, το δ.ε. είναι: $\left(\dfrac{2n\bar{X}}{q_2}, \dfrac{2n\bar{X}}{q_1} \right)$

με $q_1, q_2 > 0$ τ.ω. $q_1 < q_2$ και να ικανοποιούν την $(*)$.

(ii) Αν ενδιαφερόμαστε για το <u>βέλτιστο</u> δ.ε. δηλαδή αυτό με το <u>ελάχιστο</u> μήκος ορίζουμε:

Μήκος: $\ell = 2n\bar{X}\left(\dfrac{1}{q_1} - \dfrac{1}{q_2}\right)$.

Ζητούμε

$$\min_{q_1, q_2} \left\{ 2n\bar{X}\left(\frac{1}{q_1} - \frac{1}{q_2}\right) \right\}$$ ή

ισοδύναμα, $\min\limits_{q_1, q_2} \left\{ \dfrac{1}{q_1} - \dfrac{1}{q_2} \right\}$.

Θέτουμε:

$$\Phi(q_1, q_2, \lambda) = \left(\frac{1}{q_1} - \frac{1}{q_2}\right) - \lambda\left(F_{\chi^2_{2n}}(q_2) - F_{\chi^2_{2n}}(q_1) - 1 + \alpha\right),$$ λ: πολλαπλασιαστής

Lagrange. Οι μερικές παράγωγοι της Φ είναι:

$$\frac{\partial \Phi}{\partial q_1} = -\frac{1}{q_1^2} + \lambda f_{\chi^2_{2n}}(q_1) = 0$$

$$\frac{\partial \Phi}{\partial q_2} = \frac{1}{q_2^2} - \lambda f_{\chi^2_{2n}}(q_2) = 0$$

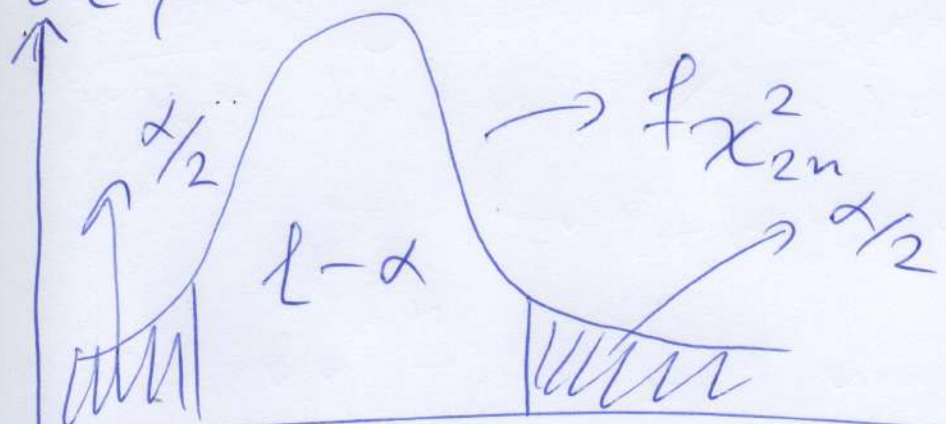
$$\Rightarrow q_1^2 f_{\chi^2_{2n}}(q_1) = q_2^2 f_{\chi^2_{2n}}(q_2)$$

$$\frac{\partial \Phi}{\partial \lambda} = 0 \Rightarrow F_{\chi^2_{2n}}(q_2) - F_{\chi^2_{2n}}(q_1) = 1-\alpha.$$

Οι δύο πρώτες εξισώσεις δίνουν την εξίσωση $q_1^2 f_{\chi^2_{2n}}(q_1) = q_2^2 f_{\chi^2_{2n}}(q_2)$ που λύνεται με αριθμητική ολοκλήρωση, λόγω της πολυπλοκότητας της $f_{\chi^2_{2n}}$. Εναλλακτικά, ένα σύνηθες δ.ε. στην πράξη, είναι εκείνο με την

ιδιότητα "των ίσων ουρών"
αντί του "ελάχιστου μήκους"! Διάστημα ίσων ουρών είναι το δ.ε. για το οποίο: $P(q_1 < Q < q_2) = 1-\alpha$ και $P(Q < q_1) = \frac{\alpha}{2}$ και

$P(Q > q_2) = \alpha/2$. Δηλαδή, η πιθανότητα, η παράμετρος Θ να μη περιέχεται στο δ.ε. <u>μοιράζεται</u> δεξιά και αριστερά του δ.ε.



