ΠΕΡΙΕΧΟΜΕΝΑ

3.4. Κριτηρια Αποδοσης Εκτιμητριων. σελ. 90.

Αμεροληψια, ασυμπτωτικη αμεροληψια, ακριβεια, συνεπεια, σχετικη αποδοση, ασυμπτωτικη σχετικη αποδοση εκτιμητριων. Πληροφορια κατα Fisher κατανομης και δειγματος, εντροπια, πληροφορια κατα Kullback-Leibler, ανισοτητα των Cramér-Frechét-Rao, αποδοση εκτιμητριας, ασυμπτωτικη κανονικοτητα, ασυμπτωτικη αποδοση, εκτιμητριες ΒΑΚ.

3.5. Περιοχες Εμπιστοσυνης Βασισμενες σε Εκτιμητριες. σελ. 113.

Στοχαστικο διαστημα εμπιστοσυνης, επιπεδο, πιθανοτητα καλυψης, συντελεστης εμπιστοσυνης, ακριβεια, αμεροληψια, συμμετρια διαστηματων εμπιστοσυνης. Ασυμπτωτικα διαστηματα εμπιστοσυνης. Κατασκευες διαστηματων εμπιστοσυνης.

3.6. Εκτιμητριες Ελαχιστων Τετραγωνων - Παλινδρομηση σελ. 121.

Γραμμικα μοντελα, απλη και πολλαπλη γραμμικη παλινδρομηση, κατασκευη εκτιμητριων ελαχιστων τετραγωνων, κανονικες εξισωσεις των ελαχιστων τετραγωνων, Θεωρημα Gauss-Markov, μοντελο αναλυσης διασπορας, διαστηματα εμπιστοσυνης.

4. Ελεγχοι Υποθεσεων σελ. 139 - 185

Στατιστικες υποθεσεις $H: \theta \in \Theta_0$ εναντι εναλλακτικων δυνατοτητων $K: \theta \in \Theta_1$, ελεγχοσυναρτηση, τυχαιοποιημενοι και μη ελεγχοι, στατιστικη ελεγχου, κρισιμη σταθερα, κρισιμη περιοχη, περιοχη αποδοχης, σφαλμα τυπου I, σφαλμα τυπου II, μετρα σφαλματων, συναρτησεις απωλειας και κινδυνου ελεγχων, ισχυς ελεγχου, μεγεθος ελεγχου, αμεροληψια ελεγχου, ομοιομορφα πλεον ισχυροι ελεγχοι - μεθοδος Neyman & Pearson, p-τιμη ελεγχου.

4.1. Ελεγχοι Neyman - Pearson. σελ. 150

Λημμα των Neyman-Pearson, κατασκευη πλεον ισχυρων ελεγχων μεγεθους α απλης υποθεσεως εναντι απλης εναλλακτικης, τυχαιοποιημενοι ελεγχοι για διακριτα μοντελα, μονοτονο πηλικο πιθανοφανειων (ΜΠΠ), ομοιομορφα πλεον ισχυροι (ΟΠΙ) ελεγχοι μεγεθους α για απλη ή μονοπλευρη συνθετο υποθεση εναντι συνθετου εναλλακτικης, ελεγχοι απλης εναντι αμφιπλευρου, ΟΠΙ αμεροληπτοι ελεγχοι, ΟΜΑ αμεροληπτα διαστηματα εμπιστοσυνης

4.2. Ελεγχοι πηλικου πιθανοφανειων (ΕΠΠ) σελ. 169

κατασκευη ΕΠΠ, ΕΠΠ οταν $\dim \Theta_0 < \dim \Theta_1$, βασικοι ελεγχοι για την κανονικη κατανομη – ελεγχος-t, ελεγχος-χ^2, δι-δειγματικος ελεγχος-t, ελεγχος-F – διαστηματα εμπιστοσυνης βασισμενα σε ΕΠΠ, ασυμπτωτικος ΕΠΠ, ασυμπτωτικος ΕΠΠ για την Πολυωνυμικη κατανομη, ελεγχος-χ^2 καλης εφαρμογης.

1. ΑΝΑΣΚΟΠΗΣΗ ΤΗΣ ΘΕΩΡΙΑΣ ΠΙΘΑΝΟΤΗΤΩΝ.

Η αβεβαιοτητα ενος στοχαστικου πειραματος $(\Omega,\mathcal{A})$, καθοριζεται στη δομη του χωρου με πιθανοτητα $(\Omega,\mathcal{A},P)$ του πειραματος, ως εξης:

(1.1) Ο δειγματικος χωρος Ω συνισταται απο ολα τα δυνατα αποτελεσματα ω του πειραματος, αντανακλα δε το ειδος της πληροφορησης που εχουμε ή ζητουμε απο το πειραμα.

(1.2) Η σ-αλγεβρα $\mathcal{A}$ των ενδεχομενων του πειραματος

(α) περιεχει τον Ω,

(β) $\forall A\in\mathcal{A} \Rightarrow A^c\in\mathcal{A}$,

(γ) $\forall (A_i)_{i=1}^{\infty}\in\mathcal{A} \Rightarrow \bigcap_{i=1}^{\infty} A_i\in\mathcal{A}$.

Η λεπτοτητα της $\mathcal{A}\subseteq\mathcal{P}_\Omega:=\{A : A\subseteq\Omega\}$, αντανακλα τη λεπτομερεια της πληροφορησης που εχουμε ή απαιτουμε απο το πειραμα μας.

(1.3) Η πιθανοσυναρτηση P του πειραματος $(\Omega,\mathcal{A})$, ειναι μια συναρτηση $P:\mathcal{A}\ni A\longmapsto P(A)\in[0,1]$,

τετοια ωστε $P(\Omega)=1$,

και $P(\bigcup_{i=1}^{\infty}A_i)=\Sigma_{i=1}^{\infty}P(A_i)$, $\forall (A_i)_{i=1}^{\infty}\in\mathcal{A}$ – ακολουθια ασυμβιβαστων (δηλ., ξενων μεταξυ τους) ενδεχομενων. Η πιθανοσυναρτηση P, επι της $\mathcal{A}$, καθοριζει την κατανομη της συνολικης "μαζας πιθανοτητος" πανω στον Ω, αντανακλα δε τη συχνοτητα με την οποια συμβαινουν ή θα μπορουσαν να συμβουν τα διαφορα ενδεχομενα $A\in\mathcal{A}$ του εν λογω πειραματος $(\Omega,\mathcal{A})$.

(1.4) Παραδειγμα: Εστω το πειραμα της ριψης ενος ζαριου μια φορα, δηλαδη, $\Omega=\{1,2,3,4,5,6\}$. Αν μας ενδιαφερουν ολα τα στοιχειωδη ενδεχομενα $\{i\}$, $i\in\Omega$, και ειναι δυνατον να τα παρατηρησουμε (π.χ., δεν ειναι αναγνωρισιμες διαφορες επιφανειες του ζαριου), τοτε $\mathcal{A}_1=\mathcal{P}_\Omega$, και το πειραμα μας ειναι το $(\Omega,\mathcal{A}_1)$. Αν ομως μας ενδιαφερουν μονο τα στοιχειωδη ενδεχομενα $\{1\}$, $\{2\}$ ή ειναι τα μονα που μπορουμε να παρατηρησουμε, τοτε $\mathcal{A}_2 = \{\{1\},\{2\},\{3,4,5,6\},\Omega, \emptyset, \{1,2\}, \{2,3,4,5,6\}, \{1,3,4,5,6\}\}$, και το πειραμα μας ειναι το $(\Omega, \mathcal{A}_2)$.

Αν ξερουμε, επι πλεον, οτι το εν λογω ζαρι ειναι αμεροληπτο, τοτε ξερουμε και την κατανομη της πιθανοτητος επι του Ω, δηλαδη ξερουμε την πιθανοσυναρτηση P, η οποια οριζεται ως εξης: $P(A) = \dfrac{\#A}{\#\Omega} = \dfrac{\#A}{6}$ $\forall A \in \mathcal{A}_1$. Στην πρωτη περιπτωση που εξετασαμε λοιπον, ο χωρος με πιθανοτητα ειναι ο $(\Omega, \mathcal{A}_1, P)$, στη δε δευτερη περιπτωση, ο αρμοδιος χωρος με πιθανοτητα ειναι ο $(\Omega, \mathcal{A}_2, P')$ οπου P' ειναι ο περιορισμος της P στην $\mathcal{A}_2 \subseteq \mathcal{A}_1$.

(1.5) Παραδειγμα: Εστω $\Omega = \mathbb{R}$, $\mathcal{B} := \sigma\{(-\infty, x], x \in \mathbb{R}\}$ – η μικροτερη σ-αλγεβρα που περιεχει ολες τις ημιευθειες $(-\infty, x]$ (Borel). Η κατ' εξοχην χρησιμοτητα του πειραματος $(\mathbb{R}, \mathcal{B})$ φαινεται στη μελετη των τυχαιων μεταβλητων, παρακατω. Επι του $(\mathbb{R}, \mathcal{B})$ μπορουμε, π.χ., να ορισουμε την πιθανοσυναρτηση

$$P(B) := \int_B f(x)\,dx \quad , \ \forall B \in \mathcal{B},$$

οπου η συναρτηση (πυκνοτητος πιθανοτητος) $f : \mathbb{R} \longrightarrow [0, +\infty)$ ολοκληρωνεται επι του $\mathbb{R}$ στο 1.

(1.6) Ασκηση. Εστω $(\Omega, \mathcal{A}, P)$ χωρος με πιθανοτητα. Δειξτε οτι:

(α) $\emptyset \in \mathcal{A}$.

(β) $\forall (A_i)_{i=1}^{\infty} \in \mathcal{A} \Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$ (με κανονα De Morgan).

(γ) $P(\emptyset) = 0$.

(δ) $P(\bigcup_{i=1}^{n} A_i) = \Sigma_{i=1}^{n} P(A_i)$ $\forall (A_i)_{i=1}^{n} \in \mathcal{A}$ ασυμβιβαστα.

(ε) $P(A^c) = 1 - P(A)$ $\forall A \in \mathcal{A}$.

(ζ) $A \subseteq B \Rightarrow P(A) \leq P(B)$ $\forall A, B \in \mathcal{A}$.

(η) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ $\forall A, B \in \mathcal{A}$.

(θ) $P(\bigcup_{i=1}^{n} A_i) \leq \Sigma_{i=1}^{n} P(A_i)$ $\forall (A_i)_{i=1}^{n} \in \mathcal{A}$ (Ανισοτητα Boole).

(ι) $P(\bigcap_{i=1}^{n} A_i) \geq \Sigma_{i=1}^{n} P(A_i) - (n-1) \quad \forall (A_i)_{i=1}^{n} \in \mathcal{A}$ (Ανισοτητα Bonferroni).

Εστω τωρα, οτι εχουμε προς εκτελεση ενα πειραμα $(\Omega, \mathcal{A})$ με πιθανοσυναρτηση $P(\cdot)$. Εστω δε, οτι μας ερχεται η επι πλεον πληροφορια οτι το αποτελεσμα $\omega \in \Omega$ του πειραματος – οποιο και αν ειναι τελικα – πρεπει (συνθηκη, δεσμευση) να ανηκη στο ενδεχομενο $E \in \mathcal{A}$, δηλαδη το E συμβαινει, τοτε ο αρχικος χωρος με πιθανοτητα $(\Omega, \mathcal{A}, P)$ μεταβαλεται στον $(E, \mathcal{A}_E, P(\cdot|E))$, οπου, $\mathcal{A}_E := \{A \cap E,\ A \in \mathcal{A}\}$, και

(1.7) $P(A|E) := \dfrac{P(A \cap E)}{P(E)} \quad \forall A \in \mathcal{A}$, εφοσον $P(E) \neq 0$,

η οποια και καλειται <u>δεσμευμενη</u> <u>πιθανοσυναρτηση</u> (ή υπο συνθηκες).

Δυο αμεσα αποτελεσματα της ανωτερω εννοιας, ειναι:

(1.8) ο <u>νομος</u> <u>της</u> <u>ολικης</u> <u>πιθανοτητος</u>: $\forall A \in \mathcal{A}$,

$$P(A) = \Sigma_{i=1}^{n} P(A|E_i)P(E_i) \quad \forall (E_i)_{i=1}^{n} \in \mathcal{A}$$ – ακολουθια <u>ασυμβιβαστων</u> και <u>συμπληρωματικων</u> (δηλ., $\bigcup_{i=1}^{n} E_i = \Omega$) ενδεχομενων.

(1.9) ο <u>τυπος</u> <u>του</u> <u>Bayes</u>: $\forall (E_i)_{i=1}^{n} \in \mathcal{A}$ – ασυμβιβαστα και συμπληρωματικα

$$P(E_j|A) = \frac{P(A|E_j)P(E_j)}{\Sigma_{i=1}^{n} P(A|E_i)P(E_i)}, \quad j=1,\dots,n \quad \forall A \in \mathcal{A} : P(A) \neq 0.$$

Η χρησιμοτητα αυτου του τυπου εγκειται στο οτι συνδεει επαγωγικα τις αιτιες $(E_j \in \mathcal{A},\ j=1,\dots,n)$ με τα αποτελεσματα τους $(A \in \mathcal{A})$, ειναι δε ενας συνδιασμος των (1.7) και (1.8).

(1.10) <u>Παραδειγμα</u>: Μια ασφαλιστικη εταιρεια ταξινομει τους πελατες της οδηγους αυτοκινητων ως Ασφαλεις, Μετριους και Κακους, ξερει δε οτι $P(A) = 0{,}2$, $P(M) = 0{,}5$ και $P(K) = 0{,}3$. Εστω E το ενδεχομενο να εχη ενας πελατης της τουλαχιστον ενα δυστυχημα κατα τη διαρκεια ενος ετους. Η εταιρεια ξερει οτι: $P(E|A) = 0{,}05$, $P(E|M) = 0{,}15$ και $P(E|K) = 0{,}3$.

Βρητε τις πιθανοτητες $P(E)$, $P(A|E^c)$.

<u>Απαντηση</u>: $P(E) = P(E|A)P(A) + P(E|M)P(M) + P(E|K)P(K) = 0{,}175$.

$$P(A|E^c) = \frac{P(E^c|A)P(A)}{P(E^c|A)P(A) + P(E^c|M)P(M) + P(E^c|K)P(K)} = \frac{0{,}190}{0{,}825} \simeq 0{,}23.$$

(1.11) <u>Ορισμοι.</u> Εστω χωρος με πιθανοτητα $(\Omega, \mathcal{A}, P)$. Τα $A, B \in \mathcal{A}$ λεγονται:

<u>θετικα συσχετισμενα</u>, αν $P(A \cap B) > P(A)P(B)$,

<u>αρνητικα συσχετισμενα</u>, αν $P(A \cap B) < P(A)P(B)$,

<u>ανεξαρτητα</u>, αν $P(A \cap B) = P(A)P(B)$.

(1.12) <u>Ορισμος.</u> Τα ενδεχομενα $A_1, \ldots, A_n \in \mathcal{A}$ λεγονται <u>ανεξαρτητα</u>, αν και μονο αν για καθε υποομαδα $A_{i_1}, \ldots, A_{i_k}$, $k \leq n$,

$$P\left(\bigcap_{j=1}^{k} A_{i_j}\right) = \prod_{j=1}^{k} P(A_{i_j}).$$

(1.13) <u>Ορισμος.</u> Μια ακολουθια ενδεχομενων $A_1, A_2, \ldots, A_n, \ldots$ λεγεται <u>ακολουθια ανεξαρτητων ενδεχομενων</u> αν και μονο αν τα ενδεχομενα καθε πεπερασμενης υποομαδος της ειναι ανεξαρτητα.

(1.14) <u>Ασκηση.</u> Εστω χωρος με πιθανοτητα $(\Omega, \mathcal{A}, P)$. Δειξτε οτι:

(α) αν $A \cap B = \emptyset$ τοτε τα $A, B \in \mathcal{A}$ ειναι <u>εξηρτημενα</u> αν $P(A), P(B) \neq 0$.

(β) αν $A, B \in \mathcal{A}$ ανεξαρτητα, τοτε A, B^c ανεξαρτητα, καθως και τα A^c, B^c.

(1.15) <u>Παραδειγμα.</u> Ριχνουμε ενα αμεροληπτο ζαρι 4 φορες ανεξαρτητα τη μια απ' την αλλη και εστω $E_i = \{$το i ζαρι ηρθε "εξι"$\}$, $i = 1, \ldots, 4$. Βρητε την πιθανοτητα του ενδεχομενου $E := \{$ηρθε τουλαχιστον ενα "εξι" στους 4 ζαριες$\}$.

<u>Απαντηση.</u> $P(E) = P\left(\bigcup_{i=1}^{4} E_i\right) = 1 - P\left(\left(\bigcup_{i=1}^{4} E_i\right)^c\right) = 1 - P\left(\bigcap_{i=1}^{4} E_i^c\right) =$

$= 1 - \prod_{i=1}^{4} P(E_i^c) = 1 - \prod_{i=1}^{4} [1 - P(E_i)] = 1 - \left(1 - \frac{1}{6}\right)^4 = 0{,}5177$.

Ο Chevalier de Meré ισχυριζοταν στους Pascal και Fermat οτι η $P(E)$ ειναι η ιδια με την πιθανοτητα ενος τουλαχιστον διπλου "εξι" σε 24 ανεξαρτητες ριψεις 2 αμεροληπτων ζαριων. Ειχε δικιο;

ΤΥΧΑΙΕΣ ΜΕΤΑΒΛΗΤΕΣ ΚΑΙ ΚΑΤΑΝΟΜΕΣ.

(1.16) Ορισμος. Εστω πειραμα $(\Omega,\mathcal{A})$, μια συναρτηση $X:\Omega\ni\omega\mapsto X(\omega)\in\mathbb{R}$, λεγεται τυχαια μεταβλητη (τ.μ.) αν και μονο αν $(X\le x):=$
$=\{\omega\in\Omega: X(\omega)\le x\}\in\mathcal{A}\ \ \forall x\in\mathbb{R}\ \ (\Leftrightarrow X^{-1}(B)\in\mathcal{A}\ \forall B\in\mathcal{B})$.

Η "συμπεριφορα" ή κατανομη μιας τυχαιας μεταβλητης X, ορισμενης στο χωρο με πιθανοτητα $(\Omega,\mathcal{A},P)$ χαρακτηριζεται πληρως απο τη συναρτηση κατανομης (σ.κ.) της

(1.17) $\quad F_X(x):=P(X\le x) \quad , x\in\mathbb{R}$.

Το μικροτερο συνολο $S_X\in\mathcal{B}$ τετοιο ωστε $P(X\in S_X)=1$ λεγεται φορεας της τ.μ. X.

(1.18) Ασκηση. Εστω τ.μ. X στο χωρο με πιθανοτητα $(\Omega,\mathcal{A},P)$. Δειξτε οτι:

(α) $(X=x)\in\mathcal{A}$ και $P(X=x)=F_X(x)-F_X(x^-)$.

(β) $(X<x)\in\mathcal{A}$ και $P(X<x)=F_X(x^-)=F_X(x)-[F_X(x)-F_X(x^-)]$.

(γ) $(X>x)\in\mathcal{A}$ και $P(X>x)=1-F_X(x)$.

(δ) $(X\ge x)\in\mathcal{A}$ και $P(X\ge x)=1-F_X(x^-)$.

(ε) $(a<X\le\beta)\in\mathcal{A}$ και $P(a<X\le\beta)=F_X(\beta)-F_X(a)\ \ \forall a,\beta\in\mathbb{R}$.

Θα ασχοληθουμε με τ.μ. που ειναι:

(1.19) διακριτες, δηλαδη ο φορεας τους ειναι διακριτο συνολο $S_X=\{x_1<x_2<\dots<x_n<\dots\}$ και η κατανομη τους καθοριζεται απο μια συναρτηση μαζας πιθανοτητος (σ.μ.π.)

$$p_X(x):=P(X=x)=F_X(x)-F_X(x^-)\ \ge 0 ,$$

για την οποια $\sum_{x\in S_X} p_X(x)=1$,

ειτε,

(1.20) (απολυτα) συνεχεις, δηλαδη υπαρχει συναρτηση πυκνοτητος $f_X:\mathbb{R}\to[0,+\infty)$ τετοια ωστε $\int_{\mathbb{R}} f_X(x)\,dx=1$ και

$$P(X\in B)=\int_B f_X(x)\,dx \quad \forall B\in\mathcal{B}.$$

Ειδικοτερα, εχουμε οτι $F_X(x)=\int_{-\infty}^{x} f_X(z)\,dz\ \ \forall x\in\mathbb{R}$ και

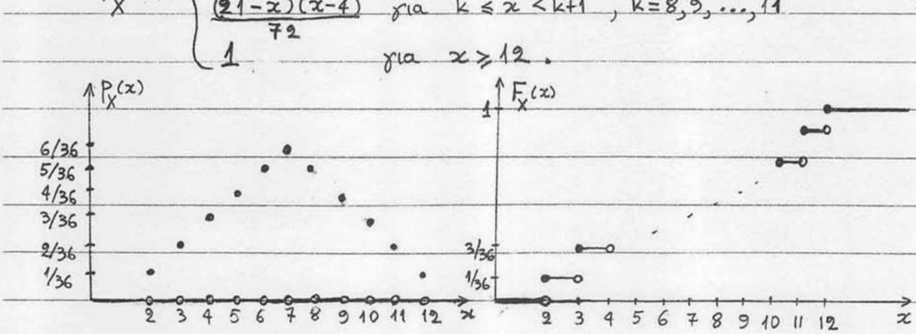
$f_X(x) = F_X'(x) \ \forall x \in \mathbb{R} \setminus N$, οπου το συνολο $N \subseteq \mathbb{R}$ ειναι αριθμησιμο.

(1.21) Παραδειγμα. Εστω το πειραμα $(\Omega, \mathcal{A})$ της ριψης δυο ζαριων, δηλαδη, $\Omega = \{(i,j),\ i,j=1,2,\dots,6\}$, $\mathcal{A} = \mathcal{P}_\Omega$. Υποθετουμε οτι τα ζαρια ειναι αμεροληπτα και τα ριχνουμε ανεξαρτητα το ενα απο το αλλο. Αρα, εχουμε επι του $(\Omega, \mathcal{A})$ την πιθανοσυναρτηση $P((i,j)) = \frac{1}{36} \ \forall (i,j) \in \Omega$. Ας υποθεσουμε οτι δεν μας ενδιαφερει το συγκεκριμμενο ζευγος τιμων (i,j) αλλα το μονο το αθροισμα τους $i+j$, δηλαδη, μας ενδιαφερει η τ.μ.
$X : \Omega \ni (i,j) \longmapsto X((i,j)) = i+j \in S_X = \{2,3,\dots,12\} \subseteq \mathbb{R}$.
Η σ.μ.π. της τ.μ. X ειναι η εξης συμμετρικη γυρω απ' το 7 συναρτηση:

$$p_X(k) = P(X=k) = \begin{cases} \frac{k-1}{36} & \text{αν } k=2,3,\dots,7 \\ \frac{13-k}{36} & \text{αν } k=8,9,\dots,12 \\ 0 & \text{παντου αλλου.} \end{cases}$$

Υπολογιζουμε ενδεικτικα την $p_X(7) = P(\{(1,6),(6,1),(2,5),(5,2),(3,4),(4,3)\}) = \frac{6}{36}$.
Η συναρτηση κατανομης της τ.μ. X ειναι τοτε η ακολουθη:

$$F_X(x) = \begin{cases} 0 & \text{αν } x<2 \\ \frac{x(x-1)}{72} & \text{για } k \le x < k+1,\ k=2,3,\dots,7 \\ \frac{(21-x)(x-4)}{72} & \text{για } k \le x < k+1,\ k=8,9,\dots,11 \\ 1 & \text{για } x \ge 12. \end{cases}$$



(1.22) Ασκηση. Εστω τ.μ. X στο χωρο με πιθανοτητα $(\Omega, \mathcal{A}, P)$, με συναρτηση κατανομης F. Δειξτε οτι:

(a) η $F(x)$, $x \in \mathbb{R}$ ειναι αυξουσα.