Υπολογισμός q_1, q_2:

$$P(Q < q_1) = \frac{\alpha}{2} \Rightarrow \int_{-\infty}^{q_1} f(q)\,dq = \frac{\alpha}{2}$$

και

$$P(Q > q_2) = \frac{\alpha}{2} \Rightarrow \int_{q_2}^{\infty} f(q)\,dq = \frac{\alpha}{2}$$

$$\left.\begin{aligned} F_{\chi^2_{2n}}(q_1) &= \frac{\alpha}{2} \\ 1 - F_{\chi^2_{2n}}(q_2) &= \frac{\alpha}{2} \end{aligned}\right\} \Rightarrow \begin{aligned} q_1 &= \chi^2_{2n,1-\alpha/2} \\ q_2 &= \chi^2_{2n,\alpha/2} \end{aligned}$$

(χρησιμοποιούνται τα άνω εκατοστιαία σημεία της χ^2_{2n}).

Άρα, το δ.ε. <u>ίσων ουρών</u> με <u>βαθμό εμπιστοσύνης $100(1-\alpha)\%$</u> είναι:

$$\left(\frac{2n\bar{X}}{\chi^2_{2n,\alpha/2}}, \frac{2n\bar{X}}{\chi^2_{2n,1-\alpha/2}} \right).$$

<u>#3</u> Έστω $X_1, \dots, X_n$ τ.δ. από $f(x;\theta) = e^{-(x-\theta)} I_{[\theta,\infty)}(x)$, $-\infty < \theta < +\infty$. Να δειχθεί ότι η $Q = 2n[X_{(1)} - \theta]$ είναι αντιστρεπτή ποσότητα.

<u>Λύση</u> Αν $Y = X_{(1)}$ τότε μπορούμε να

γράφουμε ότι:

$$f_Y(y) = n e^{-n(y-\theta)} I_{[\theta,\infty)}(x),$$

$-\infty < \theta < +\infty$.

Θέτουμε $q = 2n[X_{(1)} - \theta]$.

$$y = \frac{q+2n\theta}{2n} \Rightarrow \frac{\partial y}{\partial q} = \frac{1}{2n} \text{ και}$$

$$\frac{q+2n\theta}{2n} > \theta \Rightarrow q + 2n\theta > 2n\theta$$

$$\Rightarrow \underline{q > 0}$$

$$f_Q(q) = n e^{-n\left(\frac{q+2n\theta}{2n} - \theta\right)} \cdot \frac{1}{2n}$$

$$= n \cdot e^{-n\left(\frac{q}{2n}\right)} \cdot \frac{1}{2n} = e^{-n\left(\frac{q}{2n}\right)} \cdot \frac{1}{2}, \quad q > 0$$

Άρα $Q \sim Exp\left(\frac{1}{2}\right)$.

$$P(q_1 < Q < q_2) = P\left(q_1 < 2n\left(X_{(1)} - \theta\right) < q_2\right)$$

$$= P\left(q_1 - 2nX_{(1)} < -2n\theta < q_2 - 2nX_{(1)}\right)$$

$$= P\left(X_{(1)} - \frac{q_2}{2n} < \theta < X_{(1)} - \frac{q_1}{2n}\right).$$

<u>Μήκος</u> $\ell = X_{(1)} - \frac{q_1}{2n} - X_{(1)} + \frac{q_2}{2n}$

$\ell = \frac{1}{2n}(q_2 - q_1)$. Ελαχιστοποιώντας την παραπάνω παράσταση ή ισοδύναμα την $g(q_1, q_2) = q_2 - q_1$ με τη συνθήκη: $\int_{q_1}^{q_2} \frac{1}{2} e^{-q/2} dq = 1 - \alpha$

$\Rightarrow \left[e^{-q/2} \right]_{q_1}^{q_2} = 1 - \alpha \Rightarrow e^{-\frac{q_1}{2}} - e^{-\frac{q_2}{2}} = 1 - \alpha$ (2)

έχουμε ότι: $\frac{\partial g}{\partial q_1} = \frac{\partial q_2}{\partial q_1} - 1$ και

παραγωγίζοντας την (2), έχουμε:

$-\frac{1}{2} e^{-q_1/2} + \frac{1}{2} \frac{\partial q_2}{\partial q_1} e^{-q_2/2} = 0$

$\Rightarrow \frac{\partial q_2}{\partial q_1} = \frac{e^{-q_1/2}}{e^{-q_2/2}}$. Άρα $\frac{\partial g}{\partial q_1} = \frac{e^{-\frac{q_1}{2}}}{e^{-q_2/2}} - 1$

Επειδή $q_1 < q_2$ θα είναι $\frac{\partial g}{\partial q_1} > 0$, δηλ. η συνάρτηση g είναι γνησίως αύξουσα ως προς q_1 και το ελάχιστο λαμβάνεται στο κάτω άκρο των τιμών του q_1, εφόσον λαμβάνεται. Στην προκειμένη περίπτωση, $q_1=0$ (η $Q\sim Exp(1/2)$). Μέσω της (2) θα είναι:

$$1-e^{-q_2/2}=1-\alpha \Rightarrow q_2=-2\ln\alpha$$

Άρα, το βέλτιστο $100(1-\alpha)\%$ δ.ε. είναι το $\left(X_{(1)}+\dfrac{\ln\alpha}{n}, X_{(1)}\right)$. ☒

<u>#4</u> Συνέχεια του Παραδείγματος 2. Ενδιαφερόμαστε για το βέλτιστο δ.ε. $100(1-\alpha)\%$ της παραμέτρου μ από έναν κανονικό πληθυσμό με σ^2 άγνωστο. Η ποσότητα $Q=\dfrac{\bar{X}-\mu}{S/\sqrt{n}}\sim t_{n-1}$ είναι αντιστρεπτή ποσότητα. Επειδή η κατανομή t_n-student είναι <u>συμμετρική</u> γύρω <u>από το μηδέν</u>, όπως και η <u>τυπική</u> <u>Κανονική Κατανομή</u>, βρίσκουμε $q_1=-t_{n-1;\alpha/2}=t_{n-1;\alpha/2}$.

Το βέλτιστο δ.ε. είναι:

$$\left(\bar{X} - t_{n-1,\alpha/2}\frac{S}{\sqrt{n}},\ \bar{X} + t_{n-1;\alpha/2}\frac{S}{\sqrt{n}}\right).$$

4.3 Ύπαρξη αντιστρεπτής ποσότητας ☒

Η μέθοδος της αντιστρεπτής ποσότητας βασίζεται σε μία ποσότητα, ανεξάρτητη του θ για την παράμετρο της οποίας αναζητείται το δ.ε. Για την εύρεση της αντιστρεπτής ποσότητας στηριζόμαστε στα εξής:

(i) Προσδιορισμός ενός εκτιμητή T για την παράμετρο θ. Συνήθως επαρκής ή ΕΜΠ.

(ii) Υπολογισμός της κατανομής της τ.μ. T. Η παράμετρος θ θα είναι παράμετρος της κατανομής της T.

(iii) Αναγνώριση του ρόλου της παραμέτρου θ για την κατανομή της T.

(iv) Μετασχηματισμός τ.μ. για την T με σκοπό η νέα τ.μ. να μην εξαρτάται από την θ.

Εναλλακτικά, μία μέθοδος εύρεσης αντιστρεπτής ποσότητας, εύχρηστη, αλλά με περιορισμό μόνο για συνεχείς τ.μ. δίνεται από την παρακάτω πρόταση.