(β) $\lim\limits_{x\to-\infty} F(x)=0$, $\lim\limits_{x\to+\infty} F(x)=1$.

(γ) Η $F(x), x\in\mathbb{R}$ είναι συνεχής απ' τα δεξιά.

(1.23) Παράδειγμα. Έστω η (απόλυτα) συνεχής τ.μ. X με πυκνότητα πιθανότητος $f(x|\theta,\sigma)=\frac{1}{2\sigma}e^{-|x-\theta|/\sigma}\,1(x\in\mathbb{R})$, $\theta\in\mathbb{R}$, $\sigma>0$ — οι παράμετροι τοποθεσίας και κλίμακος της κατανομής (του Laplace).

(α) Βρήτε τη σ.κ. F της τ.μ. X, τη μέση της $m:=F^{-1}(1/2)$, και τα κάτω και άνω τεταρτημόρια $\underline{q}:=F^{-1}(1/4)$, $\bar{q}:=F^{-1}(3/4)$ αντιστοίχως.

(β) Επίσης υπολογίστε την πιθανότητα $P(|X-2|>1)$, αν $\theta=1$, $\sigma=2$.

Απάντηση: (α) Έστω $f_0(z)=\frac{1}{2}e^{-|z|}\,1(z\in\mathbb{R})$ η τυποποιημένη Laplace, τότε $f(x|\theta,\sigma)=\frac{1}{\sigma}f_0\left(\frac{x-\theta}{\sigma}\right)$.

Επίσης $F(x|\theta,\sigma)=\int_{-\infty}^{x}f(z|\theta,\sigma)\,dz=\int_{-\infty}^{x}\frac{1}{\sigma}f_0\left(\frac{z-\theta}{\sigma}\right)dz=$

$=\int_{-\infty}^{\frac{x-\theta}{\sigma}}f_0(z)\,dz = F_0\left(\frac{x-\theta}{\sigma}\right)$,

όπου $F_0(z)=\int_{-\infty}^{z}f_0(y)\,dy=\frac{1}{2}\int_{-\infty}^{z}e^{-|y|}\,dy=$

$=\frac{1}{2}e^{-|z|}\,1(z<0)+\left[1-\frac{1}{2}e^{-|z|}\right]1(z\geq 0)$, και άρα

$F(x|\theta,\sigma)=F_0\left(\frac{x-\theta}{\sigma}\right)=\frac{1}{2}\exp\left\{-\frac{|x-\theta|}{\sigma}\right\}1(x<\theta)+\left[1-\frac{1}{2}\exp\left\{-\frac{|x-\theta|}{\sigma}\right\}\right]1(x\geq\theta)$

Άρα, $F(\theta)=1/2 \Rightarrow \theta=F^{-1}(1/2)=m$,

και $\frac{1}{4}=F(\underline{q})=\frac{1}{2}\exp\left\{-\frac{1}{\sigma}|\underline{q}-\theta|\right\} \Rightarrow \underline{q}=F^{-1}(1/2)=\theta-\sigma\log 2$,

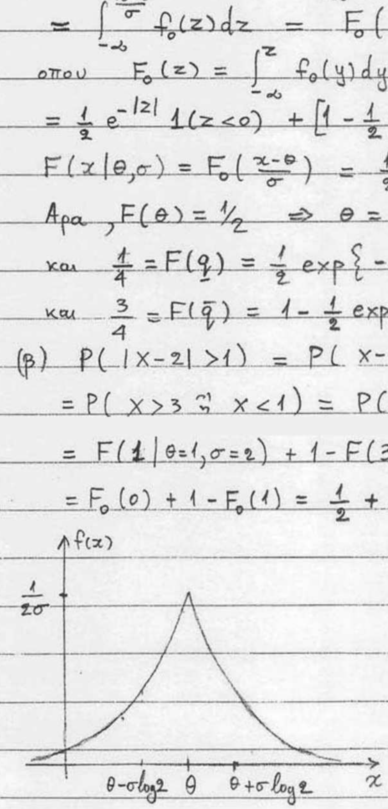
και $\frac{3}{4}=F(\bar{q})=1-\frac{1}{2}\exp\left\{-\frac{1}{\sigma}|\bar{q}-\theta|\right\} \Rightarrow \bar{q}=F^{-1}(1/4)=\theta+\sigma\log 2$.

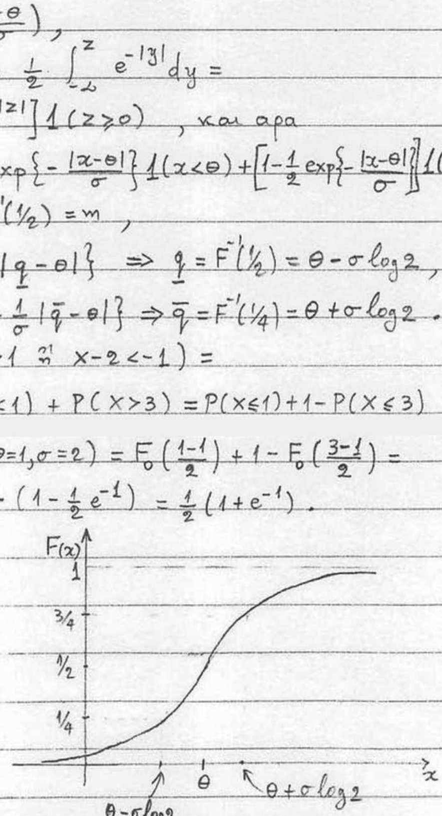
(β) $P(|X-2|>1)=P(X-2>1 \text{ ή } X-2<-1)=$

$=P(X>3 \text{ ή } X<1)=P(X<1)+P(X>3)=P(X\leq 1)+1-P(X\leq 3)$

$=F(1|\theta=1,\sigma=2)+1-F(3|\theta=1,\sigma=2)=F_0\left(\frac{1-1}{2}\right)+1-F_0\left(\frac{3-1}{2}\right)=$

$=F_0(0)+1-F_0(1)=\frac{1}{2}+1-\left(1-\frac{1}{2}e^{-1}\right)=\frac{1}{2}(1+e^{-1})$.





Η συναρτηση κατανομης F_X μιας τ.μ. X στον $(\Omega,\mathcal{A},P)$, οριζεται παντοτε και περιγραφει πληρως την πιθανοθεωρητικη συμπεριφορα της τ.μ. X, καθως και την αντιστοιχη πιθανοθεωρητικη δομη του χωρου με πιθανοτητα $(\Omega,\mathcal{A},P)$ μεσω της σχεσης, $\forall B\in\mathcal{B}$

$$P(X^{-1}(B)) = P(X\in B) = F_X\{B\} := \int_B f_X(x)\,dx \quad \text{αν } X \text{ α.συνεχης},$$
$$:= \sum_{x\in B\cap S_X} p_X(x) \quad \text{αν } X \text{ διακριτη}.$$

Συχνα ομως, ειναι χρησιμο να περιγραφουμε προσεγγιστικα τη συμπεριφορα μιας τ.μ., πιο απλα, χωρις την αναποφευκτη συγχυση που προκαλει το μεγεθος της πληροφοριας που περιεχεται στη σ.κ. F. Αυτο ειναι δυνατον να επιτευχθη, π.χ., μεσω των τριων τιμων, της μεσης και των τεταρτημοριων της F, ή μεσω των <u>ροπων</u> της F:

(1.24) <u>Ορισμος</u>: Η <u>μεση</u> <u>τιμη</u> μιας τ.μ. X οριζεται ως:

$$EX := \begin{cases} \displaystyle\int_{-\infty}^{+\infty} x f_X(x)\,dx & \text{αν η } X \text{ ειναι α.συνεχης \underline{και} } \displaystyle\int_{-\infty}^{+\infty}|x| f_X(x)\,dx<+\infty \\ \displaystyle\sum_{x\in S_X} x\,p_X(x) & \text{αν η } X \text{ ειναι διακριτη \underline{και} } \displaystyle\sum_{x\in S_X}|x|\,p_X(x)<+\infty. \end{cases}$$

(1.25) <u>Προταση</u>: Εστω τ.μ. X στον $(\Omega,\mathcal{A},P)$ και συναρτηση $g:\mathbb{R}\to\mathbb{R}$ τετοια ωστε η $Y:=g(X)$ ειναι επισης τ.μ. στον $(\Omega,\mathcal{A},P)$ (αρκει η g να ειναι κατα τμηματα συνεχης) και $E|Y|<+\infty$. Τοτε, (<u>"τυπος του αφηρημενου στατιστικου"</u>)

$$EY \equiv E\{g(X)\} = \begin{cases} \displaystyle\int_{-\infty}^{+\infty} g(x) f_X(x)\,dx & \text{αν η } X \text{ ειναι α.συνεχης} \\ \displaystyle\sum_{x\in S_X} g(x)\,p_X(x) & \text{αν η } X \text{ ειναι διακριτη.} \end{cases}$$

(1.26) <u>Ορισμος</u>: Η <u>ροπη</u> <u>k-ταξεως</u> της τ.μ. X οριζεται ως η $E(X^k)$, εφοσον $E\{|X|^k\}<+\infty$, με $k=1,2,\dots$

Η <u>κεντρικη</u> <u>ροπη</u> <u>k-ταξεως</u> της τ.μ. X οριζεται ως η $E\{(X-EX)^k\}$, εφοσον $E\{|X|^k\}<+\infty$, με $k=1,2,\dots$

Η κεντρικη ροπη 2-ταξεως καλειται <u>διασπορα</u> της τ.μ. X:

$$\mathcal{D}(X) = E(X-EX)^2 = EX^2-(EX)^2$$

(1.27) <u>Ασκηση:</u> Εστω τ.μ. X · $\forall\, \alpha, \beta \in \mathbb{R}$ – σταθερες – εχουμε:

(α) $E\{\alpha X+\beta\} = \alpha EX + \beta$ εφοσον $E|X| < +\infty$.

(β) $\mathcal{D}(\alpha X+\beta) = \alpha^2 \mathcal{D}(X)$ εφοσον $EX^2 < +\infty$.

(γ) $EX \geq \alpha$ αν $P(X \geq \alpha) = 1$. για καθοια σταθερα $\alpha \in \mathbb{R}$.

(1.28) <u>Ασκηση:</u> (Liapounov) Εστω τ.μ. X με $E|X|^s < +\infty$, τοτε

$$E|X| \leq (EX^2)^{1/2} \leq \cdots \leq (E|X|^k)^{1/k} \leq (E|X|^s)^{1/s} < +\infty \quad \forall\, k \leq s.$$

(1.29) <u>Ασκηση:</u> Εστω τ.μ. X τετοια ωστε $E|X| < +\infty$.

(α) Αν η X ειναι α. συνεχης τοτε:

$$EX = \int_0^\infty [P(X>x) - P(X<-x)]\,dx = \int_0^\infty [1-F_X(x)]\,dx - \int_{-\infty}^0 F_X(x)\,dx.$$

(β) Α η X ειναι διακριτη με φορεα $S_X = \{0, \pm1, \pm2, \ldots\}$, τοτε:

$$EX = \sum_{k=0}^\infty [P(X>k) - P(X<-k)] = \sum_{k=0}^\infty [1-F_X(k)] - \sum_{k=-1}^{-\infty} F_X(k).$$

(1.30) <u>Ορισμος:</u> Η <u>ροπογεννητρια</u> – αν υπαρχη – της τ.μ. X οριζεται ως η

$$M_X(t) := E\{e^{-tX}\}, \quad t \in I_X \subseteq \mathbb{R}$$

(1.31) <u>Προταση</u>. Η ροπογεννητρια μιας τ.μ. X, αν υπαρχη, οριζει μονοσημαντα την κατανομη της X.

(1.32) <u>Ασκηση</u>. Εστω τ.μ. X με ροπογεννητρια $M_X(t)$, $t \in I_X$.

(α) $M_{\alpha X+\beta}(t) = e^{-\beta t} M_X(\alpha t)$, $t \in \{t \in \mathbb{R} : \alpha t \in I_X\}$

(β) $\left.\frac{d^k}{dt^k}\right|_{t=0} M_X(t) = (-1)^k EX^k$, αν $E|X|^k < +\infty$ και $0 \in \mathring{I}_X$.

Στους Πινακες (1.37) και (1.38) εχουμε συγκεντρωσει τις βασικες, εν χρησει, διακριτες και απολυτα συνεχεις κατανομες, τη μεση τιμη, διασπορα και ροπογεννητρια τους.

(1.33) <u>Παραδειγμα:</u> Αν υπαρχουν βρητε:

(α) τη μεση τιμη, διασπορα και <u>τυπικη αποκλιση</u> $\sigma_X = \mathcal{D}(X)^{1/2}$ της $\mathcal{D}(n,p)$.

(β) τη ροπογεννητρια της $\mathcal{G}(\alpha,\lambda)$.

(γ) τη ροπη k-ταξεως της $\mathcal{B}(\alpha,\beta)$.

Απαντηση: (α) $EX = \sum_{k=0}^{n} k \binom{n}{k} p^k (1-p)^{n-k} = n \sum_{k=1}^{n} \binom{n-1}{k-1} p^k (1-p)^{n-k}$

$= np \sum_{m=0}^{n-1} \binom{n-1}{m} p^m (1-p)^{n-1-m} = np\,(p+1-p)^{n-1} = np$

$EX(X-1) = \sum_{k=0}^{n} k(k-1) \binom{n}{k} p^k (1-p)^{n-k} = n(n-1) \sum_{k=2}^{n} \binom{n-2}{k-2} p^k (1-p)^{n-k}$

$= n(n-1)p^2 \sum_{m=0}^{n-2} \binom{n-2}{m} p^m (1-p)^{n-2-m} = n(n-1)p^2 (p+1-p)^{n-2} = n(n-1)p^2$

$\Rightarrow EX^2 - EX = n(n-1)p^2 \Rightarrow EX^2 = np + n(n-1)p^2 \Rightarrow$

$\Rightarrow \mathcal{D}(X) = EX^2 - (EX)^2 = np + n^2p^2 - np^2 - (np)^2 = np(1-p)$

$\Rightarrow \sigma_X = \mathcal{D}(X)^{1/2} = \sqrt{np(1-p)}$.

(β) $X \sim \mathcal{G}(\alpha,\lambda)$, δηλαδη, $f_X(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} 1(x>0)$, $\alpha, \lambda > 0$,

οπου $\Gamma(\alpha) := \int_0^\infty x^{\alpha-1} e^{-x} dx$, $\alpha > 0$.

$M_X(t) = E e^{-tX} = \int_{-\infty}^{+\infty} e^{-tx} f_X(x)\, dx = \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^\infty e^{-tx} x^{\alpha-1} e^{-\lambda x} dx$

$= \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^\infty x^{\alpha-1} \exp\{-(\lambda+t)x\}\, dx = \quad (z \equiv (\lambda+t)x)$

$= [\lambda/(\lambda+t)]^\alpha \frac{1}{\Gamma(\alpha)} \int_0^\infty z^{\alpha-1} e^{-z} dz = \left(1 + \frac{t}{\lambda}\right)^{-\alpha}$, $t \in I_X \equiv (-\lambda, +\infty)$

(γ) $X \sim \mathcal{B}(\alpha,\beta)$, δηλαδη, $f_X(x) = \frac{1}{B(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1} 1(0<x<1)$,

οπου, $B(\alpha,\beta) := \int_0^1 x^{\alpha-1}(1-x)^{\beta-1} dx = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$, $\alpha, \beta > 0$.

Εφοσον $P(0<X<1) = 1 \Rightarrow E|X|^\gamma < 1 < +\infty$, δηλαδη $\exists\, EX^\gamma$ και,

$EX^\gamma = \int_{-\infty}^{+\infty} x^\gamma f_X(x)\, dx = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 x^{\alpha+\gamma-1}(1-x)^{\beta-1} dx =$

$$= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \cdot \frac{\Gamma(\alpha+\gamma)\Gamma(\beta)}{\Gamma(\alpha+\beta+\gamma)} = \frac{\Gamma(\alpha+\beta)\Gamma(\alpha+\gamma)}{\Gamma(\alpha+\beta+\gamma)\Gamma(\alpha)}, \quad \gamma > -\alpha.$$

(1.34) Προταση: Εστω α. συνεχης τ.μ. X και διαφορισιμη μονοτονη συναρτηση $g: \mathbb{R} \to \mathbb{R}$. Τοτε, η τ.μ. $Y \equiv g(X) \sim f_Y(y) = f_X(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right|$.

(1.35) Παραδειγμα: Εστω $X \sim \mathcal{G}(\alpha,\lambda)$, τοτε η $Y \equiv cX \sim \mathcal{G}(\alpha, \frac{\lambda}{c})$, $c>0$.

Απαντηση: $f_Y(y) = \frac{\lambda^\alpha}{\Gamma(\alpha)} (y/c)^{\alpha-1} e^{-\lambda(y/c)} \left|\frac{1}{c}\right| 1(\frac{y}{c}>0) = \frac{(\lambda/c)^\alpha}{\Gamma(\alpha)} y^{\alpha-1} e^{-\frac{\lambda}{c}y} 1(y>0)$.

Ειδικοτερα, αν $\alpha \in \{1, 2, \dots\}$, τοτε $Y \equiv 2\lambda X \sim \mathcal{G}(\alpha, 1/2) \stackrel{d}{=} \chi^2_{2\alpha}$.

(1.36) Ασκηση: Δειξτε οτι αν $X \sim N(\mu, \sigma^2)$ τοτε $Z := \frac{X-\mu}{\sigma} \sim N(0,1)$.

1.37. Πινακας Βασικων Διακριτων Κατανομων.

Κατανομες & παραμετροι	σ.μ.π. $p(x\mid\underline{\theta})=P_{\underline{\theta}}(X=x)$	$E_{\underline{\theta}}X$ $\mathcal{D}_{\underline{\theta}}(X)$	Ροπογεννητρια $M_X(t)=E_{\underline{\theta}}\{e^{-tX}\},\ t\in I_X(\underline{\theta})$.
Ομοιομορφη $\mathcal{U}_\Delta(n)$ $n\in\Theta=\mathbb{N}$	$p(x\mid n)=$ $\frac{1}{n}1(x\in S)$ $S\equiv\{x_1,\dots,x_n\}$	$\frac{n+1}{2}$ $\frac{n^2-1}{12}$	$\frac{e^{-t}(e^{-Nt}-1)}{e^{-t}-1}$ $t\in\mathbb{R}$
Υπεργεωμετρικη $\mathcal{YG}(K,M,n)$ $(K,M,n)\in\Theta=$ $=\mathbb{N}_0^2\times\{0,1,\dots,K\wedge M\}$	$p(x\mid K,M,n)=$ $=\frac{\binom{K}{x}\binom{M}{n-x}}{\binom{K+M}{n}}1(x\in S)$ $S\equiv\{0,1,\dots,n\}$	$\frac{nK}{K+M}$ $\frac{nKM(K+M-n)}{(K+M-1)(K+M)^2}$	$\sum_{k=0}^{n}e^{-tk}\frac{\binom{K}{k}\binom{M}{n-k}}{\binom{M+K}{n}}$ $t\in\mathbb{R}$
Διωνυμικη $\mathcal{D}(n,p)$ $(n,p)\in\Theta=$ $=\mathbb{N}\times[0,1]$	$p(x\mid n,p)=$ $=\binom{n}{x}p^x(1-p)^{n-x}1(x\in S)$ $S\equiv\{0,1,\dots,n\}$	np $np(1-p)$	$(pe^{-t}+1-p)^n$ $t\in\mathbb{R}$
Αρνητικη Διωνυμικη $\mathcal{AD}(k,p)$ $(k,p)\in\Theta=$ $=\mathbb{N}\times[0,1]$	$p(x\mid k,p)=$ $=\binom{x-1}{k-1}p^k(1-p)^{x-k}1(x\in S)$ $S\equiv\{k,k+1,\dots\}$	$\frac{k}{p}$ $\frac{k(1-p)}{p^2}$	$\left[\frac{pe^{-t}}{1-(1-p)e^{-t}}\right]^k$ $t>\log(1-p)$
Poisson $\mathcal{P}(\lambda)$ $\lambda\in\Theta=(0,+\infty)$	$p(x\mid\lambda)=$ $=e^{-\lambda}\frac{\lambda^x}{x!}1(x\in S)$ $S\equiv\{0,1,2,\dots\}$	λ λ	$\exp\{\lambda(e^{-t}-1)\}$ $t\in\mathbb{R}$.

$\mathbb{N}:=\{1,2,3,\dots\}$, $\mathbb{N}_0=\{0,1,2,3,\dots\}$, $\mathbb{R}_+=(0,+\infty)$

Σημειωσεις:
(α) Bernoulli$(p)\overset{d}{=}\mathcal{D}(1,p)$
(β) Γεωμετρικη$(p)\overset{d}{=}\mathcal{AD}(1,p)$
(γ) $X_i\sim\mathcal{D}(n_i,p)$, $i=1,\dots,n$, ανεξαρτητες $\Rightarrow\sum_{i=1}^n X_i\sim\mathcal{D}(\sum_{i=1}^n n_i,p)$.
(γ′) $X_i\sim$ Bern(p), $i=1,\dots,n$, ανεξ. $\Rightarrow\sum_{i=1}^n X_i\sim\mathcal{D}(n,p)$
(δ) $X_i\sim\mathcal{AD}(k_i,p)$, $i=1,\dots,n$, ανεξ. $\Rightarrow\sum_{i=1}^n X_i\sim\mathcal{AD}(\sum_{i=1}^n k_i,p)$
(δ′) $X_i\sim$ Γεωμ.(p), $i=1,\dots,n$, ανεξ. $\Rightarrow\sum_{i=1}^n X_i\sim\mathcal{AD}(n,p)$
(ε) $X_i\sim\mathcal{P}(\lambda_i)$, $i=1,\dots,n$, ανεξ. $\Rightarrow\sum_{i=1}^n X_i\sim\mathcal{P}(\sum_{i=1}^n\lambda_i)$
(ς) $X_1,\dots,X_n,\dots$ α.i. & ανεξ. $N\sim\mathcal{P}(\lambda)$ $\Rightarrow M_{\sum_{i=1}^N X_i}(t)=\exp\{\lambda(M_{X_1}(t)-1)\}$
(compound Poisson)

1.38. Πινακας Βασικων Συνεχων Κατανομων

Κατανομες & Παραμετροι	Πυκνοτητα Πιθανοτητος $f(x\mid\underline{\theta})$	$E_{\underline{\theta}}X$ $\mathcal{D}_{\underline{\theta}}(X)$	Ροπογεννητρια $M_X(t)=E_{\underline{\theta}}\{e^{-tX}\},\ t\in I_X(\theta)$
Ομοιομορφη $\mathcal{U}(\alpha,\beta)$ $-\infty<\alpha<\beta<+\infty$	$\frac{1}{\beta-\alpha}\,1(\alpha<x<\beta)$	$(\alpha+\beta)/2$ $(\beta-\alpha)^2/12$	$\frac{e^{-\alpha t}-e^{-\beta t}}{(\beta-\alpha)t},\ t\in\mathbb{R}$
Κανονικη $N(\mu,\sigma^2)$ $(\mu,\sigma^2)\in\mathbb{R}\times\mathbb{R}_+$	$\frac{1}{\sqrt{2\pi}\,\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}1(x\in\mathbb{R})$	μ σ^2	$e^{-\mu t+\frac{1}{2}\sigma^2t^2},\ t\in\mathbb{R}$
Γαμμα $\mathcal{G}(\alpha,\lambda)$ $(\alpha,\lambda)\in\mathbb{R}_+^2$	$\frac{\lambda^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\lambda x}1(x>0)$	α/λ α/λ^2	$\left(1+\frac{t}{\lambda}\right)^{-\alpha},\ t>-\lambda$
Βητα $\mathcal{B}(\alpha,\beta)$ $(\alpha,\beta)\in\mathbb{R}_+^2$	$\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1}1(0<x<1)$	$\alpha/(\alpha+\beta)$ $\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$	$\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)}\sum_{j=0}^{\infty}\frac{\Gamma(\alpha+j)}{\Gamma(\alpha+\beta+j)}\frac{(-t)^j}{j!}$
Weibull $W(\alpha,\lambda)$ $(\alpha,\lambda)\in\mathbb{R}_+$	$\lambda\alpha x^{\alpha-1}e^{-\lambda x^\alpha}1(x>0)$	$\Gamma(1/\alpha)/(\alpha\lambda^{1/\alpha})$ $\frac{2\Gamma(\frac{2}{\alpha})-\frac{1}{\alpha}\Gamma(\frac{1}{\alpha})^2}{\alpha\,\lambda^{2/\alpha}}$	πολυπλοκη
Cauchy $\mathcal{C}(\alpha,\beta)$ $(\alpha,\beta)\in\mathbb{R}\times\mathbb{R}_+$	$\frac{\beta}{\pi}\frac{1}{\beta^2+(x-\alpha)^2}1(x\in\mathbb{R})$	$\nexists$ $\nexists$	$\nexists$
Student - t (κεντρικη) t_n $n\in\mathbb{N}$	$\frac{\Gamma(\frac{n+1}{2})}{\sqrt{n\pi}\,\Gamma(\frac{n}{2})}\frac{1}{\left(1+\frac{x^2}{n}\right)^{\frac{n+1}{2}}}1(x\in\mathbb{R})$	0 για $n\geq 2$ $\frac{n}{n-2}$ για $n\geq 3$	πολυπλοκη
Fisher $\mathcal{F}_{k,n}$ $(k,n)\in\mathbb{N}^2$	$\frac{\Gamma(\frac{k+n}{2})}{\Gamma(\frac{k}{2})\Gamma(\frac{n}{2})}\left(\frac{k}{n}\right)^{k/2}\frac{x^{k/2-1}}{\left(1+\frac{kx}{n}\right)^{\frac{k+n}{2}}}1(x>0)$	$\frac{n}{n-2}$ για $n\geq 3$ $\frac{2n^2(k+n-2)}{k(n-2)^2(n-4)}$ για $n\geq 5$	πολυπλοκη

Σημειωσεις: (α) $\Gamma(\alpha):=\int_0^\infty x^{\alpha-1}e^{-x}dx,\ \alpha>0$. $\Gamma(\alpha+1)=\alpha\Gamma(\alpha),\ \alpha>0$. $\Gamma(\frac{1}{2})=\sqrt{\pi}$. $\Gamma(n)=(n-1)!$

(β) $\mathcal{G}(1,\lambda)\overset{d}{=}\mathcal{E}(\lambda)$ - εκθετικη, $\mathcal{G}(\frac{n}{2},\frac{1}{2})\overset{d}{=}\chi_n^2$ − χι-τετραγωνο με n βαθμους ελευθεριας

(γ) $t_1\overset{d}{=}\mathcal{C}(0,1)$, $\mathcal{U}(0,1)\overset{d}{=}\mathcal{B}(1,1)$, $\mathcal{F}_{1,n}\overset{d}{=}t_n^2$, $\mathcal{E}(\lambda)\overset{d}{=}W(1,\lambda)$.

(δ) $Z\sim N(0,1)$, $Y\sim\chi_n^2$ & Z,Y ανεξ. τοτε $\frac{\sqrt{n}\,Z}{\sqrt{Y}}\sim t_n$ - Student με n βαθ. ελευθ.

(ε) $X\sim\chi_k^2$, $Y\sim\chi_n^2$ & X,Y ανεξ. τοτε $\frac{nX}{kY}\sim\mathcal{F}_{k,n}$ - Fisher με k,n βαθ. ελευθ.

(ζ) $Z\sim N(0,1)$, τοτε $EZ^{2k}=(2k)!/(2^k k!)$, $k\in\mathbb{N}$.

(η) $X\sim\mathcal{G}(\alpha,\lambda)$, τοτε $EX^\gamma=\Gamma(\alpha+\gamma)/[\lambda^\gamma\Gamma(\alpha)]$, $\gamma>-\alpha$.

(θ) $X\sim\mathcal{B}(\alpha,\beta)$, τοτε $EX^\gamma=\Gamma(\alpha+\beta)\Gamma(\alpha+\gamma)/[\Gamma(\alpha+\beta+\gamma)\Gamma(\alpha)]$, $\gamma>-\alpha$.

(ι) $X\sim W(\alpha,\lambda)$, τοτε $EX^\gamma=(\gamma/\alpha)\Gamma(\gamma/\alpha)/\lambda^{\gamma/\alpha}$, $\gamma>0$.

(κ) $X\sim t_n$, τοτε $EX^{2k}=\frac{n\,.\,1.3.5\ldots(2k-1)}{(n-2)(n-4)\cdots(n-2k)}$, $k=1,2,\ldots,[\frac{n}{2}]$

(λ) $X\sim\mathcal{F}_{k,n}$, τοτε $EX^\gamma=\left(\frac{n}{k}\right)^\gamma\Gamma(\frac{k}{2}+\gamma)\Gamma(\frac{n}{2}-\gamma)/[\Gamma(\frac{k}{2})\Gamma(\frac{n}{2})]$

Μια κοινη μορφη των κατανομων που προαναφεραμε η οποια περιεχει και πολλες αλλες κατανομες ειναι η εξης:

(1.39) Εκθετικη Οικογενεια Κατανομων (Koopman-Pitman-Darmois):

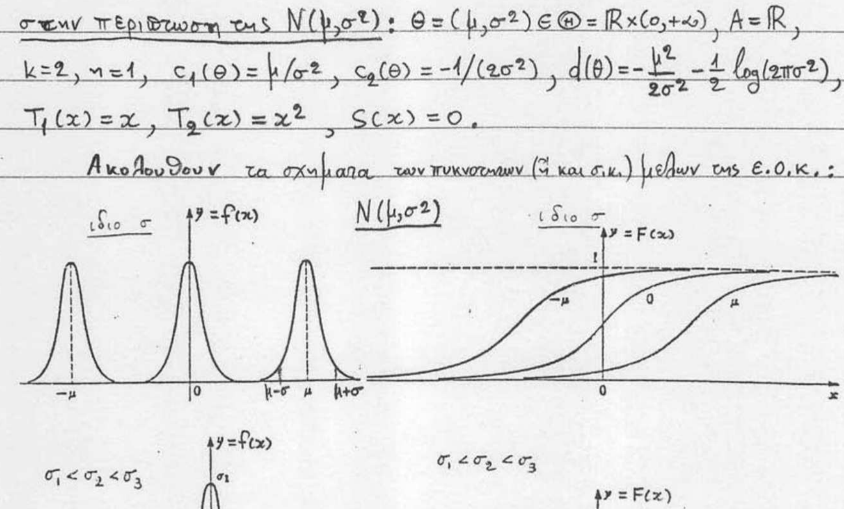
$$f(x|\theta)=\exp\left\{\sum_{i=1}^{k} c_i(\theta)T_i(x)+d(\theta)+S(x)\right\} 1(x\in A)$$

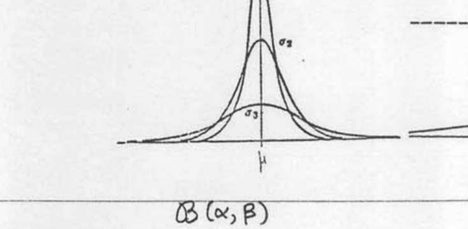
οπου, $A\subseteq\mathbb{R}^n$, $S,T_1,\dots,T_k:\mathbb{R}^n\to\mathbb{R}$ ειναι συναρτησεις μονο του x, $\theta\in\Theta\subseteq\mathbb{R}^k$, $d,c_1,\dots,c_k:\Theta\to\mathbb{R}$ ειναι συναρτησεις μονο του θ.

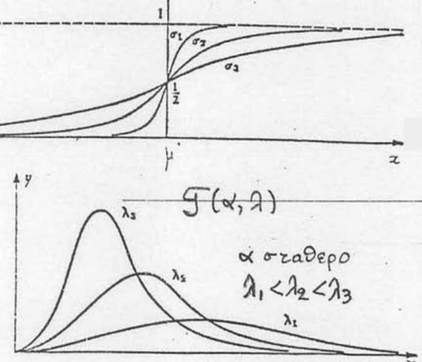
π.χ., στην περιπτωση της $\mathcal{P}(\lambda)$: $\theta\equiv\lambda\in\Theta=(0,+\infty)$, $A=\mathbb{N}_0$, $k=1, n=1$, $c(\theta)=\log\theta$, $d(\theta)=-\theta$, $T(x)=x$, $S(x)=-\log(x!)$,

στην περιπτωση της $N(\mu,\sigma^2)$: $\theta=(\mu,\sigma^2)\in\Theta=\mathbb{R}\times(0,+\infty)$, $A=\mathbb{R}$, $k=2$, $n=1$, $c_1(\theta)=\mu/\sigma^2$, $c_2(\theta)=-1/(2\sigma^2)$, $d(\theta)=-\frac{\mu^2}{2\sigma^2}-\frac{1}{2}\log(2\pi\sigma^2)$, $T_1(x)=x$, $T_2(x)=x^2$, $S(x)=0$.

Ακολουθουν τα σχηματα των πυκνοτητων (ή και σ.κ.) μελων της Ε.Ο.Κ.:







(1.40) Πινακας: τιμων της Φ (σ.κ. της $N(0,1)$), q-ποσοστημοριων ($F^{-1}(q)$), της $N(0,1)$, χ^2_k, t_k, $\mathcal{F}_{k,m}$.

Πινακας τιμων της $\Phi(z)$

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

$z(1-\alpha) := \Phi^{-1}(1-\alpha)$

$1-\alpha$	$z(1-\alpha)$	$1-\alpha$	$z(1-\alpha)$	$1-\alpha$	$z(1-\alpha)$
.50	0	.91	1.341	.995	2.576
.55	.126	.92	1.405	.999	3.090
.60	.253	.93	1.476	.9995	3.291
.65	.385	.94	1.555	.9999	3.719
.70	.524	.95	1.645	.99995	3.891
.75	.674	.96	1.751	.99999	4.265
.80	.842	.97	1.881	.999995	4.417
.85	1.036	.98	2.054	.999999	4.753
.90	1.282	.99	2.326	.9999999	5.199

$X \sim \chi^2_k$. Πινακας $F_X^{-1}(1-\alpha)$

$k \backslash \alpha$	.01	.025	.05	.10	.20
6	16.812	14.449	12.592	10.645	8.558
7	18.475	16.013	14.067	12.017	9.803
8	20.090	17.535	15.507	13.362	11.030
9	21.666	19.023	16.919	14.684	12.242
10	23.209	20.483	18.307	15.987	13.442
11	24.725	21.920	19.675	17.275	14.631
12	26.217	23.337	21.026	18.549	15.812
13	27.688	24.736	22.362	19.812	16.985
14	29.141	26.119	23.685	21.064	18.151
15	30.578	27.488	24.996	22.307	19.311
16	32.000	28.845	26.296	23.542	20.465
17	33.409	30.191	27.587	24.769	21.615
18	34.805	31.526	28.869	25.989	22.760
19	36.191	32.852	30.144	27.204	23.900
20	37.566	34.170	31.410	28.412	25.038
21	38.932	35.479	32.671	29.615	26.171
22	40.289	36.781	33.924	30.813	27.302
23	41.638	38.076	35.173	32.007	28.429
24	42.980	39.364	36.415	33.196	29.553
25	44.314	40.647	37.653	34.382	30.675
26	45.642	41.923	38.885	35.563	31.795
27	46.963	43.195	40.113	36.741	32.912
28	48.278	44.461	41.337	37.916	34.027
29	49.588	45.722	42.557	39.088	35.139
30	50.892	46.979	43.773	40.256	36.250
31	52.191	48.232	44.985	41.422	37.359
32	53.486	49.480	46.194	42.585	38.466
33	54.776	50.725	47.400	43.745	39.572
34	56.061	51.966	48.602	44.903	40.676
35	57.342	53.203	49.802	46.059	41.778
36	58.619	54.437	50.999	47.212	42.879
37	59.893	55.668	52.192	48.363	43.978
38	61.162	56.896	53.384	49.513	45.076
39	62.428	58.120	54.572	50.660	46.173
40	63.691	59.342	55.759	51.805	47.269
45	69.957	65.410	61.656	57.505	52.729
50	76.154	71.420	67.505	63.167	58.164
55	82.292	77.381	73.312	68.796	63.577
60	88.379	83.298	79.082	74.397	68.972
65	94.422	89.177	84.821	79.973	74.351
70	100.425	95.023	90.531	85.527	79.715
75	106.393	100.839	96.217	91.062	85.066
80	112.329	106.629	101.879	96.578	90.405
85	118.236	112.393	107.522	102.079	95.734
90	124.116	118.136	113.145	107.565	101.054
95	129.973	123.858	118.752	113.038	106.364
100	135.807	129.561	124.342	118.498	111.667

$X \sim \mathcal{F}_{k,m}$. Πινακας τιμων $F_X^{-1}(0.95)$

$m \backslash k$	1	2	3	4	5	6	8	12	15	20	30	60	∞
1	161·4	199·5	215·7	224·6	230·2	234·0	238·9	243·9	245·9	248·0	250·1	252·2	254·3
2	18·51	19·00	19·16	19·25	19·30	19·33	19·37	19·41	19·43	19·45	19·46	19·48	19·50
3	10·13	9·55	9·28	9·12	9·01	8·94	8·85	8·74	8·70	8·66	8·62	8·57	8·53
4	7·71	6·94	6·59	6·39	6·26	6·16	6·04	5·91	5·86	5·80	5·75	5·69	5·63
5	6·61	5·79	5·41	5·19	5·05	4·95	4·82	4·68	4·62	4·56	4·50	4·43	4·36
6	5·99	5·14	4·76	4·53	4·39	4·28	4·15	4·00	3·94	3·87	3·81	3·74	3·67
7	5·59	4·74	4·35	4·12	3·97	3·87	3·73	3·57	3·51	3·44	3·38	3·30	3·23
8	5·32	4·46	4·07	3·84	3·69	3·58	3·44	3·28	3·22	3·15	3·08	3·01	2·93
9	5·12	4·26	3·86	3·63	3·48	3·37	3·23	3·07	3·01	2·94	2·86	2·79	2·71
10	4·96	4·10	3·71	3·48	3·33	3·22	3·07	2·91	2·85	2·77	2·70	2·62	2·54
11	4·84	3·98	3·59	3·36	3·20	3·09	2·95	2·79	2·72	2·65	2·57	2·49	2·40
12	4·75	3·89	3·49	3·26	3·11	3·00	2·85	2·69	2·62	2·54	2·47	2·38	2·30
13	4·67	3·81	3·41	3·18	3·03	2·92	2·77	2·60	2·53	2·46	2·38	2·30	2·21
14	4·60	3·74	3·34	3·11	2·96	2·85	2·70	2·53	2·46	2·39	2·31	2·22	2·13
15	4·54	3·68	3·29	3·06	2·90	2·79	2·64	2·48	2·40	2·33	2·25	2·16	2·07
16	4·49	3·63	3·24	3·01	2·85	2·74	2·59	2·42	2·35	2·28	2·19	2·11	2·01
17	4·45	3·59	3·20	2·96	2·81	2·70	2·55	2·38	2·31	2·23	2·15	2·06	
18	4·41	3·55	3·16	2·93	2·77	2·66	2·51	2·34	2·27	2·19	2·11	2·02	
19	4·38	3·52	3·13	2·90	2·74	2·63	2·48	2·31	2·23	2·16	2·07	1·98	1·88
20	4·35	3·49	3·10	2·87	2·71	2·60	2·45	2·28	2·20	2·12	2·04	1·95	1·84
21	4·32	3·47	3·07	2·84	2·68	2·57	2·42	2·25	2·18	2·10	2·01	1·92	1·81
22	4·30	3·44	3·05	2·82	2·66	2·55	2·40	2·23	2·15	2·07	1·98	1·89	1·78
23	4·28	3·42	3·03	2·80	2·64	2·53	2·37	2·20	2·13	2·05	1·96	1·86	1·76
	4·26	3·40	3·01	2·78	2·62	2·51	2·36	2·18	2·11	2·03	1·94	1·84	1·73
25	4·24	3·39	2·99	2·76	2·60	2·49	2·34	2·16	2·09	2·01	1·92	1·82	1·71
26	4·23	3·37	2·98	2·74	2·59	2·47	2·32	2·15	2·07	1·99	1·90	1·80	1·69
27	4·21	3·35	2·96	2·73	2·57	2·46	2·31	2·13	2·06	1·97	1·88	1·79	1·67
28	4·20	3·34	2·95	2·71	2·56	2·45	2·29	2·12	2·04	1·96	1·87	1·77	1·65
29	4·18	3·33	2·93	2·70	2·55	2·43	2·28	2·10	2·03	1·94	1·85	1·75	1·64
30	4·17	3·32	2·92	2·69	2·53	2·42	2·27	2·09	2·01	1·93	1·84	1·74	1·62
40	4·08	3·23	2·84	2·61	2·45	2·34	2·18	2·00	1·92	1·84	1·74	1·64	1·51
60	4·00	3·15	2·76	2·53	2·37	2·25	2·10	1·92	1·84	1·75	1·65	1·53	1·39
120	3·92	3·07	2·68	2·45	2·29	2·17	2·02	1·83	1·75	1·66	1·55	1·43	1·25
∞	3·84	3·00	2·60	2·37	2·21	2·10	1·94	1·75	1·67	1·57	1·46	1·32	1·00

$X \sim t_k$. Πινακας $F_X^{-1}(1-\alpha)$

$k \backslash \alpha$	0·1	0·05	0·025	0·01	0·005
1	3·078	6·314	12·706	31·821	63·657
2	1·886	2·920	4·303	6·965	9·925
3	1·638	2·353	3·182	4·541	5·841
4	1·533	2·132	2·776	3·747	4·604
5	1·476	2·015	2·571	3·365	4·032
6	1·440	1·943	2·447	3·143	3·707
7	1·415	1·895	2·365	2·998	3·499
8	1·397	1·860	2·306	2·896	3·355
9	1·383	1·833	2·262	2·821	3·250
10	1·372	1·812	2·228	2·764	3·169
11	1·363	1·796	2·201	2·718	3·106
12	1·356	1·782	2·179	2·681	3·055
13	1·350	1·771	2·160	2·650	3·012
14	1·345	1·761	2·145	2·624	2·977
15	1·341	1·753	2·131	2·602	2·947
16	1·337	1·746	2·120	2·583	2·921
17	1·333	1·740	2·110	2·567	2·898
18	1·330	1·734	2·101	2·552	2·878
19	1·328	1·729	2·093	2·539	2·861
20	1·325	1·725	2·086	2·528	2·845
21	1·323	1·721	2·080	2·518	2·831
22	1·321	1·717	2·074	2·508	2·819
23	1·319	1·714	2·069	2·500	2·807
24	1·318	1·711	2·064	2·492	2·797
25	1·316	1·708	2·060	2·485	2·787
26	1·315	1·706	2·056	2·479	2·779
27	1·314	1·703	2·052	2·473	2·771
28	1·313	1·701	2·048	2·467	2·763
29	1·311	1·699	2·045	2·462	2·756
30	1·310	1·697	2·042	2·457	2·750
40	1·303	1·684	2·021	2·423	2·704
60	1·296	1·671	2·000	2·390	2·660
120	1·289	1·658	1·980	2·358	2·617
∞	1·282	1·645	1·960	2·326	2·576

ΔΙΑΝΥΣΜΑΤΙΚΕΣ ΤΥΧΑΙΕΣ ΜΕΤΑΒΛΗΤΕΣ.

Εστω τ.μ. $X_1, X_2, \ldots, X_n$ ορισμενες στον ιδιο χωρο με πιθανοτητα $(\Omega, \mathcal{A}, P)$. Συχνα μας ενδιαφερει να μελετησουμε μια τετοια ομαδα τυχαιων μεταβλητων $\underline{X} = (X_1, X_2, \ldots, X_n)^T : \Omega \to \mathbb{R}^n$, σε συνδιασμο, π.χ., για να διερευνησουμε τυχον συσχετισεις τους, π.χ., X_1, X_2, X_3 το βαρος, ηλικια, ζαχαρο ασθενους.

(1.41) Ορισμος. Ενα διανυσμα $\underline{X} := (X_1, \ldots, X_n)^T$, πραγματικων τ.μ. ορισμενων στον ιδιο χωρο με πιθανοτητα $(\Omega, \mathcal{A}, P)$, καλειται n-διαστατη ή διανυσματικη τυχαια μεταβλητη (δ.τ.μ.).

Παρατηρουμε, οτι μια n-διαστατη τ.μ. $\underline{X}$ ειναι μια n-διαστατη συναρτηση $\underline{X} : \Omega \ni \omega \longmapsto \underline{X}(\omega) = (X_1(\omega), \ldots, X_n(\omega))^T \in \mathbb{R}^n$, τετοια ωστε

$$(\underline{X} \leq \underline{x}) := \{\omega \in \Omega : X_i(\omega) \leq x_i, \ i = 1, \ldots, n\} = \bigcap_{i=1}^{n} \{\omega \in \Omega : X_i(\omega) \leq x_i\} \in \mathcal{A} \quad \forall \underline{x} \in \mathbb{R}^n,$$

ή ισοδυναμα $\underline{X}^{-1}(B) \in \mathcal{A} \quad \forall B \in \mathcal{B}^n := \sigma\{(-\infty, x_1] \times \cdots \times (-\infty, x_n], \ x_1, \ldots, x_n \in \mathbb{R}\}$.

(1.42) Ορισμος: Εστω δ.τ.μ. $\underline{X} = (X_1, \ldots, X_n)^T$ στον $(\Omega, \mathcal{A}, P)$. Η συναρτηση κατανομης της δ.τ.μ. $\underline{X}$ οριζεται ως εξης:

$$F_{\underline{X}}(\underline{x}) \equiv F_n(\underline{x}) \equiv F_n(x_1, \ldots, x_n) := P(X_1 \leq x_1, \ldots, X_n \leq x_n) = P\Big(\bigcap_{i=1}^{n} \{\omega \in \Omega : X_i(\omega) \leq x_i\}\Big) \quad \forall \underline{x} \in \mathbb{R}^n.$$

Παρατηρουμε οτι: $P(\underline{X} \in (\underline{\alpha}, \underline{\beta}]) \equiv P(\alpha_1 < X_1 \leq \beta_1, \ldots, \alpha_n < X_n \leq \beta_n)$ $= \Delta_{\alpha_1\beta_1} \cdots \Delta_{\alpha_n\beta_n} F_n(x_1, \ldots, x_n)$, οπου, $\Delta_{\alpha_i\beta_i} F_n(x_1, \ldots, x_n) :=$ $:= F_n(x_1, \ldots, x_{i-1}, \beta_i, x_{i+1}, \ldots, x_n) - F_n(x_1, \ldots, x_{i-1}, \alpha_i, x_{i+1}, \ldots, x_n)$,

π.χ., $P(\alpha_1 < X_1 \leq \beta_1, \alpha_2 < X_2 \leq \beta_2) = F_2(\beta_1, \beta_2) - F_2(\alpha_1, \beta_2) - F_2(\beta_1, \alpha_2) + F_2(\alpha_1, \alpha_2)$.