<u>Πρόταση</u> Έστω $X_1, \dots, X_n$ τ.δ. από πληθυσμό με συνεχή ως προς X α.σ.κ. $F(x;\theta)$. Αντιστρεπτή ποσότητα υπάρχει πάντα για τέτοιο πληθυσμό.

Θέτουμε $Y = F(X;\theta)$, $y \in [0,1]$.

$$F_Y(y) = P(Y \le y) = P(F_X(X;\theta) \le y)$$

Η F για μία συνεχής τ.μ. είναι αύξουσα συνάρτηση άρα υπάρχει η αντίστροφη συνάρτηση. Άρα,

$$P(F_X(X;\theta) \le y) = P(X \le F_X^{-1}(y))$$

$$= F_X(F_X^{-1}(y)) = y$$

Άρα $F_Y(y) = y \Rightarrow f_Y(y) = 1,\ y \in [0,1]$, Άρα, $Y \sim \mathcal{U}[0,1]$, κατανομή ανεξάρτητη της θ. Επίσης, αν $Z = -\ln Y = -\ln F(X;\theta)$ τότε $Z \sim Exp(1)$ επίσης ανεξάρτητη της θ. Αν $X_1, X_2, \ldots, X_n$ τ.δ. από έναν πληθυσμό τότε $-\ln F(X_i;\theta) \sim Exp(1)$ $\forall i$. Λόγω ανεξαρτησίας των X_i, έχουμε:

$$Q = -\sum_{i=1}^{n} \ln F(X_i;\theta) \sim Gamma(n,1).$$

Άρα Q αντιστρεπτή και ισοδύναμα, $Q' = 2Q \sim \chi^2_{2n}$ που είναι επίσης αντιστρεπτή ποσότητα. Η αντιστρεπτή ποσότητα Q ή Q' μπορεί να <u>μην</u> οδηγεί <u>πάντα</u> σε δ.ε. γιατί μπορεί να μην αντιστρέφεται ως προς θ. Αν η $F(X;\theta)$ είναι εκτός από συνεχής ως προς x και <u>μονότονη</u> ως προς θ, τότε είναι εφικτή η εύρεση Δ.Ε. $\boxtimes$

Ας δούμε ένα παράδειγμα.

Παράδειγμα $X_1, X_2, \ldots, X_n$ τ.δ. από $Beta(\theta, 1)$.

$$f(x;\theta) = \begin{cases} \theta x^{\theta-1}, & 0<x<1 \\ 0, & \text{αλλού} \end{cases}$$

Άρα $$F(x;\theta) = \begin{cases} 0, & x<0 \\ x^{\theta}, & 0<x<1 \\ 1, & x\geq 1 \end{cases}$$

$$Q = -\sum_{i=1}^{n} \ln F(X_i;\theta) = -\sum_{i=1}^{n} \ln X_i^{\theta}$$

$\sim Gamma(n,1)$

ή $$Q = -2\sum_{i=1}^{n} \ln X_i^{\theta} \sim \chi^2_{2n}.$$

$$P\left(q_1 < -2\sum_{i=1}^{n} \ln X_i^{\theta} < q_2\right) = 1-\alpha$$

$$\Rightarrow P\left(q_1 < -2\theta\sum_{i=1}^{n} \ln X_i < q_2\right) = 1-\alpha$$

$$\Rightarrow P\left(\frac{-q_1}{2\sum_{i=1}^{n} \ln X_i} < \theta < \frac{-q_2}{2\sum_{i=1}^{n} \ln X_i}\right) = 1-\alpha,$$

με $q_1 = \chi^2_{2n,1-\alpha/2}$, $q_2 = \chi^2_{2n,\alpha/2}$ λαμβάνουμε ένα $100(1-\alpha)\%$ δ.ε. ίσων ουρών. ☒

4.4 Ασυμπτωτικά δ.ε.