Αν υπαρχει συναρτηση $f_n : \mathbb{R}^n \ni \underline{x} \mapsto f(\underline{x}) \in [0, +\infty)^n$, με $\int_{\mathbb{R}^n} f_n(\underline{x}) d\underline{x} = 1$, τετοια ωστε $P(\underline{X} \in B) = \int_B f_n(\underline{x}) d\underline{x}$, ή ειδικοτερα τετοια ωστε,

$$F_{\underline{X}}(\underline{x}) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} f_n(x_1', \ldots, x_n') \, dx_1' \cdots dx_n', \quad \text{τοτε},$$

η δ.τ.μ. $\underline{X}$ λεγεται (απολυτα) συνεχης και η συναρτηση $f_n(x_1, \ldots, x_n)$ πυκνοτητα της δ.τ.μ. $\underline{X}$, ισχυει δε οτι

$$f_n(x_1, \ldots, x_n) = \frac{\partial^n}{\partial x_1 \cdots \partial x_n} F_n(x_1, \ldots, x_n) \ \forall \underline{x} \in \mathbb{R}^n \setminus N \ \text{οπου } N \subseteq \mathbb{R}^n \text{ αριθμησιμο.}$$

(1.43) <u>n-διαστατη Κανονικη</u> $N_n(\underline{\mu}, \Sigma)$, $\underline{\mu} \in \mathbb{R}^n$, $\Sigma \in \mathbb{R}^n \times \mathbb{R}^n$ — ενας <u>συμμετρικος, θετικα ορισμενος</u> πινακας.

$$f_n(\underline{x}|\underline{\mu},\Sigma) = \frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}} \exp\left\{-\frac{1}{2}(\underline{x}-\underline{\mu})^T \Sigma^{-1}(\underline{x}-\underline{\mu})\right\} 1(\underline{x} \in \mathbb{R}^n).$$

Για $n=2$, δηλαδη, $\underline{X} \sim N_2(\underline{\mu}, \Sigma) \equiv N_2(\mu_1, \mu_2 | \sigma_1^2, \sigma_2^2 | \rho)$,

$$f_2(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2 - 2\rho\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_2-\mu_2}{\sigma_2}\right)\right]\right\},$$

$x_1, x_2 \in \mathbb{R}$, με παραμετρους $\mu_1, \mu_2 \in \mathbb{R}$, $\sigma_1^2, \sigma_2^2 > 0$, $\rho \in [-1, 1]$.

(1.44) <u>n-διαστατη Βητα</u> $\mathcal{B}(\alpha_1, \ldots, \alpha_n, \alpha_{n+1})$, $\alpha_1, \ldots, \alpha_n, \alpha_{n+1} > 0$, με

$$f_n(\underline{x} \mid \underline{\alpha}) = \frac{\Gamma(\alpha_1 + \cdots + \alpha_n + \alpha_{n+1})}{\Gamma(\alpha_1)\cdots\Gamma(\alpha_n)\Gamma(\alpha_{n+1})} x_1^{\alpha_1 - 1} \cdots x_n^{\alpha_n - 1}(1 - \Sigma_{i=1}^n x_i)^{\alpha_{n+1}-1} \cdot$$

$$\cdot 1(0 < x_1, \ldots, x_n < 1)\, 1(\Sigma_{i=1}^n x_i < 1).$$

(1.45) <u>n-διαστατη Ομοιομορφη</u> $\mathcal{U}_n(B)$, $B \in \mathcal{B}^n$, με

$$f_n(\underline{x}|B) = \frac{1}{|B|} 1(\underline{x} \in B).$$

(1.46) <u>n-διαστατη Γαμμα</u> $\mathcal{G}_n(\alpha_1, \ldots, \alpha_n, \lambda_1, \ldots, \lambda_n)$, $\alpha_i, \lambda_i > 0$, $i = 1, \ldots, n$, με

$$f_n(\underline{x}|\underline{\alpha}, \underline{\lambda}) = \prod_{i=1}^n \left\{ \frac{\lambda_i^{\alpha_i}}{\Gamma(\alpha_i)} x_i^{\alpha_i - 1} e^{-\lambda_i x_i} 1(x_i > 0) \right\}.$$

Αν, τωρα, $X_1, \ldots, X_n$ διακριτες τ.μ. με φορεις S_{X_i}, $i = 1, \ldots, n$ τοτε, η δ.τ.μ. $\underline{X} = (X_1, \ldots, X_n)^T$ ειναι διακριτη, με φορεα $S_{\underline{X}} = S_{X_1} \times \cdots \times S_{X_n}$ — αριθμησιμο συνολο. Η μελετη λοιπον, των διακριτων δ.τ.μ. δεν διαφερει απο την μελετη των μονοδιαστατων διακριτων τ.μ.. Η σ.μ.π. μιας διακριτης δ.τ.μ. οριζεται ως: $P_{\underline{X}}(\underline{x}) := P(\underline{X} = \underline{x}) \equiv P(X_1 = x_1, \ldots, X_n = x_n)$, $\underline{x} \in S_{\underline{X}}$. Εχουμε δε οτι $P(\underline{X} \in A) = \Sigma_{\underline{x} \in A} P_{\underline{X}}(\underline{x})$ $\forall A \subseteq S_{\underline{X}}$.

Χαρακτηριστικο παραδειγμα διακριτης δ.τ.μ. ειναι η γενικευση της $\mathcal{B}(n,p)$:

(1.47) <u>n-διαστατη Πολυωνυμικη</u> $\Pi(p_1, \ldots, p_n, m)$, $m \in \mathbb{N}$ και $p_i \in [0, 1]$, $i = 1, \ldots, n$ και $\Sigma_{i=1}^n p_i \leq 1$.

$$P_n(\underline{x}|\underline{p}, m) = \frac{m!}{x_1! \cdots x_n! (m - \Sigma_{i=1}^n x_i)!} p_1^{x_1} \cdots p_n^{x_n} (1 - \Sigma_{i=1}^n p_i)^{(m - \Sigma_{i=1}^n x_i)} \cdot$$

$$\cdot 1(0 \leq x_1, \ldots, x_n \leq m)\, 1(\Sigma_{i=1}^n x_i \leq m).$$

Παρατηρουμε οτι $\Pi(p_1, m) \stackrel{d}{=} \mathcal{B}(m, p_1)$.

(1.48) Ορισμοί: Εστω η δ.τ.μ. $\underline{X} = (X_1, \ldots, X_n)^T$ στον $(\Omega, \mathcal{A}, P)$.

(α) η μεση τιμη $E\underline{X} := (EX_1, \ldots, EX_n)^T \in \mathbb{R}^n$

(β) η συνδιακυμανση των τ.μ. X_1, X_2 : $cov(X_1, X_2) := E\{(X_1 - EX_1)(X_2 - EX_2)\}$

(γ) ο συντελεστης συνδιακυμανσης των τ.μ. X_1, X_2 : $\rho \equiv \rho(X_1, X_2) :=$

$$:= \frac{cov(X_1, X_2)}{\sqrt{\mathcal{D}(X_1)\mathcal{D}(X_2)}} \in [-1, 1] \quad \text{(από ανισοτητα Cauchy-Schwartz)}$$

(δ) ο πινακας διασπορας της δ.τ.μ. $\underline{X} = (X_1, \ldots, X_n)^T$:

$$\Sigma \equiv E\{(\underline{X} - E\underline{X})(\underline{X} - E\underline{X})^T\} = [cov(X_i, X_j)] \in \mathbb{R}^n \times \mathbb{R}^n,$$

ειναι συμμετρικος, θετικα ορισμενος πινακας (Ασκηση!)

(1.49) Ασκηση: Εστω τ.μ. X, Y ορισμενες στο $(\Omega, \mathcal{A}, P)$, δειξτε οτι:

(α) $cov(X, Y) = EXY - EXEY$.

(β) $cov(X, X) = \mathcal{D}(X)$.

(γ) $cov(\alpha X + \beta, \gamma Y + \delta) = \alpha\gamma\, cov(X, Y)$, $\forall \alpha, \beta, \gamma, \delta \in \mathbb{R}$ σταθερες.

(δ) $\rho(\alpha X + \beta, \gamma Y + \delta) = sgn(\alpha\gamma)\rho(X, Y)$, $\forall \alpha, \beta, \gamma, \delta \in \mathbb{R}$ σταθερες.

(ε) $\rho(X, Y) = \pm 1 \Rightarrow P(Y = \pm X + c) = 1$ για καποια σταθερα $c \in \mathbb{R}$.

(ζ) Αν $(X, Y) \sim N_2(\mu_x, \mu_y \mid \sigma_x^2, \sigma_y^2 \mid \rho)$, δειξτε οτι:

$EX = \mu_x$, $EY = \mu_y$, $\mathcal{D}(X) = \sigma_x^2$, $\mathcal{D}(Y) = \sigma_y^2$, $\rho(X, Y) = \rho$.

(η) $cov(X, Y) = \frac{\partial^2}{\partial t \partial s}\Big|_{t=s=0} M_{X,Y}(t, s) - M_X'(0) M_Y'(0)$, οπου

$M_{X,Y}(t, s) := E \exp\{-tX - sY\}$ η ροπογεννητρια της (X, Y).

Αν $cov(X, Y) = 0$, οι τ.μ. X, Y λεγονται ασυσχετιστες.

(1.50) Παραδειγμα: Εστω η δ.τ.μ. (X, Y) με πυκνοτητα κατανομης:

$f_2(x, y) = y\, e^{-(x+1)y}\, 1(x > 0) 1(y > 0)$. Υπολογιστε:

(α) Τη σ.κ. $F_2(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_2(x', y')\, dx' dy'$

(β) $P(\frac{1}{2} < X \leq 1, \frac{3}{4} < Y \leq 3)$

(γ) $P(X \leq Y)$

(δ) Τις πυκνοτητες των περιθωριων κατανομων των τ.μ. X και Y,

$f_X(x) = \int_{-\infty}^{+\infty} f_2(x, y)\, dy$ και $f_Y(y) = \int_{-\infty}^{+\infty} f_2(x, y)\, dx$.

<u>Απαντηση</u>: (α) $F_2(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y} f_2(z,w)\,dz\,dw=\int_0^y w e^{-w}\left(\int_0^x e^{-wz}dz\right)dw$

$=\int_0^y e^{-w}(1-e^{-wx})\,dw = 1-e^{-y}-\int_0^y e^{-(1+x)w}\,dw=$

$= x/(1+x) - e^{-y} - e^{-(1+x)y}(1+x)^{-1},\quad x,y>0.$

(β) $P(\tfrac12<X\le 1,\ \tfrac34<Y\le 3)=F_2(1,3)-F_2(1,\tfrac34)-F_2(\tfrac12,3)+F_2(\tfrac12,\tfrac34)$

(γ) $P(X\le Y)=\iint_{x\le y} f_2(x,y)\,dx\,dy=\int_0^{\infty}\left(\int_0^y f_2(x,y)\,dx\right)dy=$

$=\int_0^{\infty} y e^{-y}\left(\int_0^y e^{-xy}dx\right)dy=\int_0^{\infty} e^{-y}(1-e^{-y^2})\,dy=1-\int_0^{\infty}e^{-y-y^2}dy$

$=1-e^{1/4}\int_0^{\infty}e^{-(y+1/2)^2}dy=1-\sqrt{\pi\sqrt{e}}\,P(N(-\tfrac14,\tfrac12)>0)=$

$=1-\sqrt{\pi\sqrt{e}}\,P(N(0,1)>\tfrac{1}{2\sqrt2})=1-\sqrt{\pi\sqrt{e}}\left(1-\Phi(\tfrac{1}{2\sqrt2})\right)\approx 0{,}174.$

(δ) $f_X(x)=\int_{-\infty}^{\infty} f_2(x,y)\,dy=\int_0^{\infty} y e^{-(x+1)y}dy=(x+1)^{-2}\Gamma(2)=(x+1)^{-2},\ x>0$

$f_Y(y)=\int_{-\infty}^{+\infty} f_2(x,y)\,dx=e^{-y}\int_0^{\infty} y e^{-yx}dx=e^{-y},\ y>0 \Rightarrow Y\sim\mathcal{E}(\lambda=1).$

(1.51) <u>Ορισμος</u>. Οι τ.μ. $X_1,\dots,X_n$ ορισμενες στο ιδιο $(\Omega,\mathcal{A},P)$, καλουνται <u>ανεξαρτητες</u> ανν $\forall B_1,\dots,B_n\in\mathcal{B}$ τα ενδεχομενα $(X_1\in B_1),\dots,(X_n\in B_n)$ ειναι ανεξαρτητα, δηλαδη $P(X_1\in B_1,\dots,X_n\in B_n)=\prod_{i=1}^{n}P(X_i\in B_i)$.

Ισοδυναμα οι $\underline{X}=(X_1,\dots,X_n)$ ειναι ανεξαρτητες αν και μονο αν:

$F_{\underline{X}}(\underline{x})=\prod_{i=1}^{n}F_{X_i}(x_i)\quad \forall \underline{x}\in\mathbb{R}^n$

ή $\begin{cases} f_{\underline{X}}(\underline{x})=\prod_{i=1}^{n} f_{X_i}(x_i) & \forall \underline{x}\in\mathbb{R}^n \text{ αν ειναι απολ. συνεχεις} \\ p_{\underline{X}}(\underline{x})=\prod_{i=1}^{n} p_{X_i}(x_i) & \forall \underline{x}\in\mathbb{R}^n \text{ αν ειναι διακριτες} \end{cases}$

ή $M_{\underline{X}}(\underline{t})=\prod_{i=1}^{n}M_{X_i}(t_i)\quad \forall \underline{t}\in I_{\underline{X}}\subseteq\mathbb{R}^n$

ή $g_i(X_i),\ i=1,\dots,n$ ανεξαρτητες $\forall g_1,\dots,g_n:\mathbb{R}\to\mathbb{R}:\ g_i(X_i)$ τ.μ. $\forall i=1,\dots,n$.

ή $E\{\prod_{i=1}^{n} g_i(X_i)\}=\prod_{i=1}^{n}E\{g_i(X_i)\}$ $\forall g_1,\dots,g_n$ οπως πριν με $E|g_i(X_i)|<+\infty,\ i=1,\dots,n$.

(1.52) <u>Ασκηση</u>. (α) Εστω τ.μ. X,Y ανεξαρτητες δειξτε οτι $cov(X,Y)=0$.

(β) Εστω οτι η δ.τ.μ. $(X_1,\dots,X_n)\sim N_n(\underline{\mu},\Sigma)$. Δειξτε οτι οι $X_1,\dots,X_n$ ειναι ανεξαρτητες αν και μονο αν ο πινακας Σ ειναι διαγωνιος.

(Εν γενει ομως $cov(X,Y)=0$ <u>δεν</u> συνεπαγεται οτι οι X,Y ειναι ανεξαρτητες.)

(1.53) <u>Προταση</u>: Εστω ανεξαρτητες, τ.μ. $X_1,\dots,X_n$, δειξτε οτι

$$M_{\sum_{i=1}^{n}X_i}(t)=\prod_{i=1}^{n}M_{X_i}(t)\quad \forall\, t\in\bigcap_{i=1}^{n}I_{X_i}.$$

<u>Αποδ.</u> $M_{\Sigma_i^n X_i}(t) = E\{e^{-t\Sigma_{i=1}^n X_i}\} = E\{\prod_{i=1}^n e^{-tX_i}\} = \prod_{i=1}^n E\{e^{-tX_i}\} = \prod_{i=1}^n M_{X_i}(t)$.

(1.54) <u>Ασκηση</u>: Εστω $X_1,\dots,X_n$ <u>ανεξαρτητες</u> τ.μ. στον $(\Omega,\mathcal{A},P)$. Δειξτε οτι:

(α) Αν $X_i \sim \mathcal{D}(n_i,p)$, $i=1,\dots,n$ τοτε $\Sigma_{i=1}^n X_i \sim \mathcal{D}(\Sigma_{i=1}^n n_i, p)$.

(β) Αν $X_i \sim \mathcal{AD}(k_i,p)$, $i=1,\dots,n$ τοτε $\Sigma_{i=1}^n X_i \sim \mathcal{AD}(\Sigma_{i=1}^n k_i, p)$.

(γ) Αν $X_i \sim \mathcal{P}(\lambda_i)$, $i=1,\dots,n$ τοτε $\Sigma_{i=1}^n X_i \sim \mathcal{P}(\Sigma_{i=1}^n \lambda_i)$.

(δ) Αν $X_i \sim N(\mu_i,\sigma_i^2)$, $i=1,\dots,n$ τοτε $\Sigma_{i=1}^n X_i \sim N(\Sigma_{i=1}^n \mu_i, \Sigma_{i=1}^n \sigma_i^2)$.

(ε) Αν $X_i \sim \mathcal{G}(\alpha_i,\lambda)$, $i=1,\dots,n$ τοτε $\Sigma_{i=1}^n X_i \sim \mathcal{G}(\Sigma_{i=1}^n \alpha_i, \lambda)$.

(ζ) Αν $X_i \sim \chi^2_{n_i}$, $i=1,\dots,n$ τοτε $\Sigma_{i=1}^n X_i \sim \chi^2_{\Sigma_{i=1}^n n_i}$.

(η) Αν $X_i \sim \mathcal{C}(\alpha_i,\beta_i)$, $i=1,\dots,n$ τοτε $\Sigma_{i=1}^n X_i \sim \mathcal{C}(\Sigma_{i=1}^n \alpha_i, \Sigma_{i=1}^n \beta_i)$.

Ενδεικτικα αποδεικνυουμε τις (γ) και (δ):

(γ) $M_{\Sigma_{i=1}^n X_i}(t) = \prod_{i=1}^n M_{X_i}(t) = \prod_{i=1}^n \exp\{\lambda_i(e^{-t}-1)\} = \exp\{(\Sigma_{i=1}^n \lambda_i)(e^{-t}-1)\}$

αρα απο το μονοσημαντο της ροπογεννητριας $\Sigma_{i=1}^n X_i \sim \mathcal{P}(\Sigma_{i=1}^n \lambda_i)$.

(δ) οπως και πριν: $M_{\Sigma_{i=1}^n X_i}(t) = \prod_{i=1}^n \exp\{-\mu_i t + \frac{1}{2}\sigma_i^2 t^2\} = \exp\{-t(\Sigma_{i=1}^n \mu_i) + \frac{1}{2}t^2(\Sigma_{i=1}^n \sigma_i^2)\}$.

Τα αποτελεσματα της (1.53) μπορουν να αποδειχθουν και χωρις τη χρηση ροπογεννητριων, αλλα με τη χρηση της <u>συνελιξης</u>:

[Εστω τ.μ. X, Y με πυκνοτητα $f_2(x,y)$:

$$f_{X+Y}(z) = \frac{d}{dz}F_{X+Y}(z) = \frac{d}{dz}P(X+Y\le z) = \frac{d}{dz}\iint_{x+y\le z} f_2(x,y)\,dx\,dy =$$

$$= \frac{d}{dz}\int_{-\infty}^{+\infty}\left(\int_{-\infty}^{z-y} f_2(x,y)\,dx\right)dy = \int_{-\infty}^{+\infty} f_2(z-y,y)\,dy,$$

και αρα αν οι X, Y ειναι ανεξαρτητες, τοτε:

$$(1.55)\quad f_{X+Y}(z) = \int_{-\infty}^{+\infty} f_X(z-y) f_Y(y)\,dy = \int_{-\infty}^{+\infty} f_Y(z-x) f_X(x)\,dx,$$

η οποια καλειται <u>συνελιξη</u> των f_X, f_Y.

Στην περιπτωση διακριτων ανεξαρτητων τ.μ. X, Y εχουμε:

$$P(X+Y=z) = \sum_{x\in S_X} P(X+Y=z, X=x) = \sum_{x\in S_X} P(Y=z-x, X=x)$$

$$= \sum_{x\in S_X} P(Y=z-x)P(X=x), \quad \text{και αρα:}$$

$$(1.56)\quad p_{X+Y}(z) = \sum_{x\in S_X} p_Y(z-x)p_X(x) = \sum_{y\in S_Y} p_X(z-y)p_Y(y),\ \forall z\in S_{X+Y}.$$

Επαναλαβατε την (1.54), χρησιμοποιοντας τις (1.55) ή (1.56) και επαγωγη.

(1.57) <u>Ασκηση</u>: Εστω ανεξαρτητες τ.μ. $X_1, \ldots, X_n$ και εστω

$X_{(1)} := \min\{X_1, \ldots, X_n\}$, $X_{(n)} := \max\{X_1, \ldots, X_n\}$. Δειξτε οτι:

(α) $F_{X_{(n)}}(z) = \prod_{i=1}^{n} F_{X_i}(z)$.

(β) $1 - F_{X_{(1)}}(z) = \prod_{i=1}^{n} [1 - F_{X_i}(z)]$.

(γ) Αν $X_i \sim W(\alpha, \lambda_i), i=1,\ldots,n$, τοτε $X_{(1)} \sim W(\alpha, \sum_{i=1}^{n} \lambda_i)$.

(δ) Αν $X_i \sim \mathcal{U}(0,1)$, $i=1,\ldots,n$, τοτε $X_{(1)} \sim \mathcal{B}(1,n)$ και $X_{(n)} \sim \mathcal{B}(n,1)$.

(1.58) <u>Προταση</u>: Εστω απολ. συνεχης δ.τ.μ. $\underline{X} := (X_1, X_2)$ με πυκνοτητα $f_{\underline{X}}(\underline{x})$, $\underline{x} \in \mathbb{R}^2$, και συναρτησεις $g_1, g_2 : \mathbb{R}^2 \to \mathbb{R}$ τετοιες ωστε:

(i) οι εξισωσεις $\begin{cases} y_1 = g_1(x_1, x_2) \\ y_2 = g_2(x_1, x_2) \end{cases}$ εχουν μοναδικη λυση $\begin{cases} x_1 = h_1(y_1, y_2) \\ x_2 = h_2(y_1, y_2) \end{cases}$,

(ii) υπαρχουν οι παραγωγοι $\frac{\partial}{\partial x_j} g_i(x_1, x_2)$, $i, j = 1,2$ $\forall\, x_1, x_2 \in \mathbb{R}$, και

$$\text{(iii)}\quad J(x_1, x_2) := \begin{vmatrix} \frac{\partial}{\partial x_1} g_1(x_1,x_2) & \frac{\partial}{\partial x_2} g_1(x_1,x_2) \\ \frac{\partial}{\partial x_1} g_2(x_1,x_2) & \frac{\partial}{\partial x_2} g_2(x_1,x_2) \end{vmatrix} \neq 0 .$$

Τοτε, υπαρχει η πυκνοτητα $f_{\underline{Y}}(\underline{y})$ της δ.τ.μ. $\underline{Y} = (Y_1, Y_2) := (g_1(X_1,X_2), g_2(X_1,X_2))$, και,

$$f_{\underline{Y}}(y_1, y_2) = \left. \frac{f_{\underline{X}}(x_1, x_2)}{|J(x_1, x_2)|} \right|_{x_i = h_i(y_1, y_2),\ i=1,2} \qquad \forall\, \underline{y} \in \mathbb{R}^2 .$$

(1.59) <u>Παραδειγμα</u>: Εστω τ.μ. $X \sim N(0,1)$ και $Y \sim \chi^2_n$ ανεξαρτητες.

Δειξτε οτι η τ.μ. $Z := \sqrt{n}\, X / \sqrt{Y} \sim t_n$.

<u>Αποδ.</u> Οριζουμε τη βοηθητικη τ.μ. $W := (Y/n)^{1/2}$.

Εχουμε, $X = ZW$, $Y = nW^2$ και αρα $|J|^{-1} = 2n\omega^2$, δηλαδη

$$f_{Z,W}(z,\omega) = f_{X,Y}(z\omega, n\omega^2)\, 2n\omega^2 = f_X(z\omega) f_Y(n\omega^2)\, 2n\omega^2 =$$

$$= \Gamma(\tfrac{n}{2})^{-1} \pi^{-1/2} n^{n/2} 2^{-(n-1)/2} \omega^n \exp\{-\tfrac{1}{2}(n+z^2)\omega^2\} \mathbb{1}(z \in \mathbb{R}) \mathbb{1}(\omega > 0).$$

Αρα, $f_Z(z) = \int_{-\infty}^{+\infty} f_{Z,W}(z,\omega)\, d\omega = (n\pi)^{-1/2} \Gamma(\frac{n}{2})^{-1} (1 + \frac{z^2}{n})^{-\frac{n+1}{2}}\, 2\int_0^\infty \xi^n e^{-\xi^2} d\xi$,

οπου $\xi \equiv ((n+z^2)/2)^{1/2} \omega$. Επισης $2\int_0^\infty \xi^n e^{-\xi^2} d\xi = \Gamma(\frac{n+1}{2})$ (θετε $\zeta = \xi^2$).

$$\text{Αρα,}\quad f_Z(z) = \frac{\Gamma(\frac{n+1}{2})}{\sqrt{n\pi}\, \Gamma(\frac{n}{2})} \frac{1}{\left(1 + \frac{z^2}{n}\right)^{\frac{n+1}{2}}} \mathbb{1}(z \in \mathbb{R}) .$$

(1.60) <u>Παράδειγμα</u>: Έστω τ.μ. $X \sim \chi^2_k$ και $Y \sim \chi^2_n$ ανεξάρτητες.
Δείξτε ότι η τ.μ. $Z := \frac{n}{k}\frac{X}{Y} \sim F_{k,n}$.
<u>Απόδ.</u> Θέτουμε $W := Y$ και έχουμε $f_{X,Y}(x,y) = f_X(x) f_Y(y)$,
$X = \frac{k}{n} Z W$, $Y = W$, δηλαδή $|J|^{-1} = \frac{k}{n}\omega$. Άρα,
$$f_{Z,W}(z,\omega) = 2^{-(k+n)/2}\left[\Gamma\left(\tfrac{k}{2}\right)\Gamma\left(\tfrac{n}{2}\right)\right]^{-1}\left(\tfrac{k}{n}\right)^{k/2} z^{k/2-1}\omega^{n/2}\exp\left\{-\left(\tfrac{k}{n}z+1\right)\omega/2\right\}1(z,\omega>0),$$
$$\Rightarrow f_Z(z) = \int_{-\infty}^{+\infty} f_{Z,W}(z,\omega)\,d\omega = \frac{2^{-(k+n)/2}}{\Gamma(\frac{k}{2})\Gamma(\frac{n}{2})}\left(\frac{n}{k}\right)^{k/2} z^{k/2-1}\int_0^\infty \omega^{\frac{k+n}{2}-1}\exp\left\{-\tfrac{1}{2}\left(\tfrac{k}{n}z+1\right)\omega\right\}d\omega$$
$$= \frac{\Gamma(\frac{k+n}{2})}{\Gamma(\frac{k}{2})\Gamma(\frac{n}{2})}\left(\frac{n}{k}\right)^{k/2} z^{k/2-1}\left(1+\frac{k}{n}z\right)^{-(k+n)/2},$$ δηλαδή, $Z \sim F_{k,n}$.

(1.61) <u>Παράδειγμα</u>: Έστω οι τ.μ. $X_i \sim N(\mu_i, \sigma_i^2)$, $i=1,\dots,n$ ανεξάρτητες.
Δείξτε ότι: (α) $Y_i := \left(\frac{X_i-\mu_i}{\sigma_i}\right)^2 \sim \chi^2_1$, $i=1,\dots,n$ ανεξάρτητες.
(β) $\Sigma_{i=1}^n Y_i \sim \chi^2_n$.
<u>Απόδ.</u> (α) Από τις (1.36), (1.51) $Z_i := \frac{X_i-\mu_i}{\sigma_i} \sim N(0,1)$, $i=1,\dots,n$ ανεξάρτητες.
Τώρα, αν $Z \sim N(0,1)$, τότε $f_{|Z|}(\omega) = \frac{d}{d\omega}P(|Z|\le\omega) = \frac{d}{d\omega}P(-\omega\le Z\le\omega)$
$= \frac{d}{d\omega}\{F_Z(\omega) - F_Z(-\omega)\} = f_Z(\omega) + f_Z(-\omega) = [\phi(\omega)+\phi(-\omega)]1(\omega>0) = 2\phi(\omega)1(\omega>0)$,
και άρα από την (1.34), $Y = |Z|^2 \sim f_Y(y) = 2\phi(\sqrt{y})\frac{1}{2\sqrt{y}} \Rightarrow Y \sim \mathcal{G}(\frac{1}{2},\frac{1}{2}) \overset{d}{=} \chi^2_1$.
(β) Άρα από την (1.54 ζ) $\Sigma_{i=1}^n Y_i \sim \chi^2_n$.

(1.62) <u>Παράδειγμα</u>: Έστω οι ανεξάρτητες και ισόνομες τ.μ. $X_1,\dots,X_n \sim N(\mu,\sigma^2)$.
Ορίζουμε: $\overline{X}_n := \frac{1}{n}\Sigma_{i=1}^n X_i$ και $S_n^2 := \frac{1}{n-1}\Sigma_{i=1}^n (X_i - \overline{X}_n)^2$. Δείξτε ότι:
(α) $\overline{X}_n \sim N(\mu, \frac{\sigma^2}{n})$, $(n-1)S_n^2/\sigma^2 \sim \chi^2_{n-1}$ <u>ανεξάρτητες</u>.
(β) $T_n := \frac{\sqrt{n}(\overline{X}_n - \mu)}{S_n} \sim t_{n-1}$.
<u>Απόδ.</u> (α) Θέτουμε $\underline{\alpha}_1 := \left(\frac{1}{\sqrt{n}},\dots,\frac{1}{\sqrt{n}}\right)^T \in \mathbb{R}^n$, και για $k=2,3,\dots,n$
θέτουμε: $\underline{\alpha}_k := \Big(\underbrace{-\frac{1}{\sqrt{k(k-1)}},\dots,-\frac{1}{\sqrt{k(k-1)}}}_{k-1 \text{ θέσεις}}, \sqrt{\frac{k-1}{k}}, 0,\dots,0\Big)^T \in \mathbb{R}^n$.

Παρατηρούμε ότι $\underline{\alpha}_k^T\underline{\alpha}_m = \delta_{km}$ ($=0$ αν $k\ne m$ και 1 αν $k=m$), $\forall k,m=1,\dots,n$.
Θέτουμε $\underline{X} := (X_1,\dots,X_n)^T$ και $\underline{Y} := (Y_1,\dots,Y_n)^T$, με $Y_i := \underline{\alpha}_i^T\underline{X}$, $i=1,\dots,n$,
δηλαδή, $\underline{Y} = \underline{\underline{A}}\,\underline{X}$, όπου $\underline{\underline{A}}$ ο ορθοκανονικός πίνακας με k-γραμμή
το διάνυσμα-γραμμή $\underline{\alpha}_k^T$, $k=1,2,\dots,n$. ($A^T = A^{-1}$)

Τωρα, εχουμε, $M_{Y_i}(t) = E\{e^{-tY_i}\} = E\{e^{-t\Sigma_{j=1}^n \alpha_{ij}X_j}\} =$

$= E\{\prod_{j=1}^n e^{-\alpha_{ij}tX_i}\} = \prod_{j=1}^n M_{X_i}(\alpha_{ij}t) = \exp\{-\mu t\Sigma_{j=1}^n \alpha_{ij} + \frac{1}{2}\sigma^2 t^2 \Sigma_{j=1}^n \alpha_{ij}^2\}$,

οπου για $i=1$: $\Sigma_{j=1}^n \alpha_{1j} = \Sigma_{j=1}^n \frac{1}{\sqrt{n}} = \sqrt{n}$, $\Sigma_{j=1}^n \alpha_{1j}^2 = \Sigma_{i=1}^n \frac{1}{n} = 1$,

αρα, $M_{Y_1}(t) = \exp\{-t\mu\sqrt{n} + \frac{1}{2}\sigma^2 t\} \Rightarrow Y_1 \quad N(\mu\sqrt{n}, \sigma^2)$.

Για $i \geq 2$, $\Sigma_{j=1}^n \alpha_{ij} = \sqrt{n}\,\Sigma_{j=1}^n \alpha_{ij}\alpha_{1j} = \sqrt{n}\,\underline{\alpha}_i^T\underline{\alpha}_1 = 0$,

και, $\Sigma_{j=1}^n \alpha_{ij}^2 = \underline{\alpha}_i^T\underline{\alpha}_i = 1$, δηλαδη, $M_{Y_i}(t) = \exp\{\frac{1}{2}\sigma^2 t^2\}$

και αρα $Y_2, Y_3, \ldots, Y_n \sim N(0, \sigma^2)$.

Επισης, $\text{cov}(Y_i, Y_j) = \text{cov}(\Sigma_{k=1}^n \alpha_{ik}X_k, \Sigma_{\ell=1}^n \alpha_{j\ell}X_\ell) =$

$= \Sigma_{k=1}^n \Sigma_{\ell}^n \alpha_{ik}\alpha_{j\ell}\,\text{cov}(X_k, X_\ell) = \Sigma_{k=1}^n \Sigma_{\ell=1}^n \alpha_{ik}\alpha_{j\ell}\sigma^2\delta_{k\ell} =$

$= \sigma^2 \Sigma_{k=1}^n \alpha_{ik}\alpha_{jk} = \sigma^2 \underline{\alpha}_i^T\underline{\alpha}_j = 0$ αν $i \neq j$ $\forall i,j = 1, \ldots, n$.

Εχουμε, λοιπον, οτι : $Y_1 \sim N(\mu\sqrt{n}, \sigma^2)$, $Y_2, Y_3, \ldots, Y_n \sim N(0, \sigma^2)$

και ειναι <u>ασυσχετιστες</u> και αρα απο την <u>(1.52 β)</u> <u>ανεξαρτητες</u>.

Τωρα, εχουμε $Y_1 = \Sigma_{i=1}^n \frac{1}{\sqrt{n}} X_i = \sqrt{n}\,\bar{X}_n \Rightarrow \bar{X}_n = \frac{1}{\sqrt{n}} Y_1 \sim N(\mu, \frac{\sigma^2}{n})$.

Επισης, εφοσον $\underline{\alpha}_i^T\underline{\alpha}_j = \delta_{ij}$ εχουμε οτι $\underline{\underline{A}}^T\underline{\underline{A}} = \underline{\underline{I}}$ ο μοναδιαιος πινακας,

και αρα $\underline{Y}^T\underline{Y} = (\underline{\underline{A}}\,\underline{X})^T(\underline{\underline{A}}\,\underline{X}) = \underline{X}^T\underline{\underline{A}}^T\underline{\underline{A}}\,\underline{X} = \underline{X}^T\underline{\underline{I}}\,\underline{X} = \underline{X}^T\underline{X}$, δηλαδη,

$(n-1)S_n^2 = \Sigma_{i=1}^n (X_i - \bar{X}_n)^2 = \Sigma_{i=1}^n X_i^2 + n\bar{X}_n^2 - 2\bar{X}_n\Sigma_{i=1}^n X_i =$

$= \Sigma_{i=1}^n X_i^2 + n\bar{X}_n^2 - 2n\bar{X}_n^2 = \Sigma_{i=1}^n X_i^2 - n\bar{X}_n^2 = \underline{X}^T\underline{X} - (\sqrt{n}\bar{X}_n)^2$

$= \underline{Y}^T\underline{Y} - Y_1^2 = \Sigma_{i=1}^n Y_i^2 - Y_1^2 = \Sigma_{i=2}^n Y_i^2$ (βλ. και (1.61)).

Αρα, $(n-1)S_n^2/\sigma^2 \sim \chi_{n-1}^2$ <u>ανεξαρτητη</u> της $\bar{X}_n \sim N(\mu, \frac{\sigma^2}{n})$.

(β) Εχουμε λοιπον οτι

$\sqrt{n}(\bar{X}_n - \mu)/\sigma \sim N(0,1)$ <u>ανεξαρτητη</u> της $(n-1)S_n^2/\sigma^2 \sim \chi_{n-1}^2$,

και αρα $T := \frac{\sqrt{n}(\bar{X}_n - \mu)}{S_n} = \frac{\sqrt{n}(\bar{X}_n - \mu)/\sigma}{\sqrt{\frac{(n-1)S_n^2/\sigma^2}{n-1}}} \sim t_{n-1}$ (απο (1.59)).

(1.63) <u>Ασκηση</u>: Εστω $X_1, \ldots, X_n \sim N(\mu, \sigma^2)$ ανεξαρτητες και ανεξαρτητες των ανεξαρτητων $Y_1, \ldots, Y_m \sim N(\nu, \tau^2)$.

Εστω, $S_X^2 := \frac{1}{n-1}\Sigma_{i=1}^n (X_i - \bar{X}_n)^2$, $S_Y^2 := \frac{1}{m-1}\Sigma_{i=1}^m (Y_i - \bar{Y}_m)^2$.

Δειξτε οτι :

$$F := \frac{\tau^2}{\sigma^2} \cdot \frac{S_X^2}{S_Y^2} \sim F_{n-1, m-1}.$$

ΔΕΣΜΕΥΜΕΝΕΣ ΤΥΧΑΙΕΣ ΜΕΤΑΒΛΗΤΕΣ ΚΑΙ ΚΑΤΑΝΟΜΕΣ.

Θα εργασθουμε με δυο τ.μ. X, Y : η γενικευση, των οσων εκθεσουμε, σε μεγαλυτερη διασταση ειναι αμεση.

Εστω λοιπον διακριτες τ.μ. X, Y με σ.μ.π. $p_2(x,y)$, η

$$P(Y=y \mid X=x) = \frac{P(X=x, Y=y)}{P(X=x)} = \frac{p_2(x,y)}{p_X(x)} \quad \forall x : p_X(x) \neq 0.$$

Οριζουμε λοιπον τη <u>δεσμευμενη σ.μ.π.</u> της τ.μ. Y δεδομενου οτι $X=x$, ως

$$(1.64)\quad p(y|x) := \frac{p_2(x,y)}{p_X(x)}, \quad x \in S_X, \quad y \in S_Y,$$

και παρατηρουμε οτι: $p(y|x) \geq 0$ και $\sum_{y\in S_Y} p(y|x) = \frac{\sum_{y\in S_Y} p_2(x,y)}{p_X(x)} = \frac{p_X(x)}{p_X(x)} = 1$, και αρα $Y \mid X=x \sim p(y|x)$ καλα ορισμενη σ.μ.π.

Αναλογα, στην περιπτωση που οι τ.μ. X, Y ειναι απολ. συνεχεις με πυκνοτητα $f_2(x,y)$, $x, y \in \mathbb{R}$, οριζουμε τη <u>δεσμευμενη πυκνοτητα</u>:

$$(1.65)\quad f(y|x) := \frac{f_2(x,y)}{f_X(x)}, \quad \forall x \in \mathbb{R} : f_X(x) \neq 0 \text{ και } \forall y \in \mathbb{R},$$

Παρατηρουμε δε οτι: $f(y|x) \geq 0$ και $\int_{-\infty}^{+\infty} f(y|x)dy = \frac{\int_{-\infty}^{+\infty} f_2(x,y)dy}{f_X(x)} = \frac{f_X(x)}{f_X(x)} = 1$, και αρα $Y \mid X=x \sim f(y|x)$ καλα ορισμενη πυκνοτητα.

Επισης, οριζουμε τη <u>συναρτηση</u> <u>της δεσμευμενης</u> <u>μεσης τιμης</u>:

$$(1.66)\quad \mu(x) \equiv E(Y|X=x) := \begin{cases} \sum_{y\in S_Y} y\, p(y|x) & \text{εφοσον } \sum_{y\in S_Y} |y| p(y|x) < +\infty \\ \text{ή} & \\ \int_{-\infty}^{+\infty} y f(y|x) dy & \text{εφοσον } \int_{-\infty}^{+\infty} |y| f(y|x) dy < +\infty, \end{cases}$$

$\forall x \in \mathbb{R} : p_X(x) \neq 0$ ή $f_X(x) \neq 0$ αντιστοιχως,

καθως επισης και τη <u>συναρτηση</u> <u>της δεσμευμενης</u> <u>διασπορας</u>:

$$(1.67)\quad \sigma^2(x) \equiv \mathcal{D}(Y|X=x) := E\{(Y-\mu(x))^2 \mid X=x\} \quad \forall x \in \mathbb{R} : p_X(x) \neq 0$$

ή $f_X(x) \neq 0$ και εφοσον $E(Y^2|X=x) < +\infty$.

Εφοσον οι συναρτησεις $\mu(x)$, $\sigma^2(x)$ ειναι καλα ορισμενες, θεωρουμε τις αντιστοιχες τ.μ.:

$$(1.68)\quad E(Y|X) := \mu(X),$$

$$(1.69)\quad \mathcal{D}(Y|X) := \sigma^2(X) = E(Y^2|X) - \{E(Y|X)\}^2 \quad (\text{ασκηση!})$$

και έχουμε:

(1.70) <u>Πρόταση</u>: (α) $E\{E(Y|X)\} = E\{\mu(X)\} = EY$,

(β) $\mathcal{D}(Y) = E\{\mathcal{D}(Y|X)\} + \mathcal{D}(E(Y|X)) = E\{\sigma^2(X)\} + \mathcal{D}(\mu(X))$.

<u>Απόδ.</u> (α) <u>για διακριτές τ.μ.</u>: $E(\mu(X)) = \sum_{x\in S_X} \mu(x) p_X(x) =$

$= \sum_{x\in S_X} \left(\sum_{y\in S_Y} y\, p(y|x)\right) p_X(x) = \sum_{y\in S_Y} y \sum_{x\in S_X} p(y|x) p_X(x) =$

$= \sum_{y\in S_Y} y \sum_{x\in S_X} p_2(x,y) = \sum_{y\in S_Y} y\, p_Y(y) = EY$.

<u>για απολ. συνεχείς τ.μ.</u>: $E(\mu(X)) = \int_{-\infty}^{+\infty} \mu(x) f_X(x)\,dx =$

$= \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{+\infty} y\, f(y|x)\,dy\right) f_X(x)\,dx = \int_{-\infty}^{+\infty} y \left(\int_{-\infty}^{+\infty} f(y|x) f(x)\,dx\right) dy$

$= \int_{-\infty}^{+\infty} y \left(\int_{-\infty}^{+\infty} f_2(x,y)\,dx\right) dy = \int_{-\infty}^{+\infty} y\, f_Y(y)\,dy = EY$.

(β) $E\{\mathcal{D}(Y|X)\} = E\{E(Y^2|X) - (E(Y|X))^2\} =$

$= E\{E(Y^2|X)\} - E\{E(Y|X)\}^2 = EY^2 - E\{\mu(X)\}^2 =$

$= EY^2 - (EY)^2 + \{E(\mu(X))\}^2 - E\{\mu(X)\}^2 = \mathcal{D}(Y) - \mathcal{D}(\mu(X))$.

Ο γενικευμένος <u>τύπος της ολικής πιθανότητος</u> έπεται από τον (1.70α):

(1.71) $P((X,Y)\in B) = E\{1((X,Y)\in B)\} = E\{E(1((X,Y)\in B)|X)\} =$

$= \int_{-\infty}^{+\infty} E[1((X,Y)\in B)|X=x] f_X(x)\,dx =$

$= \int_{-\infty}^{+\infty} P((x,Y)\in B)|X=x) f_X(x)\,dx = \int_{-\infty}^{+\infty} \left\{\int_{(x,y)\in B} f(y|x)\,dy\right\} f_X(x)\,dx$,

$\forall B \in \mathcal{B}^2$.

Έχουμε επίσης το τους γενικευμένους <u>τύπους του Bayes</u>:

(1.72) $p(x|y) = \dfrac{p(y|x) p_X(x)}{\sum_{z\in S_X} p(y|z) p_X(z)}$, για διακριτές τ.μ.,

(1.73) $f(x|y) = \dfrac{f(y|x) f_X(x)}{\int_{-\infty}^{+\infty} f(y|z) f_X(z)\,dz}$, για απολ. συνεχείς τ.μ.,

(1.74) $f(x|k) = \dfrac{p(k|x) f_X(x)}{\int_{-\infty}^{+\infty} p(k|z) f_X(z)\,dz}$, για συνδιασμό διακριτής και απολ. συνεχούς τ.μ..

(1.75) <u>Άσκηση</u>. Αποδείξτε τους (1.72), (1.73) και (1.74).

(<u>Υπόδ.</u> για τον (1.74) κάντε χρήση τον (1.71).)

(1.76) <u>Παραδειγμα</u>: Εστω $X \mid \Lambda=\lambda \sim \text{Poisson}(\lambda)$, δηλαδη $p(x|\lambda)=\frac{e^{-\lambda}\lambda^{x}}{x!}1(x\in\mathbb{N}_0)$, και εστω οτι η <u>a priori</u>" κατανομη της Λ ειναι η $\mathcal{G}(\alpha,\mu)$, δηλαδη, $f_\Lambda(\lambda)=\frac{\mu^\alpha}{\Gamma(\alpha)}\lambda^{\alpha-1}e^{-\mu\lambda}1(\lambda>0)$. Δειξτε οτι η <u>"a posteriori"</u> κατανομη της $\Lambda \mid X=x$ ειναι $\mathcal{G}(\alpha+x,\mu+1)$, δηλαδη,

$$f(\lambda|x)=\frac{(\mu+1)^{\alpha+x}}{\Gamma(\alpha+x)}\lambda^{\alpha+x-1}e^{-(\mu+1)\lambda}1(\lambda>0).$$

<u>Αποδ.</u> Απο την (1.74), $f(\lambda|x)=\dfrac{p(x|\lambda)f_\Lambda(\lambda)}{\int_{-\infty}^{+\infty}p(x|s)f_\Lambda(s)\,ds}$,

οπου,

$$\int_{-\infty}^{+\infty}\frac{e^{-s}s^{x}}{x!}1(x\in\mathbb{N}_0)\frac{\mu^\alpha}{\Gamma(\alpha)}s^{\alpha-1}e^{-\mu s}1(s>0)\,ds=1(x\in\mathbb{N}_0)\frac{\mu^\alpha}{\Gamma(\alpha)x!}\int_0^\infty s^{\alpha+x-1}\exp\{-(\mu+1)s\}\,ds$$

$$=1(x\in\mathbb{N}_0)\frac{\mu^\alpha(\mu+1)^{-(\alpha+x)}}{\Gamma(\alpha)x!}\Gamma(\alpha+x)\ .$$ Αρα,

$$f(\lambda|x)=e^{-\lambda}\lambda^{x}\cancel{\Gamma(\alpha)}(\mu+1)^{\alpha+x}\cancel{\mu^{-\alpha}}\Gamma(\alpha+x)^{-1}\cancel{\mu^{\alpha}}\cancel{\Gamma(\alpha)^{-1}}\lambda^{\alpha-1}e^{-\mu\lambda}1(\lambda>0)=$$

$$=\frac{(\mu+1)^{\alpha+x}}{\Gamma(\alpha+x)}\lambda^{\alpha+x-1}e^{-(\mu+1)\lambda}1(\lambda>0).$$

(1.77) <u>Παραδειγμα</u>: Εστω $X_1,X_2,\ldots$ ανεξαρτητες με κοινη μεση τιμη και ανεξαρτητες της ακεραιας μη αρνητικης τ.μ. N. Δειξτε οτι:

$$E\{\Sigma_{i=1}^{N}X_i\}=(EN)(EX_1).$$

<u>Αποδειξη</u>: $E\{\Sigma_{i=1}^{N}X_i\}=E\{E\{\Sigma_{i=1}^{N}X_i\,|\,N\}\}=\Sigma_{n=0}^{\infty}E\{\Sigma_{i=1}^{N}X_i\,|\,N=n\}p_N(n)$

$=\Sigma_{n=0}^{\infty}E\{\Sigma_{i=1}^{n}X_i\,|\,N=n\}p_N(n)=\Sigma_{n=0}^{\infty}E\{\Sigma_{i=1}^{n}X_i\}p_N(n)=$

$=\Sigma_{n=0}^{\infty}(\Sigma_{i=1}^{n}EX_i)p_N(n)=(EX_1)\Sigma_{n=0}^{\infty}np_N(n)=(EX_1)(EN)$

(1.78) <u>Ασκηση</u>: Εστω $X_1,X_2,\ldots$ ανεξαρτητες και ισονομες τ.μ. και ανεξαρτητες της $N\sim\mathcal{P}(\lambda)$. Δειξτε οτι $M_{\Sigma_{i=1}^{N}X_i}(t)=\exp\{\lambda[M_{X_1}(t)-1]\}$.