Μέχρι τώρα παρουσιάστηκαν ακριβείς μέθοδοι για την κατασκευή δ.ε. για την παράμετρο θ, ή την παραμετρική συνάρτηση $g(\theta)$. Στην ακριβή περίπτωση, η πιθανότητα:

$$P_\theta\{L(\underset{\sim}{X}) < g(\theta) < U(\underset{\sim}{X})\} = 1-\alpha$$

του ορισμού για συνεχή τ.μ. είναι εφικτό να ισούται με $1-\alpha$ <u>ακριβώς</u> δηλ. $P_\theta\{L(\underset{\sim}{X}) < g(\theta) < U(\underset{\sim}{X})\} = 1-\alpha$. Το δ.ε. λέγεται ακριβές. Δεν είναι πάντα εφικτό, οι κατανομές των εκτιμητών να προσδιορίζονται με ακρίβεια, αλλά προσεγγιστικά. Κατά συνέπεια, η πιθανότητα δεν είναι ίση με $1-\alpha$, αλλά κατά προσέγγιση, δηλαδή:

$$P_\theta\{L(\underset{\sim}{X}) < g(\theta) < U(\underset{\sim}{X})\} \simeq 1-\alpha$$

Τα παραπάνω δ.ε. καλούνται ασυμπτωτικά δ.ε. κυρίως διότι οι προσεγγιστικές κατανομές των εκτιμητών T προκύπτουν για μεγάλα δείγματα δηλ. θεωρητικά για $n \to \infty$. Η πιο συνήθης μέθοδος για προσεγγιστική εύρεση της κατανομής μιας σ.σ. είναι η ασυμπτωτική κατανομή του ΕΜΠ. Με τη βοήθεια της ασυμπτωτικής κατανομής, υπολογίζεται η παραπάνω πιθανότητα, προσεγγιστικά.

Έχουμε δει ότι, κάτω από τις συνθήκες ομαλότητας για τον ΕΜΠ, $\hat{\theta}_n$ ισχύει: $\sqrt{n}\left(\hat{\theta}_n - \theta\right) \xrightarrow{K} N\left(0, \frac{1}{I_X(\theta)}\right)$.

Μέσω της Κανονικής Κατανομής:

$$P\left(-z_{\alpha/2} < \frac{\sqrt{n}\left(\hat{\theta}_n - \theta\right)}{\sqrt{1/I_X(\theta)}} < z_{\alpha/2}\right) \simeq 1-\alpha$$

και λύνοντας, ως προς θ, θα προκύψει το $(1-\alpha)100\%$ δ.ε. για το θ. Εκτός από τον αριθμητή του κλάσματος:

$$\frac{\sqrt{n}\,(\hat{\theta}_n-\theta)}{\sqrt{1/I_X(\theta)}}$$

και το μέτρο πληροφορίας $I_X(\theta)$ είναι επίσης συνάρτηση του θ, με αποτέλεσμα η παραπάνω επίλυση ως προς θ, να είναι ορισμένες φορές δύσκολη ή αδύνατη. Τότε, αντί $I_X(\theta)$ χρησιμοποιούμε τον ΕΜΠ της ποσότητας αυτής, $\widehat{I_X(\theta)}$. ⊠

<u>Παράδειγμα 1</u>: Έστω $X_1, \ldots, X_n$ τ.δ. από την Ομοιόμορφη κατανομή στο διάστημα $(0,\theta)$. Βρείτε το ασυμπτωτικό $100(1-\alpha)\%$ δ.ε. για την παράμετρο θ.

<u>Λύση</u> Έχουμε δει ότι:
ΕΜΠ του θ: $\hat{\theta} = X_{(n)}$. Επίσης
$I_X(\theta) = 1/\theta^2$. Για την ασυμπτωτική
κατανομή του ΕΜΠ είναι:

$$\hat{\theta} \sim N\left(\theta, \frac{1}{n I_X(\theta)}\right) \text{ ή } \hat{\theta} \sim N\left(\theta, \frac{\theta^2}{n}\right).$$

$$\frac{\hat{\theta}-\theta}{\sqrt{\theta^2/n}} \sim N(0,1).$$ Για το $100(1-\alpha)\%$