(1.79) <u>Ασκηση</u>: Εστω οτι η δ.τ.μ. $(X,Y)\sim N_2(\mu_x,\mu_y\,|\,\sigma_x^2,\sigma_y^2\,|\,\rho)$. Δειξτε οτι:

(α) $Y|X=x\ \sim N\left(\mu_y+\rho\frac{\sigma_y}{\sigma_x}(x-\mu_x),\ \sigma_y^2(1-\rho^2)\right)$,

(β) $E(Y|X)=\mu_y+\rho\frac{\sigma_y}{\sigma_x}(X-\mu_x)$,

(γ) $E(Y-E(Y|X))^2=\sigma_y^2(1-\rho^2)$.

ΟΡΙΑΚΑ ΘΕΩΡΗΜΑΤΑ.

Θα θεωρουμε ακολουθιες τ.μ. $X_1, \dots, X_n, \dots$ οριςμενων στον ιδιο χωρο με πιθανοτητα $(\Omega, \mathcal{A}, P)$.

(1.80) Ορισμοι: Λεμε οτι η ακολουθια τ.μ. $X_1, X_2, \dots, X_n, \dots$

(i) συγκλινει σχεδον βεβαιως (σ.β.) στην τ.μ. X : $X_n \xrightarrow[n\to\infty]{\sigma.\beta.} X$,
αν $\forall \omega \in \Omega \setminus N$ $X_n(\omega) \xrightarrow[n\to\infty]{} X(\omega)$ και $P(N)=0$.

(ii) συγκλινει κατα πιθανοτητα στην τ.μ. X : $X_n \xrightarrow[n\to\infty]{P} X$,
αν $P(|X_n - X| > \varepsilon) \xrightarrow[n\to\infty]{} 0$ $\forall \varepsilon > 0$.

(iii) συγκλινει κατα κατανομη στην τ.μ. X : $X_n \xrightarrow[n\to\infty]{d} X$,
αν $F_{X_n}(x) \xrightarrow[n\to\infty]{} F_X(x)$ $\forall x \in \mathbb{R}$ σημειο συνεχειας της F_X.

(iv) συγκλινει κατα μεση ταξεως $p = 1, 2, \dots$ στην τ.μ. X : $X_n \xrightarrow[n\to\infty]{L_p} X$,
αν $E|X_n - X|^p \xrightarrow[n\to\infty]{} 0$.

(1.81) Προταση. $X_n \xrightarrow[n\to\infty]{\sigma.\beta.} X \Leftrightarrow \lim_{n\to\infty} P(|X_m - X| < \varepsilon \ \forall m \geq n) = 1 \ \forall \varepsilon > 0$.

(1.82) Λημμα. Εστω τ.μ. X και μη φθινουσα συναρτηση $\varphi : (0, +\infty) \to (0, +\infty)$, τετοια ωστε η $\varphi(X)$ ειναι τ.μ.. Τοτε,
$P(|X| > \varepsilon) \leq E\{\varphi(|X|)\}/\varphi(\varepsilon)$ $\forall \varepsilon > 0$.

Αποδ. $E\{\varphi(|X|)\} = \int_{\mathbb{R}} \varphi(|x|) f(x) dx \geq \int_{|x|>\varepsilon} \varphi(|x|) f(x) dx \geq$
$\geq \varphi(\varepsilon) \int_{|x|>\varepsilon} f(x) dx = \varphi(\varepsilon) P(|X| > \varepsilon)$.

(1.83) Πορισμα. (α) $P(|X| > \varepsilon) \leq E|X|^p/\varepsilon^p$ $\forall \varepsilon > 0, p > 0$ (Ανισοτ. Markov).
(β) $P(|X - EX| > \varepsilon) \leq \mathcal{D}(X)/\varepsilon^2$ $\forall \varepsilon > 0$ (Ανισοτητα Chebyshev).

(1.84) Προταση: Εστω ακολουθια τ.μ. $X_1, X_2, \dots, X_n, \dots$, εχουμε:

(α) $X_n \xrightarrow{\sigma.\beta.} X \Rightarrow X_n \xrightarrow{P} X$

(β) $X_n \xrightarrow{P} X \Rightarrow X_n \xrightarrow{d} X$

(γ) $X_n \xrightarrow{L_p} X \Rightarrow X_n \xrightarrow{P} X$ (για $p > 0$)

(δ) $X_n \xrightarrow{d} c \Rightarrow X_n \xrightarrow{P} c$, οπου c σταθερα.

Αποδ. Η (α) επεται απο την (1.81) και η (γ) απο την (1.83α).

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

$\Rightarrow \forall \varepsilon,\delta>0 \ \exists n_0 : P(|X_n-X|>\varepsilon)<\delta \ \forall n>n_0$.

Αρα $\forall \varepsilon,\delta>0 \ \exists n_0 : \forall n>n_0 \ F_{X_n}(x) = P(X_n - X \le x-X) =$

$= P(\{X_n-X\le x-X\}\cap\{X_n-X\ge-\varepsilon\}) + P(\{X_n-X\le x-X\}\cap\{X_n-X<-\varepsilon\})$

$\le P(-\varepsilon\le x-X) + P(X_n-X<-\varepsilon) < F_X(x+\varepsilon)+\delta$

και ομοια $F_X(x-\varepsilon) < F_{X_n}(x)+\delta$.

Αρα $\forall \varepsilon,\delta>0 \ \exists n_0 : F_X(x-\varepsilon)-\delta < F_{X_n}(x) < F_X(x+\varepsilon)+\delta \ \forall n>n_0 \Rightarrow$

$\Rightarrow \forall \varepsilon,\delta>0 \ \exists n_0 : |F_{X_n}(x)-F_X(x)| < \delta + \max\{F_X(x+\varepsilon)-F_X(x), F_X(x)-F_X(x-\varepsilon)\}$

$\xrightarrow[\delta,\varepsilon\to0]{} F_X(x)-F_X(x^-)=0 \ \forall x$ σημειο συνεχειας της F_X.

(δ) $X_n \xrightarrow{d} c \Rightarrow F_{X_n}(x) \xrightarrow[n\to\infty]{} \begin{cases} 1 & \text{αν } x>c \\ 0 & \text{αν } x<c \end{cases}$. Αρα $\forall\varepsilon>0$

$P(|X_n-c|>\varepsilon) = P(X_n-c>\varepsilon)+P(X_n-c<-\varepsilon) =$

$= 1-F_{X_n}(c+\varepsilon) + P(X_n<c-\varepsilon) \le 1-F_{X_n}(c+\varepsilon)+F_{X_n}(c-\varepsilon)$

$\xrightarrow[n\to\infty]{} 1-1+0=0$.

Παραθετουμε τωρα τα δυο βασικοτερα θεωρηματα της Θεωριας Πιθανοτητων – τον (ασθενη & ισχυρο) Νομο των Μεγαλων Αριθμων (ΝΜΑ) και το Κεντρικο Οριακο Θεωρημα (ΚΟΘ):

(1.85) Ορισμος: Λεμε οτι εχουμε μια ακολουθια ανεξαρτητων τ.μ. $X_1,\dots,X_n,\dots$ αν και μονο αν $\forall n\in\{2,3,\dots\}$ και διαφορετικους μεταξυ τους δεικτες $i_1,\dots,i_n\in\mathbb{N}$ οι τ.μ. $X_{i_1},\dots,X_{i_n}$ ειναι ανεξαρτητες.

(1.86) Θεωρημα (ΝΜΑ – Khintchin). Εστω ακολουθια ανεξαρτητων και ισονομων (α.ι.) τ.μ. $X_1,\dots,X_n,\dots$ τετοιες ωστε $E|X_1|<+\infty$, τοτε:

$\bar{X}_n := \frac{1}{n}\sum_{i=1}^n X_i \xrightarrow[n\to\infty]{P} EX_1$.

Αποδ. με την επιπλεον (για ευκολια) υποθεση οτι $EX^2<+\infty$:

Απο την ανισοτητα του Chebyshev, $\forall\varepsilon>0 \ P(|\bar{X}_n-EX_1|\ge\varepsilon) \le \frac{\mathcal{D}(\bar{X}_n)}{\varepsilon^2} =$

$= \frac{\mathcal{D}(X_1)}{\varepsilon^2 n} \xrightarrow[n\to\infty]{} 0$.

Με τις υποθεσεις του (1.86) εχουμε επισης το ισχυροτερο αποτελεσμα:

(1.87) (Ισχυρος ΝΜΑ) $\overline{X}_n \xrightarrow[n\to\infty]{\sigma.\beta.} EX_1$,

το οποιο οφειλεται στον Kolmogorov, το παραθετουμε δε χωρις αποδειξη.

(1.88) Θεωρημα (ΚΟΘ). Εστω ακολουθια ανεξαρτητων και ισονομων (α.ι.) τ.μ. $X_1, \ldots, X_n, \ldots$ τετοιες ωστε $EX_1^2 < +\infty$, τοτε:

$$\frac{\sqrt{n}\,(\overline{X}_n - EX_1)}{\sqrt{\mathcal{D}(X_1)}} \xrightarrow[n\to\infty]{d} N(0,1) .$$

Αποδ. Θετουμε $\mu := EX_1$, $\sigma^2 := \mathcal{D}(X_1)$, $Z_i := \frac{X_i - \mu}{\sigma}$, $i = 1, 2, \ldots$. Εχουμε οτι οι $Z_1, \ldots, Z_n, \ldots$ ειναι α.ι. με $EZ_1 = 0$, $\mathcal{D}(Z_1) = 1$ και $\frac{\sqrt{n}\,(\overline{X}_n - \mu)}{\sigma} = \sqrt{n}\,\overline{Z}_n = \sum_{i=1}^n \frac{1}{\sqrt{n}} Z_i$. Αρκει να δειξουμε η ροπογεννητρια του $\sqrt{n}\,\overline{Z}_n$ συγκλινει στη ροπογεννητρια της $N(0,1)$:

$$M_{\sqrt{n}\,\overline{Z}_n}(t) = E\{\exp\{-t\textstyle\sum_{i=1}^n \frac{1}{\sqrt{n}} Z_i\}\} = E\{\prod_{i=1}^n \exp\{-(t/\sqrt{n}) Z_i\}\} =$$
$$= \prod_{i=1}^n M_{Z_i}(t/\sqrt{n}) = \left(M_{Z_1}(\tfrac{t}{\sqrt{n}})\right)^n . \quad \text{Αρα,}$$

$$\log M_{\sqrt{n}\,\overline{Z}_n}(t) = n \log M_{Z_1}(\tfrac{t}{\sqrt{n}}) = n \Big\{ \log M_{Z_1}(0) + \frac{t}{\sqrt{n}} \frac{M'_{Z_1}(0)}{M_{Z_1}(0)} +$$
$$+ \frac{1}{2} \frac{t^2}{n} \frac{M''_{Z_1}(0) M_{Z_1}(0) - (M'_{Z_1}(0))^2}{(M_{Z_1}(0))^2} + o(\tfrac{1}{n}) \Big\} = \frac{1}{2}\frac{t^2}{n} + o(\tfrac{1}{n}) ,$$

εφοσον $M_{Z_1}(0) = 1$, $M'_{Z_1}(0) = -EZ_1 = 0$, $M''_{Z_1}(0) = EZ_1^2 = \mathcal{D}(Z_1) = 1$.

Αρα, $M_{\sqrt{n}\,\overline{Z}_n}(t) \xrightarrow[n\to\infty]{} \exp\{\frac{1}{2}t^2\} = M_{N(0,1)}(t)$.

(1.89) Παραδειγμα: Εστω $X_1, \ldots, X_n, \ldots$ α.ι. τ.μ. με αγνωστη κοινη σ.κ. $F(x)$, $x \in \mathbb{R}$. Οριζουμε την εμπειρικη σ.κ.

(1.90) $\mathbb{F}_n(x) := \frac{1}{n} \sum_{i=1}^n 1(X_i \le x) = \frac{1}{n} \#\{X_i \le x, i = 1, \ldots, n\}$, $\forall x \in \mathbb{R}$,

η οποια ειναι μια κλιμακωτη συναρτηση, με πηδηματα μεγεθους $\frac{1}{n}$ στις θεσεις $X_1, \ldots, X_n$. Δειξτε οτι: $\forall x \in \mathbb{R}$ σταθερο:

(α) $n\mathbb{F}_n(x) \sim \mathcal{D}(n,p)$, με $p = E\{\mathbb{F}_n(x)\} = F(x)$,

(β) $\mathbb{F}_n(x) \xrightarrow{P} F(x)$ καθως το $n \to \infty$,

(γ) $\sqrt{n}\,(\mathbb{F}_n(x) - F(x)) \xrightarrow{d} N(0, F(x)[1-F(x)])$ καθως το $n \to \infty$.

Αποδ. (α) Εστω $Y_i := 1(X_i \le x)$, $i = 1, 2, \ldots$. Εχουμε οτι $Y_1, \ldots, Y_n, \ldots$ α.ι. Bernoulli (p), με $p = E\{1(X_1 \le x)\} = F(x)$.

Αρα, $\forall\, n\in\mathbb{N}$, $n\mathbb{F}_n(x)=\sum_{i=1}^n Y_i \sim \mathcal{B}(n,p)$.

(β) $\mathbb{F}_n(x)=\frac{1}{n}\sum_{i=1}^n Y_i \xrightarrow[n\to\infty]{P} EY_1=p=F(x)$, από τον ΝΜΑ.

(γ) $\dfrac{\sqrt{n}\,(\mathbb{F}_n(x)-F(x))}{\sqrt{F(x)(1-F(x))}}=\dfrac{\sqrt{n}\,(\overline{Y}_n-EY_1)}{\sqrt{\mathcal{D}(Y_1)}}\xrightarrow[n\to\infty]{d} N(0,1)$, από το ΚΟΘ.

Παραθετουμε, χωρις αποδειξη, το ΚΟΘ στη διασταση-2:

(1.91) <u>Θεωρημα</u>: Εστω $(X_1,Y_1),\dots,(X_n,Y_n)$ α.ι. δ.τ.μ. με $EX_1^2, EY_1^2<+\infty$, και εστω $\mu_x:=EX_1$, $\mu_y:=EY_1$, $\sigma_x^2:=\mathcal{D}(X_1)$, $\sigma_y^2:=\mathcal{D}(Y_1)$, $\rho:=\dfrac{Cov(X_1,Y_1)}{\sigma_x\sigma_y}$.
Εχουμε:

$$P\left(\frac{\sqrt{n}\,(\overline{X}_n-\mu_x)}{\sigma_x}\le x,\ \frac{\sqrt{n}\,(\overline{Y}_n-\mu_y)}{\sigma_y}\le y\right)\xrightarrow[n\to\infty]{} P(X\le x, Y\le y),$$

οπου η δ.τ.μ. $(X,Y)\sim N_2(0,0\,|\,1,1\,|\,\rho)$.

(1.92) <u>Θεωρημα</u> (Cramer–Wold). Εστω η δ.τ.μ. $\underline{X}$ και η ακολουθια δ.τ.μ. $\underline{X}_1,\dots,\underline{X}_n,\dots$ στο $\mathbb{R}^k$: $\underline{X}_n\xrightarrow[n\to\infty]{d}\underline{X} \Leftrightarrow \underline{\lambda}^T\underline{X}_n\xrightarrow[n\to\infty]{d}\underline{\lambda}^T\underline{X}\ \ \forall\,\underline{\lambda}\in\mathbb{R}^k$.

Παραθετουμε τωρα ενα εξαιρετικα χρησιμο για μας, αποτελεσμα:

(1.93) <u>Θεωρημα</u> (Slutsky). Εστω ακολουθιες τ.μ. $X_1,\dots,X_n,\dots$ και $Y_1,\dots,Y_n,\dots$, τ.μ. X και σταθερα $c\in\mathbb{R}$ τετοιες ωστε: $X_n\xrightarrow[n\to\infty]{d}X$ και $Y_n\xrightarrow{P}c$. Τοτε (X_n, Y_n ακομη και εξαρτημενες):

(α) $X_n+Y_n\xrightarrow[n\to\infty]{d}X+c$,

(β) $X_nY_n\xrightarrow[n\to\infty]{d}cX$,

(γ) $\dfrac{X_n}{Y_n}\xrightarrow[n\to\infty]{d}\dfrac{X}{c}$ αν $c\neq 0$.

<u>Αποδ.</u> (α) $F_{X_n+Y_n}(z)=P(X_n+Y_n\le z, |Y_n-c|\le\varepsilon)+P(X_n+Y_n\le z, |Y_n-c|>\varepsilon)$
$\le P(X_n+Y_n\le z, -\varepsilon\le Y_n-c)+P(|Y_n-c|>\varepsilon)\le F_{X_n}(z-c+\varepsilon)+P(|Y_n-c|>\varepsilon)\ \ \forall\varepsilon>0$.
Επισης, $F_{X_n}(z-c-\varepsilon)=P(X_n\le z-c-\varepsilon, |Y_n-c|\le\varepsilon)+P(X_n\le z-c-\varepsilon, |Y_n-c|>\varepsilon)$
$\le P(X_n\le z-c-\varepsilon, Y_n-c\le\varepsilon)+P(|Y_n-c|>\varepsilon)\le F_{X_n+Y_n}(z)+P(|Y_n-c|>\varepsilon)\ \ \forall\varepsilon>0$.
Αρα, $\forall\varepsilon>0$ $F_{X_n}(z-c-\varepsilon)-P(|Y_n-c|>\varepsilon)\le F_{X_n+Y_n}(z)\le F_{X_n}(z-c+\varepsilon)+P(|Y_n-c|>\varepsilon)$.

Οποτε, $F_{X+c}(z-\varepsilon)=F_X(z-c-\varepsilon)\le \liminf\limits_{n\to\infty} F_{X_n+Y_n}(z) \le \limsup\limits_{n\to\infty} F_{X_n+Y_n}(z) \le F_X(z-c+\varepsilon)=F_{X+c}(z+\varepsilon)$ $\forall \varepsilon>0$, τετοιο ωστε $z\pm\varepsilon$ σημ. συνεχ. της F_{X+c}. Τωρα, z σημ. συνεχ. της F_{X+c} και αρα $\forall \epsilon>0\ \exists\,\delta>0 : 0<\varepsilon=|z\pm\varepsilon-z|<\delta \Rightarrow |F_{X+c}(z\pm\varepsilon)-F_{X+c}(z)|<\epsilon$

$\Rightarrow \forall \epsilon>0\ \exists\,\delta>0 : F_{X+c}(z)-\epsilon \le \underline{\lim}\, F_{X_n+Y_n}(z) \le \overline{\lim}\, F_{X_n+Y_n}(z) \le F_{X+c}(z)+\epsilon \quad (*)$

$\forall\, z\pm\varepsilon \in (z-\delta, z+\delta)$ σημεια συνεχειας της F_{X+c}. Το οτι υπαρχουν τετοια σημεια $z\pm\varepsilon$ εξασφαλιζεται απ' το οτι τα σημ. ασυνεχειας μιας συναρτησης κατανομης ειναι (το πολυ) αριθμησιμα το πληθος. Απο την $(*)$ λοιπον, στελνοντας $\epsilon \searrow 0$, εχουμε οτι $\exists \lim\limits_{n\to\infty} F_{X_n+Y_n}(z) = F_{X+c}(z)$ $\forall\, z$ σημ. συνεχ. της F_{X+c}, ο.ε.δ.

Οι αποδειξεις των (β) και (γ) ειναι αναλογες και αφινονται ως ασκησεις.

Ισχυει επισης το ακολουθο γενικοτερο θεωρημα (δεν διδουμε εδω αποδειξη):

(1.94) Θεωρημα (Mann-Wald) Εστω ακολουθια δ. τ.μ. $\underline{X}_1, \dots, \underline{X}_n, \dots$ (στον $\mathbb{R}^k$) στο ιδιο χωρο πιθανοτητα, δ.τ.μ. $\underline{X}$ και συνεχης συναρτηση g στον $\mathbb{R}^k$.

Τοτε: (α) $\underline{X}_n \xrightarrow{\sigma.\beta.} \underline{X} \Rightarrow g(\underline{X}_n) \xrightarrow{\sigma.\beta.} g(\underline{X})$,

(β) $\underline{X}_n \xrightarrow{P} \underline{X} \Rightarrow g(\underline{X}_n) \xrightarrow{P} g(\underline{X})$,

(γ) $\underline{X}_n \xrightarrow{d} \underline{X} \Rightarrow g(\underline{X}_n) \xrightarrow{d} g(\underline{X})$, καθως το $n\to\infty$.

Το ακολουθο θεωρημα ειναι εξαιρετικα χρησιμο για τη σταθεροποιηση της ασυμπτωτικης διασπορας συναρτησεων τ.μ..

(1.95) Θεωρημα: Εστω ακολουθια τ.μ. $X_1, \dots, X_n, \dots$ τετοια ωστε

$\alpha_n(X_n - m) \xrightarrow[n\to\infty]{d} N(0,\tau^2)$, οπου $\tau^2>0$, $m, \alpha_1, \alpha_2, \dots \in \mathbb{R}$ σταθερες με $\alpha_n \xrightarrow[n\to\infty]{} +\infty$. Επισης, εστω συναρτηση $g:\mathbb{R}\to\mathbb{R}$ συνεχως διαφορισιμη στη θεση $m\in\mathbb{R}$, με $g'(m)\neq 0$. Τοτε,

$$\alpha_n\left(g(X_n)-g(m)\right) \xrightarrow[n\to\infty]{d} N\left(0, [g'(m)]^2\tau^2\right).$$

Αποδ. Απο (1.93) $X_n - m = \frac{1}{\alpha_n}\left[\alpha_n(X_n-m)\right] \xrightarrow{d} 0$ και αρα (απ' το (1.84δ)) $X_n \xrightarrow{P} m$. Επισης απ' το Θεωρημα της μεσης τιμης:

$$g(X_n)-g(m) = g'(X_n^*)(X_n-m) \quad \text{με } |X_n^*-m|<|X_n-m|$$

και αρα (εφοσον $X_n \xrightarrow{P} m$) $X_n^* \xrightarrow{P} m \Rightarrow g'(X_n^*) \xrightarrow{P} g(m)$, απο το (1.94β). Αρα, απ' το Θεωρημα του Slutsky:

$$\alpha_n[g(X_n)-g(m)] = g'(X_n^*)\left[\alpha_n(X_n-m)\right] \to g'(m)N(0,\tau^2) \stackrel{d}{=}$$
$$\stackrel{d}{=} N\left(0, [g'(m)]^2\tau^2\right).$$

(1.96) Παραδειγμα (συνεχεια του (1.89)) Δειξτε οτι $\forall x \in \mathbb{R}$ σταθερο:

$$2\sqrt{n}\left[\text{τοξημ}\sqrt{\mathbb{F}_n(x)} - \text{τοξημ}\sqrt{F(x)}\right] \xrightarrow[n\to\infty]{d} N(0,1).$$

Απαντηση: Το αποτελεσμα επεται απο το (1.89γ) και το (1.95) με $g(z) = 2\,\text{τοξημ}\sqrt{z}$, με $g'(z) = (z(1-z))^{-1/2}$ και αρα

$$\sqrt{n}\left[g(\mathbb{F}_n(x)) - g(F(x))\right] \xrightarrow[n\to\infty]{d} N(0, [g'(F(x))]^2 F(x)[1-F(x)]) \stackrel{d}{=} N(0,1).$$

(1.97) Παραδειγμα. Εστω $X_1, \dots, X_n, \dots$ α.ι. με $\mu \equiv EX_1$, $\sigma^2 \equiv \mathcal{D}(X_1)$, $\mu_4 \equiv E(X_1-\mu)^4 < +$

Δειξτε οτι, για $\bar{X}_n$ και S_n^2 οπως στο (1.62), εχουμε:

(α) $\sqrt{n}(S_n^2 - \sigma^2) \xrightarrow[n\to\infty]{d} N(0, \mu_4 - \sigma^4)$

(β) $\sqrt{n}(S_n - \sigma) \xrightarrow[n\to\infty]{d} N\left(0, \frac{\mu_4-\sigma^4}{4\sigma^2}\right)$

Αποδ (α) $(n-1)S_n^2 = \Sigma_{i=1}^n (X_i - \bar{X}_n)^2 = \Sigma_{i=1}^n [(X_i-\mu) - (\bar{X}_n - \mu)]^2 =$

$= \Sigma_{i=1}^n \{(X_i-\mu)^2 + (\bar{X}_n-\mu)^2 - 2(\bar{X}_n-\mu)(X_i-\mu)\} = \Sigma_{i=1}^n (X_i-\mu)^2 + n(\bar{X}_n-\mu)^2$

$- 2n(\bar{X}_n-\mu)^2 = \Sigma_{i=1}^n (X_i-\mu)^2 - n(\bar{X}_n-\mu)^2$, δηλαδη,

(1.98) $\Sigma_{i=1}^n (X_i - \bar{X}_n)^2 = \Sigma_{i=1}^n (X_i-\mu)^2 - n(\bar{X}_n-\mu)^2.$

Αρα, $\sqrt{n}(S_n^2 - \sigma^2) = \frac{n}{n-1}\left\{\sqrt{n}(\hat{\sigma}_n^2 - \sigma^2) - \frac{1}{\sqrt{n}}[\sqrt{n}(\bar{X}_n-\mu)]^2 + \frac{1}{\sqrt{n}}\sigma^2\right\}$ (*)

οπου, $\hat{\sigma}_n^2 := \frac{1}{n}\Sigma_{i=1}^n (X_i-\mu)^2$.

Τωρα, εφοσον $Y_i \equiv (X_i-\mu)^2$, $i=1,2,\dots$ α.ι. με $EY_1 = \sigma^2$ και $\mathcal{D}(Y_1) = \mu_4 - \sigma^4$, εχουμε απο το ΚΟΘ, οτι:

$\sqrt{n}(\hat{\sigma}_n^2 - \sigma^2) \xrightarrow[n\to\infty]{d} N(0, \mu_4 - \sigma^4)$. (**)

Επισης, απο το ΚΟΘ, $\sqrt{n}(\bar{X}_n-\mu) \xrightarrow{d} N(0,\sigma^2) \stackrel{d}{=} \sigma N(0,1)$

και αρα, απο το (1.94γ), εχουμε: $[\sqrt{n}(\bar{X}_n-\mu)]^2 \xrightarrow{d} \sigma^2 N(0,1)^2 = \sigma^2 \chi_1^2$

και αρα απο το θεωρημα του Slutsky και απο την (1.84δ):

$\frac{1}{\sqrt{n}}[\sqrt{n}(\bar{X}_n-\mu)]^2 \xrightarrow{d} 0 \Rightarrow \frac{1}{\sqrt{n}}[\sqrt{n}(\bar{X}_n-\mu)]^2 \xrightarrow{P} 0$. (***)

Τελος, απο το θεωρημα του Slutsky και τις σχεσεις με "αστρα", εχουμε: $\sqrt{n}(S_n^2 - \sigma^2) \xrightarrow{d} N(0, \mu_4 - \sigma^4)$

(β) Το αποτελεσμα επεται απο (1.95) με $g(z) := \sqrt{z}$, $g'(z) = (2z^{1/2})^{-1}$.

(1.99) Ασκηση: Εστω $X_1, \dots, X_n, \dots$ α.ι. $\mathcal{P}(\lambda)$. Δειξτε οτι:

$$2\sqrt{n}(\sqrt{\bar{X}_n} - \sqrt{\lambda}) \xrightarrow[n\to\infty]{d} N(0,1).$$

2. ΒΑΣΙΚΕΣ ΑΡΧΕΣ ΤΗΣ ΣΤΑΤΙΣΤΙΚΗΣ ΣΥΜΠΕΡΑΣΜΑΤΟΛΟΓΙΑΣ.

Εστω χωρος με πιθανοτητα $(\Omega, \mathcal{A}, P)$ με αγνωστη πιθανοσυναρτηση $P \in \mathcal{P}$ — καποια οικογενεια πιθανοσυναρτησεων.

Στα πλαισια της παραμετρικης στατιστικης υποθετουμε οτι η οικογενεια $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$ με $\Theta \subseteq \mathbb{R}^k$, $k \in \mathbb{N}$, οπου η καθε πιθανοσυναρτηση P_θ εξαρταται μεν απο την αγνωστη παραμετρο $\theta \in \Theta$ αλλα κατα τα αλλα ειναι γνωστη συναρτηση.

(2.1) Παραδειγμα: (α) $P_\theta(B) = \int_B (2\pi\sigma^2)^{-1} \exp\{-(x-\mu)^2/2\sigma^2\}\, dx$, $B \in \mathcal{B}$, οπου $\theta = (\mu, \sigma^2) \in \Theta = \mathbb{R} \times (0, +\infty) \subseteq \mathbb{R}^2$.

(β) $P_\theta(B) = \int_B \{(2\pi)^n |\Sigma|\}^{-1/2} \exp\{-(\underline{x}-\underline{\mu})^T \Sigma^{-1} (\underline{x}-\underline{\mu})/2\}\, d\underline{x}$, $B \in \mathcal{B}^n$, οπου $\theta = (\underline{\mu}, \Sigma) = (\mu_1, \dots, \mu_n, \sigma_1^2, \dots, \sigma_n^2, \rho_{ij}$ με $i, j = 1, \dots, n$ και $i < j)$
$\in \Theta = \mathbb{R}^n \times (0, +\infty)^n \times [-1, 1]^{n(n-1)/2} \subseteq \mathbb{R}^{n(n+3)/2}$.

(γ) $P_\theta(B) = \sum_{x \in B \cap \mathbb{N}_0} e^{-\lambda} \lambda^x / x!$, $B \subseteq \mathbb{R}$, οπου $\theta = \lambda \in \Theta = (0, +\infty)$.

(δ) $P_\theta(B) = (1-\alpha) \int_B \lambda \exp\{-2\lambda |x-m|\}\, dx + \alpha \sum_{x \in B \cap \mathbb{N}} p(1-p)^{x-1}$, $B \in \mathcal{B}$, οπου $\theta = (\lambda, m, p, \alpha) \in \Theta = (0, +\infty) \times \mathbb{R} \times [0, 1]^2$.

Οι βασικοι σκοποι της παραμετρικης στατιστικης ειναι
η εκτιμηση της αγνωστης παραμετρου $\theta_0 \in \Theta \subseteq \mathbb{R}^k$,
η κατασκευη (στοχαστικων) περιοχων εμπιστοσυνης οι οποιες να περιεχουν την αγνωστη παραμετρο $\theta_0 \in \Theta$ με προκαθορισμενη πιθανοτητα $1-\alpha$, $\alpha \in (0,1)$,
η κατασκευη ελεγχων της υποθεσεως $H: \theta_0 \in \Theta_0$ εναντι καποιας εναλλακτικης $K: \theta_0 \in \Theta_1$, οπου $\Theta_0, \Theta_1 \subseteq \Theta$ και $\Theta_0 \cap \Theta_1 = \emptyset$,
καθως και την αναπτυξη κριτηριων καταλληλοτητος των εκτιμητριων, διαστηματων ή γενικοτερα περιοχων εμπιστοσυνης και των ελεγχων υποθεσεων. Οι μεθοδοι της στατιστικης αναπτυσσονται ετσι ωστε να ειναι βελτιστοι (ή σχεδον) οσον αφορα σ' αυτα τα προκαθορισμενα κριτηρια καταλληλοτητος.

Εστω τυχαια μεταβλητη $X : \Omega \longrightarrow \mathbb{R}^m$, με συναρτηση κατανομης $F_\theta(x) \equiv F(x|\theta) := P_\theta(X \le x)$, $\underline{x} \in \mathbb{R}^m$, $\theta \in \Theta \subseteq \mathbb{R}^k$. Σε καθε πιθανοσυναρτηση $P_\theta \in \mathcal{P} \equiv \mathcal{P}_\Theta := \{P_\theta, \theta \in \Theta\}$ αντιστοιχουμε, κατ' αυτον τον τροπο, μια συναρτηση κατανομης $F_\theta \in \mathcal{F} \equiv \mathcal{F}_\Theta := \{F_\theta, \theta \in \Theta\}$. Υπαρχει μια ενα-προς-ενα αντιστοιχια μεταξυ των χωρων $\mathcal{P}$ και $\mathcal{F}$ και μπορουμε να τους ταυτιζουμε.

(2.2) <u>Ορισμος.</u> Η παραμετριση των χωρων $\mathcal{P}_\Theta$ ή $\mathcal{F}_\Theta$ των παραμετρικων μοντελων της αβεβαιοτητος του πειραματος $(\Omega, \mathcal{A})$ λεγεται <u>προσδιορισιμη</u> αν και μονο αν $\theta_1 \ne \theta_2 \Rightarrow P_{\theta_1} \ne P_{\theta_2}$ $\forall\, \theta_1, \theta_2 \in \Theta$.

(2.3) <u>Ορισμος</u>. Τα παραμετρικα μοντελα $\mathcal{P}_\Theta$ ή $\mathcal{F}_\Theta$ λεγονται <u>ομαλα μοντελα</u> αν και μονο αν

$\forall\, P_\theta \in \mathcal{P}_\Theta$ $\exists\, f_\theta(\cdot) \equiv f(\cdot|\theta)$ πυκνοτητα πιθανοτητος της αντιστοιχης F_θ,

ή $\forall\, P_\theta \in \mathcal{P}_\Theta$ $\exists\, p_\theta(\cdot) \equiv p(\cdot|\theta)$ συναρτηση μαζας πιθανοτητος της αντιστοιχης F_θ, με κοινο αριθμησιμο φορεα $S := \{x_1, x_2, \dots\}$ ο οποιος <u>δεν</u> εξαρταται απο την παραμετρο $\theta \in \Theta$.

Σε οσα ακολουθουν, δεχομαστε οτι η παραμετριση που εχουμε κανει ειναι <u>προσδιορισιμη</u> και οτι τα παραμετρικα μοντελα που εχουμε ειναι <u>ομαλα</u>, δηλαδη, ειτε ολα διακριτα με κοινο φορεα, ανεξαρτητο της αγνωστης παραμετρου, ειτε ολα απολυτα συνεχη.

Εστω λοιπον τ.μ. X στο χωρο με πιθανοτητα $(\Omega, \mathcal{A}, P)$, για τον οποιο υποθετουμε οτι το μοντελο P της αβεβαιοτητος του πειραματος $(\Omega, \mathcal{A})$ ειναι παραμετρικο, δηλαδη, $P = P_{\theta_0} \in \mathcal{P}_\Theta$ ή ισοδυναμα $F = F_{\theta_0} \in \mathcal{F}_\Theta$, για καποιο <u>αγνωστο</u> $\theta_0 \in \Theta$. Σκοπος μας ειναι η συναγωγη συμπερασματων (εκτιμησεις, ελεγχοι, περιοχες εμπιστοσυνης) για την <u>αγνωστη</u> παραμετρο και αρα ο προσδιορισμος του νομου F της αβεβαιοτητος του πειραματος. Η στατιστικη, με την οποια θα ασχοληθουμε εδω, επιτυγχανει αυτο το σκοπο με την επαγωγη, απο συγκεκριμενα δεδομενα — παρατηρησεις — $X(\omega) = (X_1(\omega), \dots, X_n(\omega))$ απο την εκτελεση $\omega \in \Omega$ του πειραματος $(\Omega, \mathcal{A})$, με $X \sim F_\theta$, στο προσδιορισμο του (γενικου) μοντελου F_θ.

Εδω, θεωρουμε τις τ.μ. $X_1, \dots, X_n$ ανεξαρτητες και ισονομες (α.ι.), δηλαδη, $F_{\underline{X}}(\underline{x}|\theta) = \prod_{i=1}^{n} F_{X_i}(x_i|\theta) = \prod_{i=1}^{n} F_{X_1}(x_i|\theta)$, $x_1, \dots, x_n \in \mathbb{R}^m$ και συνηθως θα εχουμε $m=1$, οι δε στατιστικες μεθοδοι που θα αναπτυξουμε θα ισχυουν και για $m>1$, με απολυτα φυσικες επεκτασεις.

(2.4) Σε οτι ακολουθει, λοιπον, θεωρουμε: ανεξαρτητες και ισονομες τ.μ. $X_1, \dots, X_n \sim F_\theta$, $\theta \in \Theta \subseteq \mathbb{R}^k$, ορισμενες στον ιδιο χωρο με πιθανοτητα $(\Omega, \mathcal{A}, F_\theta)$, οπου η σ.κ. $F_\theta \in \mathcal{F}_\Theta := \{F_\theta, \theta \in \Theta\}$ — μια ομαλη παραμετρικη οικογενεια κατανομων της οποιας η παραμετριση ειναι προσδιορισιμη. Τις τ.μ. $X_1, \dots, X_n$ καλουμαι ανεξαρτητο στοχαστικο δειγμα απο την παραμετρικη κατανομη F_θ, με αγνωστη, προσδιορισιμη, παραμετρο $\theta \in \Theta$. Το υποσυνολο $\Theta \subseteq \mathbb{R}^k$ καλουμαι παραμετρικο χωρο.

Στην πραξη, τα δεδομενα ή το δειγμα ειναι καποιες $n \in \mathbb{N}$ — μεγεθος δειγματος — συγκεκριμμενες ποσοτητες $\underline{x} := (x_1, \dots, x_n)$ τις οποιες και θεωρουμε σαν το συγκεκριμενο αποτελεσμα $\omega \in \Omega$ του πειραματος $(\Omega, \mathcal{A})$, δηλαδη, θεωρουμε οτι $\underline{X}(\omega) := (X_1(\omega), \dots, X_n(\omega)) = (x_1, \dots, x_n) = \underline{x}$, για καποιο $\omega \in \Omega$ το οποιο συνεβη.

Οι στατιστικες μεθοδοι τις οποιες θα αναπτυξουμε θα αφορουν τα γενικα στοχαστικα ανεξαρτητα δειγματα $(X_1, \dots, X_n)$, και θα ειναι λοιπον εφαρμοσιμες για καθε συγκεκριμενη πραγματοποιηση τους $(x_1, \dots, x_n) = (X_1(\omega), \dots, X_n(\omega))$, $\omega \in \Omega$. Ειναι αυτη ακριβως η θεωρηση της αντιπροσωπευτικοτητος του δειγματος που επιτρεπει τη στατιστικη επαγωγη απο το συγκεκριμενο δειγμα στο γενικο πληθυσμο ή μοντελο F_θ που το δημιουργησε.

(2.5) Καθε συναρτηση μονο του στοχαστικου δειγματος (ανεξαρτητη, δηλαδη, των παραμετρων) λεγεται <u>στατιστικη</u> <u>συναρτηση</u>.

Οι εκτιμητριες της παραμετρου $\theta \in \Theta$, οι ελεγχοι υποθεσεων σχετικα μετην παραμετρο θ και οι περιοχες εμπιστοσυνης για την παραμετρο θ, ειναι ολες μια ή περισσοτερες στατιστικες συναρτησεις και αρα εν γενει τυχαιες μεταβλητες και οι ιδιες.

(2.6) <u>Παραδειγμα</u>: Εστω στοχαστικο δειγμα $X_1,\dots,X_n$ α.κ. $N(\theta,1)$, δηλαδη, $\mathcal{F}_\Theta = \{ F_\theta(\cdot) = \Phi(\cdot - \theta), \ \theta \in \Theta = \mathbb{R} \}$.

Παρατηρουμε οτι η παραμετριση ειναι προσδιορισιμη και η οικογενεια των παραμετρικων μοντελων $\mathcal{F}_\Theta$ που θεωρουμε ειναι ομαλη.

(a) <u>Εκτιμητικη</u>: Μας ενδιαφερει να προσδιορισουμε την αληθη αγνωστη τιμη θ_0 της παραμετρου $\theta \in \Theta$.

Θα αναπτυξουμε αργοτερα διαφορες μεθοδους, και κριτηρια καταλληλοτητος, για την κατασκευη εκτιμητριων. Προς το παρον απευθυνομενοι στη διαισθηση μονο, παρατηρουμε οτι εφοσον το θ ειναι η μεση τιμη της κατανομη F_θ, μια λογικη εκτιμητρια της αγνωστης θ_0 ειναι ο μεσος ορος του δειγματος:

(2.7) $T_1(\underline{X}) \equiv T_1(X_1,\dots,X_n) := \overline{X}_n := \frac{1}{n}\sum_{i=1}^{n} X_i$.

Φυσικα με την ιδια λογικη, εφοσον $\theta = F_\theta^{-1}(1/2)$, δηλαδη ειναι η μεση της κατανομης F_θ, μια λογικη εκτιμητρια της αγνωστης παραμετρου θ ειναι η μεση του δειγματος:

(2.8) $$T_2(\underline{X}) \equiv \hat{X}_n := \begin{cases} X_{nk}, & \text{οπου } k = [n/2]+1, \text{ αν } n \text{ περιττος}, \\ \frac{1}{2}\{X_{nk} + X_{n(k+1)}\}, & \text{οπου } k = n/2, \text{ αν } n \text{ αρτιος}, \end{cases}$$

οπου X_{ni} ειναι η i διατεταγμενη στατιστικη, δηλαδη, η i μικροτερη απο τις $X_1, \dots, X_n$, $i=1,2,\dots,n$.
Το ποια απ' τις δυο ειναι "καλυτερη" εκτιμητρια και με ποια κριτηρια θα το συζητησουμε αργοτερα.

(β) Διαστηματα Εμπιστοσυνης: Θελουμε να βρουμε ενα στοχαστικο διαστημα $[L(\underline{X}), U(\underline{X})]$, μικρου μηκους, τετοιο ωστε:

$$P_{\theta_0}(\theta_0 \in [L,U]) \geq 1-\alpha, \quad \alpha \in [0,1].$$

Προφανως αυτη η μεθοδος εκτιμησης ειναι γενικοτερη της σημειακης εκτιμησης του μερους (α), διδει δε περισσοτερη πληροφορηση για την αγνωστη παραμετρο θ_0.

Μπορουμε να βασισουμε την κατασκευη ενος διαστηματος εμπιστοσυνης σε μια σημειακη εκτιμητρια, π.χ., την (2.7), ως εξης: εστω $L(\underline{X}) := \overline{X}_n - c_\alpha$, $U(\underline{X}) := \overline{X}_n - c_\alpha$, εχουμε

$$P_{\theta_0}(\theta_0 \in [\overline{X}_n - c_\alpha, \overline{X}_n + c_\alpha]) \geq 1-\alpha \iff$$
$$P_{\theta_0}(-c_\alpha \leq \overline{X}_n - \theta_0 \leq c_\alpha) \geq 1-\alpha \iff P_{\theta_0}(|\overline{X}_n - \theta_0| \leq c_\alpha) \geq 1-\alpha$$
$$\iff P_{\theta_0}(|\sqrt{n}(\overline{X}_n - \theta_0)| \leq c_\alpha\sqrt{n}) \geq 1-\alpha \iff P(|N(0,1)| \leq c_\alpha\sqrt{n}) \geq 1-\alpha$$
$$\iff \Phi(c_\alpha\sqrt{n}) - \Phi(-c_\alpha\sqrt{n}) \geq 1-\alpha \iff 2\Phi(c_\alpha\sqrt{n}) - 1 \geq 1-\alpha \iff$$
$$\iff \Phi(c_\alpha\sqrt{n}) \geq 1-\alpha/2 \iff c_\alpha \geq \Phi^{-1}(1-\alpha/2)/\sqrt{n} = z(1-\alpha/2)/\sqrt{n}.$$

Το μικροτερου λοιπον μηκους (συμμετρικο) διαστημα εμπιστοσυνης, με προκαθορισμενο επιπεδο εμπιστοσυνης $1-\alpha$, ειναι το $\left[\overline{X}_n - \frac{z(1-\alpha/2)}{\sqrt{n}}, \overline{X}_n - \frac{z(1-\alpha/2)}{\sqrt{n}}\right]$.

(γ) Ελεγχος Υποθεσεως: Εστω οτι υποψιαζομαστε οτι $\theta_0 = 0$, και θελουμε να ελεγξουμε την υποθεση $H: \theta = 0$ εναντι της εναλλακτικης $K: \theta = 1$. Μια διαισθητικα ικανοποιητικη στατιστικη συναρτηση ελεγχου ειναι η

$$\phi(\underline{X}) = \begin{cases} 1, \text{ δηλαδη, απορριπτουμε την } H \text{ αν } \overline{X}_n > c \\ 0, \text{ δηλαδη, δεχομαστε την } H \text{ αν } \overline{X}_n \leq c. \end{cases}$$

Ειναι δυνατον να γινουν δυο ειδων λαθη με μια τετοια αποφαση:
Λαθος τυπου I: να απορριψουμε την H ενω ειναι η ορθη,

και αυτο μπορει να συμβη με πιθανοτητα $P_{\theta=0}(\bar{X}_n > c)$, ή
λαθος τυπου II: να δεχθουμε την H ενω η K ειναι η ορθη,
και αυτο μπορει να συμβη με πιθανοτητα $P_{\theta=1}(\bar{X} \le c)$.
Συνηθως διαλεγουμε τη κρισιμη σταθερα c ουτως
ωστε η πιθανοτητα λαθους τυπου I να ειναι προκαθορισμενη, δηλαδη, $P_{\theta=0}(\bar{X}_n > c) = \alpha$, για καποιο $\alpha \in (0,1)$,
δηλαδη $P(N(0,1) > c\sqrt{n}) = \alpha \Leftrightarrow 1 - \Phi(c\sqrt{n}) = \alpha \Leftrightarrow$
$\Leftrightarrow \Phi(c\sqrt{n}) = 1-\alpha \Leftrightarrow c = c_\alpha = z(1-\alpha)/\sqrt{n}$. Το λαθος
τυπου II, κρατιεται χαμηλο μεσω καλης εκλογης της
στατιστικης συναρτησης του ελεγχου – την οποια εδω προκαθωρισαμε διαισθητικα μονο. Αρα ενας ελεγχος μεγεθους $\alpha \in (0,1)$
της $H: \theta=0$ vs $K: \theta=1$ ειναι ο $\phi(\underline{X}) = 1(\bar{X}_n > z(1-\alpha)/\sqrt{n})$,
θα δουμε μαλιστα αργοτερα οτι αυτος ειναι ο βελτιστος του μεγεθους
του, δηλαδη ελαχιστοποιει το λαθος τυπου II, παντα, για το μοντελο
που εχουμε υποθεσει και την υποθεση που ελεγχουμε εδω.

ΣΤΟΙΧΕΙΑ ΘΕΩΡΙΑΣ ΑΠΟΦΑΣΕΩΝ

Θα εισαγουμε τωρα ορισμενες εννοιες απο την θεωρια αποφασεων, οι οποιες ειναι χρησιμες για τη συνθετικη οργανωση των διαφορων κριτηριων καταλληλοτητος πολλων μαζυ στατιστικων μεθοδων.