δ.ε. θα είναι: $$P\left(-z_{\alpha/2} \leq \frac{\hat{\theta}-\theta}{\sqrt{\theta^2/n}} \leq z_{\alpha/2}\right) = 1-\alpha$$

$$\Rightarrow P\left(-z_{\alpha/2} \leq \frac{\hat{\theta}}{\theta/\sqrt{n}} - \frac{\theta}{\theta/\sqrt{n}} \leq z_{\alpha/2}\right) = 1-\alpha$$

$$\Rightarrow P\left(-z_{\alpha/2} \leq \frac{\hat{\theta}}{\theta/\sqrt{n}} - \sqrt{n} \leq z_{\alpha/2}\right) = 1-\alpha$$

$$\Rightarrow P\left(\sqrt{n} - z_{\alpha/2} \leq \frac{\sqrt{n}\,\hat{\theta}}{\theta} \leq \sqrt{n} + z_{\alpha/2}\right) = 1-\alpha$$

$$\Rightarrow P\left(\frac{1}{\sqrt{n}+z_{\alpha/2}} \leq \frac{\theta}{\sqrt{n}\,\hat{\theta}} \leq \frac{1}{\sqrt{n}-z_{\alpha/2}}\right) = 1-\alpha$$

$$\Rightarrow P\left(\frac{\sqrt{n}\,\hat{\theta}}{\sqrt{n}+z_{\alpha/2}} \le \theta \le \frac{\sqrt{n}\cdot\hat{\theta}}{\sqrt{n}-z_{\alpha/2}}\right)=1-\alpha$$

Άρα το $100(1-\alpha)\%$ δ.ε. για το θ είναι το:

$$\left(\frac{\sqrt{n}\,\hat{\theta}}{\sqrt{n}+z_{\alpha/2}}, \frac{\sqrt{n}\,\hat{\theta}}{\sqrt{n}-z_{\alpha/2}}\right).$$

Αντικαθιστώντας $\hat{\theta}=X_{(n)}$, θα είναι:

$$\left(\frac{\sqrt{n}\,X_{(n)}}{\sqrt{n}+z_{\alpha/2}}, \frac{\sqrt{n}\,X_{(n)}}{\sqrt{n}-z_{\alpha/2}}\right). \boxtimes$$

<u>Παράδειγμα 2</u> Έστω $X_1,\dots,X_n$ τ.δ. ένα τ.δ. από κατανομή Poisson με παράμετρο λ. Βρείτε το ασυμπτωτικό $100(1-\alpha)\%$ δ.ε. για την παράμετρο λ.

<u>Λύση</u> Θέτουμε $\lambda=\theta$. Έχουμε δει ότι $\hat{\lambda}=\bar{X}=\hat{\theta}$ και $I_X(\theta)=\frac{1}{\theta}$.

Άρα, η ασυμπτωτική κατανομή του ΕΜΠ είναι $\hat{\theta} \sim N\left(\theta, \frac{1}{\theta}\right)$ ή

$$\frac{\hat{\theta}-\theta}{\sqrt{\theta/n}} \sim N(0,1).$$ Για το δ.ε

$$P\left(-z_{\alpha/2} \le \frac{\hat{\theta}-\theta}{\sqrt{\theta/n}} \le z_{\alpha/2}\right) = 1-\alpha$$

Επειδή ως προς θ δεν είναι εύκολο να λυθεί, αντικαθιστούμε στον παρονομαστή, το $I_X(\theta)$ με τον ΕΜΠ του θ, δηλαδή $\widehat{I_X(\theta)} = 1/\hat{\theta}$.

Άρα,

$$P\left(-z_{\alpha/2} \le \frac{\hat{\theta}-\theta}{\sqrt{\hat{\theta}/n}} \le z_{\alpha/2}\right) = 1-\alpha$$

$$\Rightarrow P\left(-z_{\alpha/2}\sqrt{\frac{\hat{\theta}}{n}} \le \hat{\theta}-\theta \le z_{\alpha/2}\sqrt{\frac{\hat{\theta}}{n}}\right) = 1-\alpha$$

$$\Rightarrow P\left(-\hat{\theta}-z_{\alpha/2}\sqrt{\frac{\hat{\theta}}{n}} \le -\theta \le -\hat{\theta}+z_{\alpha/2}\sqrt{\frac{\hat{\theta}}{n}}\right) = 1-\alpha$$

$$\Rightarrow P\left(\hat{\theta}-z_{\alpha/2}\sqrt{\frac{\hat{\theta}}{n}} \le \theta \le \hat{\theta}+z_{\alpha/2}\sqrt{\frac{\hat{\theta}}{n}}\right) = 1-\alpha$$

Άρα, το ασυμπτωτικό $100(1-\alpha)\%$ δ.ε. για το θ είναι:

$$\left(\hat{\theta}-z_{\alpha/2}\sqrt{\frac{\hat{\theta}}{n}}\,,\ \hat{\theta}+z_{\alpha/2}\sqrt{\frac{\hat{\theta}}{n}}\right)$$

όπου $\hat{\theta}=\overline{X}$. ⊠

<u>Άσκηση</u> Έστω $f(x;\theta)=\theta x^{\theta-1}$, $0\le x\le 1$ και τ.δ. μεγέθους n από την $f(x;\theta)$.

(i) Υπολογίστε την ΕΜΠ του θ, $\hat{\theta}$.

(ii) Επιβεβαιώστε ότι $Q=\dfrac{2n\theta}{\hat{\theta}}$ είναι αντιστρεπτή ποσότητα και με τη βοήθειά της, να υπολογίσετε το $100(1-\alpha)\%$ δ.ε. <u>ίσων ουρών</u> για το θ.

(iii) Με τη βοήθεια του $\hat{\theta}$ δώστε το <u>ασυμπτωτικό</u> $100(1-\alpha)\%$ δ.ε. για το θ.

(iv) Εφαρμογή: Για δείγμα μεγέθους $n=40$, από την παραπάνω κατανομή, οι τιμές του δείγματος είναι:

x_i: 0.47 0.97 0.52 0.72 0.85
0.74 0.76 1 0.79 0.53 0.39
0.95 0.99 0.67 0.63 0.78 0.74
0.99 0.61 0.86 0.86 0.59 0.80
0.72 0.30 0.88 0.65 0.97 0.90
0.55 0.80 0.87 0.69 0.60 0.43
0.88 0.76 0.32 0.85 0.91

Δώστε το 95% δ.ε. των ερωτημάτων (ii), (iii) για το παραπάνω δείγμα.

<u>Λύση</u> (i) $\ell(\theta) = \ln L(\theta)$

$$= \ln \prod_{i=1}^{n} \theta x_i^{\theta-1} = \ln\left(\theta^n \left(\prod_{i=1}^{n} x_i\right)^{\theta-1}\right)$$

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

$$\frac{\partial \ell(\theta)}{\partial \theta} = \frac{n}{\theta} + \sum_{i=1}^{n} \ln(x_i) = 0$$

$$\Rightarrow \theta = -\frac{n}{\sum_{i=1}^{n} \ln(x_i)} > 0 \text{ γιατί } \ln(x_i) < 0$$

$$\frac{\partial^2 \ell}{\partial \theta^2} = -\frac{n}{\theta^2} < 0 \Rightarrow \hat{\theta} = -\frac{n}{\sum_{i=1}^{n} \ln(x_i)}$$

(ii) $Q = \dfrac{2n\theta}{-n \Big/ \sum_{i=1}^{n} \ln(x_i)} = -2\theta \sum_{i=1}^{n} \ln(x_i)$

Όμως $-\ln(X) \sim Exp(\theta)$

$-\ln(X_i) \sim Exp(\theta)\ \forall i$

$\Rightarrow -\sum_{i=1}^{n} \ln(X_i) \sim Gamma(n, \theta)$.

Αν $W = -\sum_{i=1}^{n} \ln(X_i)$ τότε $Q = 2W\theta$

$\sim Gamma(n, \frac{1}{2})$. Άρα $Q \sim Gamma(n, 1/2)$

$\equiv \chi^2_{2n}$. Συνεπώς, το δ.ε. ίσων ουρών

είναι: $P\left(\chi^2_{2n, 1-\alpha/2} < Q < \chi^2_{2n, \alpha/2}\right)$

$= 1-\alpha \Rightarrow P\left(\chi^2_{2n, 1-\alpha/2} < 2\theta W < \chi^2_{2n, \alpha/2}\right)$

$= 1-\alpha$

$\Rightarrow P\left(\dfrac{1}{2W}\chi^2_{2n, 1-\alpha/2} < \theta < \dfrac{1}{2W}\chi^2_{2n, \alpha/2}\right)$

$= 1-\alpha$.

Άρα, το $100(1-\alpha)\%$ δ.ε. είναι:

$$\left(\frac{1}{-2\sum_{i=1}^{n}\ln(X_i)}\chi^2_{2n,1-\alpha/2},\ \frac{1}{-2\sum_{i=1}^{n}\ln(X_i)}\chi^2_{2n,\alpha/2}\right).$$

(iii) $\hat{\theta} = -\frac{n}{\sum_{i=1}^{n}\ln(X_i)}$, $\hat{\theta} \sim N\left(\theta, \frac{1}{n I_X(\theta)}\right)$

όπου $I_X(\theta) = -E\left(\frac{\partial^2}{\partial\theta}\ln f(X;\theta)\right)$

$= \ldots = 1/\theta^2$.

Άρα $\hat{\theta} \sim N(\theta, \theta^2/n)$ οπότε για το ασυμπτωτικό δ.ε. θα έχουμε:

$$\frac{\hat{\theta}-\theta}{\theta/\sqrt{n}} \sim N(0,1).$$

$$P\left(-z_{\alpha/2} < \frac{\hat{\theta}-\theta}{\theta/\sqrt{n}} < z_{\alpha/2}\right) = 1-\alpha$$

$$P\left(\sqrt{n} - z_{\alpha/2} < \frac{\sqrt{n}\,\hat{\theta}}{\theta} < \sqrt{n} + z_{\alpha/2}\right)$$

$= 1-\alpha$

$$\Rightarrow P\left(-\frac{n}{\sum_{i=1}^{n} \ln(X_i)\left(1+z_{\alpha/2}/\sqrt{n}\right)} < \theta\right.$$

$$\left. < -\frac{n}{\sum_{i=1}^{n} \ln(X_i)\left(1-z_{\alpha/2}/\sqrt{n}\right)}\right)$$

$= 1-\alpha$. Άρα, το προσεγγιστικό $100(1-\alpha)\%$ δ.ε. για το θ είναι:

$$\left(-\frac{n}{\sum_{i=1}^{n} \ln(X_i)\left(1+\frac{z_{\alpha/2}}{\sqrt{n}}\right)}, -\frac{n}{\sum_{i=1}^{n} \ln(X_i)\left(1-\frac{z_{\alpha/2}}{\sqrt{n}}\right)}\right)$$

(iv) Το δ.ε. ίσων ουρών για το (ii) είναι:

$$\left(\frac{1}{-2\sum_{i=1}^{n} \ln(X_i)}\chi^2_{2n,1-\alpha/2}, \frac{1}{-2\sum_{i=1}^{n} \ln(X_i)}\chi^2_{2n,\frac{\alpha}{2}}\right)$$

Για τις τιμές του δείγματος, $n=40$,

$$\chi^2_{2n,1-\alpha/2} = \chi^2_{80,0.975} = 57.153$$

$$\chi^2_{2n,\alpha/2} = \chi^2_{80,0.025} = 106.629$$

και $-\sum_{i=1}^{40} \ln(X_i) = 14.06$ βρίσκουμε:

$$\left(\frac{1}{2\times 14.06}\cdot 57.153, \frac{1}{2\times 14.06}\cdot 106.629\right)$$

$= (2.03, 3.79)$. Το ασυμπτωτικό δ.ε. του (iii) ήταν:

$$\left(\frac{40}{14.06\cdot(1+1.96/\sqrt{40})}, \frac{40}{14.06\cdot\left(1-\frac{1.96}{\sqrt{40}}\right)}\right)$$

$= (2.17, 4.12)$. ☒

Ζ