Θα συμβολιζουμε με $\mathcal{X}$ τον δειγματικο χωρο και για μας
εδω $\mathcal{X} \subseteq \mathbb{R}^n$, και με A τον χωρο των αποφασεων, που
στην περιπτωση της σημειακης εκτιμητικης ειναι $A \subseteq \Theta$,
στην περιπτωση των περιοχων εμπιστοσυνης ειναι $A \subseteq P_\Omega$
και στην περιπτωση των ελεγχων αποφασεως ειναι $A = [0,1]$.
Καθε συναρτηση $d: \mathcal{X} \ni \underline{x} \longmapsto d(\underline{x}) \in A$ λεγεται
συναρτηση αποφασεως. Για παραδειγμα στην περιπτωση του
(2.6α) $d(\underline{X}) = T_1(\underline{X})$, με $a = d(\underline{x}) = T_1(\underline{x}) = \frac{1}{n}\sum_{i=1}^n x_i \in A = \Theta$,
στην περιπτωση του (2.6β) $a = d(\underline{x}) = \left[\bar{x}_n - \frac{z(1-\alpha/2)}{\sqrt{n}}, \bar{x}_n + \frac{z(1-\alpha/2)}{\sqrt{n}}\right]$
$\in A \subseteq P_\Theta$, στην περιπτωση του (2.6γ) $a = d(\underline{x}) = \phi(x) \in A = [0,1]$.

(2.9) <u>Ορισμος.</u> Καθε συναρτηση $L : \Theta \times A \ni (\theta, a) \longmapsto L(\theta, a) \in [0, +\infty)$ καλειται <u>συναρτηση απωλειας</u> του προβληματος.

Για τα στατιστικα προβληματα, τα $a \in A$ που μας ενδιαφερουν ειναι της μορφης $a = d(\underline{x}) = d(\underline{X}(\omega))$ για $\omega \in \Omega$. Αρα, εν γενει, μας ενδιαφερουν οι τ.μ. $L(\theta, d(\underline{X}))$.

Για να συγκρινουμε λοιπον την αποδοση δυο συναρτησεων αποφασεως d_1, d_2, πρεπει να συγκρινουμε δυο τυχαιες μεταβλητες, τις $L(\theta, d_1(\underline{X}))$ και $L(\theta, d_2(\underline{X}))$. Ενας απλος, και συχνα αποτελεσματικος τροπος, για να γινη μια τετοια συγκριση ειναι να συγκρινουμε τις μεσες τιμες (αν υπαρχουν) των κατανομων αυτων των τ.μ.

(2.10) <u>Ορισμος.</u> Η συναρτηση $R(\theta, d) := E_\theta\{L(\theta, d(\underline{X}))\} \in [0, +\infty)$, για $\theta \in \Theta$, $d \in \mathcal{D}$ — το συνολο των συναρτησεων αποφασεως του προβληματος — καλειται <u>συναρτηση κινδυνου</u> του προβληματος.

(2.11) <u>Παραδειγμα: Εκτιμητικη:</u> Εστω $\underline{X}_1, \ldots, \underline{X}_n$ α.ι. $f(\underline{x}|\underline{\theta})$, $\underline{x} \in \mathbb{R}^m$, $\underline{\theta} \in \Theta \subseteq \mathbb{R}^k$. Το προβλημα ειναι η εκτιμηση της αγνωστης παραμετρου $\underline{\theta} \in \Theta$, και αρα ο χωρος των συναρτησεων αποφασεως ειναι το συνολο $\mathcal{D}$ ολων των στατιστικων συναρτησεων $d : \mathcal{X} \subseteq \mathbb{R}^{mn} \longrightarrow A = \Theta \subseteq \mathbb{R}^k$.

Για συναρτηση απωλειας συχνα χρησιμοποιειται η τετραγωνικη μορφη: $L(\underline{\theta}, d(\underline{x})) := (d(\underline{x}) - \underline{\theta})^T \underline{\underline{W}} (d(\underline{x}) - \underline{\theta})$, οπου ο $k \times k$ πινακας ειναι συμμετρικος θετικα ορισμενος,

με αντιστοιχη συναρτηση κινδυνου:

$$R(\theta, d) = E_\theta\{L(\underline{\theta}, d(\underline{X}))\} = E_\theta\{(d(\underline{X}) - \underline{\theta})^T \underline{W} (d(\underline{X}) - \underline{\theta})\}$$

$$= \int_{\mathbb{R}^m} (d(\underline{x}) - \underline{\theta})^T \underline{W} (d(\underline{x}) - \underline{\theta}) f(\underline{x}|\underline{\theta}) d\underline{x}.$$

Στην ενδιαφερουσα περιπτωση που $k = 1$,

$L(\theta, d(\underline{x})) := (d(\underline{x}) - \theta)^2$, με αντιστοιχη συν. κινδυνου:

(2.12) $$R(\theta, d) = E_\theta(d(\underline{X}) - \theta)^2 =: MT\Sigma_\theta(d),$$

το <u>μεσο τετραγωνικο σφαλμα</u>.

(2.13) <u>Λημμα:</u> $MT\Sigma_\theta(d) = \mathcal{D}_\theta(d(\underline{X})) + [E_\theta(d(\underline{X})) - \theta]^2 \ \forall (\theta, d) \in \Theta \times \mathcal{D}$.

Δηλαδη, ο κινδυνος, μετρουμενος με το ΜΤΣ, ισουται με το αθροισμα της $\mathcal{D}_\theta(d(\underline{X}))$, που μετρα την ελλειψη ακριβειας της εκτιμητριας $d(\underline{X})$ και της $[b_\theta(d)]^2$, οπου η μεροληψια (ή βια - bías)

(2.14) $b_\theta(d) := E_\theta[d(\underline{X})] - \theta$, μετρα το συστηματικο σφαλμα στην εκτιμηση.

Αποδ. $ΜΤΣ_\theta(d) = E_\theta(d(\underline{X})-\theta)^2 = E_\theta\{(d(\underline{X})-E_\theta d(\underline{X})) + b_\theta(d)\}^2$
$= E_\theta\{(d(\underline{X})-E_\theta d(\underline{X}))^2 + b_\theta(d)^2 + 2b_\theta(d)[d(\underline{X})-E_\theta d(\underline{X})]\} =$
$= \mathcal{D}_\theta(d(\underline{X})) + [b_\theta(d)]^2 + 2b_\theta(d)[E_\theta d(\underline{X}) - E_\theta d(\underline{X})] = \mathcal{D}_\theta(d(\underline{X})) + [b_\theta(d)]^2.$

(2.15) Ορισμος: Αν $E_\theta d(\underline{X}) = \theta \ \forall \theta \in \Theta$, δηλαδη, $b_\theta(d) = 0 \ \forall \theta\in\Theta$, η εκτιμητρια d καλειται αμεροληπτη για την θ.
Το συνολο των αμεροληπτων εκτιμητριων για την παραμετρο $\theta \in \Theta$, συμβολιζουμε με $\mathcal{U}_\Theta$ (unbiased estimators), και εχουμε οτι

(2.16) $ΜΤΣ_\theta(d) = \mathcal{D}_\theta(d(\underline{X})) \quad \forall\theta\in\Theta, \ \forall d \in \mathcal{U}_\Theta$.

(2.17) Παραδειγμα: Εστω $X_1, \dots, X_n$ α.ι. $\mathcal{U}(0,\theta)$, $\theta\in\Theta = (0,+\infty) \subseteq \mathbb{R}$.
Εκτιμουμαι την αγνωστη παραμετρο θ με την διαισθητικα καταλληλη εκτιμητρια $d(\underline{X}) := X_{nn} = \max\{X_i, \ i=1,\dots,n\}$.

(α) Δειξτε οτι η εκτιμητρια d δεν ειναι αμεροληπτη και βρητε τη συναρτηση μεροληψιας της $b_\theta(d)$, $\theta\in\Theta$.

(β) Υπολογιστε τη συναρτηση ακριβειας $\mathcal{D}_\theta(d(\underline{X}))$, $\theta\in\Theta$, της εκτιμητριας d.

(γ) Υπολογιστε τη συναρτηση του μεσου τετραγωνικου σφαλματος $ΜΤΣ_\theta(d)$, $\theta\in\Theta$ της εκτιμητριας d.

(δ) Χρησιμοποιειστε για συναρτηση απωλειας την $L(\theta, d(\underline{x})) :=$
$:= |d(\underline{x}) - \theta|$, $\underline{x} \in \mathcal{X} = (0,\theta)^n \subseteq \mathbb{R}^n$, $\theta \in \Theta = (0,+\infty)$,
και υπολογιστε την αντιστοιχη συναρτηση κινδυνου $R(\theta,d)$, $\theta\in\Theta$, για την εκτιμητρια d που επιλεξαμε.

Απαντηση: Εστω $Z_i := X_i/\theta$, $i=1,\dots,n$, τοτε $Z_1,\dots,Z_n$ α.ι. $\mathcal{U}(0,1)$, και αρα, απο την (1.57δ), $Z_{nn} = X_{nn}/\theta \sim \mathcal{B}(n,1)$, δηλ.,
$X_{nn} \sim f(x|\theta) = n\theta^{-n}x^{n-1}\mathbb{1}(0<x<\theta)$,

(α) $E_\theta X_{nn} = \int_{-\infty}^{+\infty} x f(x|\theta)\,dx = \int_0^\theta n\theta^{-n} x^n\,dx = n\theta \int_0^1 z^n dz =$

$= \frac{n\theta}{n+1} \quad \forall \theta \in \Theta$ και άρα

$b_\theta(d) = E_\theta X_{nn} - \theta = \frac{n\theta}{n+1} - \theta = -\frac{\theta}{n+1} \quad \forall \theta \in \Theta.$

(β) $\mathcal{D}_\theta(d(\underline{X})) = E_\theta X_{nn}^2 - (E_\theta X_{nn})^2$, και εφόσον,

$E_\theta X_{nn}^2 = \int_0^\theta n\theta^{-n} x^{n+1} dx = n\theta^2 \int_0^1 z^{n+1} dz = \frac{n\theta^2}{n+2}$, έχουμε

$\mathcal{D}_\theta(d(\underline{X})) = \frac{n\theta^2}{n+2} - \frac{n^2\theta^2}{(n+1)^2} = \frac{n\theta^2}{(n+1)^2(n+2)} \quad \forall \theta \in \Theta.$

(γ) Άρα, από το (2.13),

$MT\Sigma_\theta(d) = \mathcal{D}_\theta(d(\underline{X})) + [b_\theta(d)]^2 = \frac{n\theta^2}{(n+1)^2(n+2)} + \frac{\theta^2}{(n+1)^2} =$

$= \frac{2\theta^2}{(n+1)(n+2)} \quad \forall \theta \in \Theta.$

(δ) $R(\theta,d) = E_\theta|d(\underline{X})-\theta| = E_\theta|X_{nn}-\theta| = \int_0^\theta |x-\theta|\, n\theta^{-n} x^{n-1} dx$

$= \int_0^\theta (\theta - x)\, n\theta^{-n} x^{n-1} dx = n\theta \int_0^1 z^{n-1}(1-z)\,dz =$

$= n\theta\, B(n,2) = n\theta \frac{\Gamma(n)\Gamma(2)}{\Gamma(n+2)} = n\theta \frac{(n-1)!\,1!}{(n+1)!} =$

$= \frac{\theta}{n+1} \quad \forall \theta \in \Theta.$

Συνήθως το Μ.Τ.Σ. είναι ευκολότερο να υπολογισθή απ' τις άλλες μορφές κινδύνου.

(2.18) <u>Άσκηση</u>: Έστω $X_1, \ldots, X_n$ α.ι. $N(\theta, \sigma_0^2)$, με σ_0 γνωστό και άγνωστη παράμετρο $\theta \in \Theta = \mathbb{R}$. Ως εκτιμήτρια της θ επιλέγουμε την $d(\underline{X}) := \overline{X}_n$, όπως στο παράδειγμα (2.6 α). Δείξτε ότι:

(α) $d \in \mathcal{U}_\Theta$ και επίσης ότι η συνάρτηση $MT\Sigma_\theta(d) = \frac{\sigma_0^2}{n}$ – σταθερά,

(β) $R(\theta,d) = \sqrt{\frac{2}{\pi}}\frac{\sigma_0}{\sqrt{n}}$ – σταθερά, αν χρησιμοποιηθή η συνάρτηση απώλειας της (2.17 δ).

(2.19) <u>Άσκηση</u>: Επαναλάβετε το Παράδειγμα (2.17) με:

$X_1, \ldots, X_n$ α.ι. $\mathcal{U}(\theta, 1)$ και εκτιμήτρια

$d(\underline{X}) := X_{n1} = \min\{X_i,\ i=1,\ldots,n\}$.

(2.20) Παραδειγμα. Ελεγχος Υποθεσεως: Εστω $\underline{X}_1, \ldots, \underline{X}_n$ α.ι. $f(\underline{x}|\underline{\theta})$, $\underline{x}\in\mathbb{R}^m$, $\underline{\theta}\in\Theta\subseteq\mathbb{R}^k$. Το προβλημα ειναι ο ελεγχος της υποθεσεως $H:\underline{\theta}\in\Theta_0$ vs $K:\underline{\theta}\in\Theta_1$ με $\Theta_0\cap\Theta_1=\emptyset$ και $\Theta_1, \Theta_2\subseteq\Theta$, δηλαδη, πρεπει να αποφασισουμε το αν δεχομαστε οτι η H ειναι ορθη οταν εχουμε να διαλεξουμε αναμεσα σ'αυτην και την K – την εναλλακτικη υποθεση. Αρα ο χωρος $\mathcal{D}$ των συναρτησεων αποφασεως αποτελειται απο ολες τις στατιστικες συναρτησεις $d:\mathcal{X}\subseteq\mathbb{R}^{mn}\longrightarrow\mathcal{A}=[0,1]$, δηλαδη, αν $d(\underline{x})=0$ τοτε δεχομαστε ανεπιφυλακτα την H, αν $d(\underline{x})=1$ τοτε απορριπτουμε ανεπιφυλακτα την H εναντι της εναλλακτικης της K, εαν δε $0<d(\underline{x})<1$ τοτε η αποφαση μας ειναι τυχαιοποιημενη, δηλαδη, μια τ.μ. η οποια διαλεγει την H με πιθανοτητα $1-d(\underline{x})$ και την K με πιθανοτητα $d(\underline{x})$, και αποφασιζουμε λοιπον με την διεξαγωγη του συμπληρωματικου πειραματος της παρατηρησης μιας Bernoulli$(p=d(\underline{x}))$. (Υπαρχουν, φυσικα, και αλλοι τροποι να γινη ο ελεγχος, αυτος ειναι ο κλασσικος.)

Ως συναρτηση απωλειας για το προβλημα χρησιμοποιειται συνηθως η εξης:

$$L(\theta, d(\underline{x})) = \begin{cases} 0 & \text{αν ελημφθη η σωστη αποφαση} \\ 1 & \text{αν ελημφθη καποια λανθασμενη αποφαση,} \end{cases}$$

με αντιστοιχη συναρτηση κινδυνου

$$R(\underline{\theta}, d) = E_{\underline{\theta}} L(\underline{\theta}, d(\underline{X})) =$$

$$= \left\{ \begin{array}{ll} E_{\underline{\theta}}\, d(\underline{X}) & \text{αν } \underline{\theta}\in\Theta_0 \\ E_{\underline{\theta}}[1-d(\underline{X})] & \text{αν } \underline{\theta}\in\Theta_1 \end{array} \right\} = \begin{cases} P_{\underline{\theta}}\,(\text{απορριψης της } H) & \text{αν } \underline{\theta}\in\Theta_0, \\ P_{\underline{\theta}}\,(\text{αποδοχης της } H) & \text{αν } \underline{\theta}\in\Theta_1. \end{cases}$$

Συνηθως, ο ελεγχος $d(\underline{X}) = d_\alpha(\underline{X})$, $\alpha\in[0,1]$, δηλαδη, εξαρταται απο την προκαθορισμενη πιθανοτητα α του

λαθος τυπου I $:= \sup\limits_{\underline{\theta}\in\Theta_0} R(\underline{\theta}, d) = \sup\limits_{\underline{\theta}\in\Theta_0} E_{\underline{\theta}} d_\alpha(\underline{X}) \leq \alpha$.

Ζητουμε δε να βρουμε εναν ελεγχο d_α, μεγεθους α, ο οποιος να ελαχιστοποιη το λαθος τυπου II $:= R(\underline{\theta}, d) = 1 - E_{\underline{\theta}} d(\underline{X})$ $\forall\underline{\theta}\in\Theta_1$, ει δυνατον.

Θα δουμε αργοτερα οτι το προβλημα της κατασκευης περιοχης επιθυμοσυνης της αγνωστης παραμετρου $\underline{\theta} \in \Theta$, ειναι, κατα μια εννοια, δυϊκο του προβληματος ελεγχου υποθεσεως σχετικης με την $\underline{\theta}$, και ως εκ τουτου δεν ειναι απαραιτητο να μελετησουμε την δομη του κινδυνου του ξεχωριστα.

Για ενα συγκεκριμενο παραδειγμα, συνεχιζουμε το (2.17),

(2.21) Παραδειγμα. Εστω $\underline{X}_1, \ldots, \underline{X}_n$ a.r. $\mathcal{U}(0,\theta)$, $\theta \in \Theta = (0,+\infty)$. Το προβλημα μας ειναι ο ελεγχος της υποθεσεως $H: \theta \leq 1$ vs $K: \theta > 1$, δηλαδη, $\Theta_0 = (0,1]$, $\Theta_1 = (1,+\infty)$.

Με μια στατιστικη μεθοδολογια, την οποια δεν θα περιγραψουμε τωρα, καταληγουμε στην ελεγχοσυναρτηση:

$d(\underline{X}) = 1(X_{nn} > c)$, οπου η κρισιμη σταθερα c οριζεται μεσω της σχεσεως:

$\sup\limits_{\theta \in \Theta_0} R(\theta, d) = \sup\limits_{\theta \in \Theta_0} E_\theta d(\underline{X}) = \alpha$. Εχουμε λοιπον,

$$R(\theta,d) = E_\theta d(\underline{X}) = E_\theta\{1(X_{nn} > c)\} = P_\theta(X_{nn} > c) =$$
$$= \int_c^\infty f(x|\theta)\,dx = \int_c^\theta n\theta^{-n} x^{n-1}\,dx =$$
$$= n\int_{c/\theta}^1 z^{n-1}\,dz = \left[1 - \left(\tfrac{c}{\theta}\right)^n\right] 1(c \leq \theta) \uparrow_\theta \Rightarrow$$

$$\Rightarrow \sup_{\theta \in \Theta_0} R(\theta,d) = \sup_{0<\theta\leq 1}\left[1 - \left(\tfrac{c}{\theta}\right)^n\right] 1(c \leq \theta) = (1-c^n)\,1(c \leq 1)$$

$= \alpha$, το λαθος τυπου I. Αρα, $c_\alpha = (1-\alpha)^{1/n}$ και ο ελεγχος μεγεθους $\alpha \in (0,1)$ που χρησιμοποιουμε ειναι ο εξης:

$$d_\alpha(\underline{X}) = 1(X_{nn} > (1-\alpha)^{1/n}).$$

Η συναρτηση κινδυνου του d_α για $\theta \in \Theta_1$ (Λαθος τυπου II) ειναι:

$$R(\theta, d_\alpha) = E_\theta[1 - d_\alpha(\underline{X})] = E_\theta\{1(X_{nn} \leq (1-\alpha)^{1/n})\} =$$
$$= P_\theta(X_{nn} \leq (1-\alpha)^{1/n}) = \int_0^{(1-\alpha)^{1/n}} n\theta^{-n} x^{n-1}\,dx =$$
$$= \int_0^{\frac{(1-\alpha)^{1/n}}{\theta}} n z^{n-1}\,dz = \frac{1-\alpha}{\theta^n}, \quad \theta \in \Theta_1 = (1,+\infty).$$

Βλεπουμε λοιπον, οτι ο κινδυνος λαθους, ειτε τυπου I ειτε τυπου II, μειουται οσο τα θ του Θ_0 και του Θ_1 απομακρυνονται μεταξύ τους, δηλαδή, η επιλογή που έχουμε να κάνουμε γίνεται πιο ξεκάθαρη.

(2.22) <u>Άσκηση</u>. Έστω $X_1, X_2, \ldots, X_n$ α.ι. $N(\theta, \sigma_0^2)$, με σ_0 γνωστό και άγνωστη παράμετρο $\theta \in \Theta = \mathbb{R}$. Θέλουμε να ελέγξουμε σε επίπεδο $\alpha = 0{,}05$ την υπόθεση $H: \theta \leq 0$ vs $K: \theta > 0$. Χρησιμοποιείστε την ελεγχοσυνάρτηση: $d_\alpha(\underline{X}) = 1(\bar{X}_n > c_\alpha)$.
(α) Βρήτε την κρίσιμη σταθερά c_α.
(β) Υπολογίστε την συνάρτηση κινδύνου $R(\theta, d_\alpha)$ για $\theta \in \Theta_1$.

(2.23) <u>Άσκηση</u>. Επαναλάβετε το Παράδειγμα (2.21) με: $X_1, \ldots, X_n$ α.ι. $\mathcal{U}(\theta, 1)$, $\theta \in \Theta = (-\infty, 1)$. Θέλουμε να ελέγξουμε σε επίπεδο $\alpha = 0{,}05$ την υπόθεση $H: \theta \in [0, 1)$ vs $K: \theta < 0$. Χρησιμοποιείστε την ελεγχοσυνάρτηση $d(\underline{X}) = 1(X_{n1} < c)$.

3. ΕΚΤΙΜΗΤΙΚΗ

Σε πρωτο σταδιο, θα δουμε τις δυνατοτητες που μας παρεχει η θεωρια αποφασεων για την κατασκευη εκτιμητριων καθως επισης και τα ορια των δυνατοτητων αυτης της μεθοδου. Οι εκτιμητριες που βασιζονται σ'αυτες τις ιδεες ειναι κυριως τριων ειδων:

(i) Εκτιμητριες minimax,

(ii) Εκτιμητριες Bayes,

(iii) Ομοιομορφα αμεροληπτες εκτιμητριες ελαχιστης διασπορας (ΟΑΕΕΔ)

Κατοπιν, θα πλησιασουμε το προβλημα της εκτιμησης απο μια αλλη σκοπια. Θα βρουμε δηλαδη μεθοδους κατασκευης εκτιμητριων $\hat{\theta}$ της αγνωστης παραμετρου θ_0 τετοιες ωστε να ελαχιστοποιουν καποιο μετρο αποστασης αναμεσα στην $F(\cdot|\hat{\theta})$ και $F(\cdot|\theta_0)$ ή να μεγιστοποιουν καποιο μετρο καλης εφαρμογης του μοντελου $F(\cdot|\theta)$ στα δεδομενα $x_1, \dots, x_n$. Οι κυριες εκπροσωποι αυτου του τροπου κατασκευης εκτιμητριων ειναι οι:

(iv) Εκτιμητριες μεγιστης πιθανοφανειας (ΕΜΠ).

(v) Εκτιμητριες ελαχιστων τετραγωνων, που εφαρμοζουμε εδω στα γραμμικα μοντελα.

Επισης θα αναφερθουμε σε ενα εκπροσωπο της μεθοδου της αντικαταστασης, συγκεκριμενα στις

(vi) Εκτιμητριες της μεθοδου των ροπων.

Οι καλες — βασει κριτηριων που θα ορισουμε — ιδιοτητες των εκτιμητριων (iv) και (v), αποδεικνυονται αφου ορισθη η μεθοδος κατασκευης τους, εν αντιθεσει προς τις (i),(ii),(iii) που κατασκευαζονται βασει προκαθορισμενων — και λιγοτερο φιλοδοξων — κριτηριων καταλληλοτητος.

Θα ασχοληθουμε εδω, λογω της σπουδαιοτητος και αντιπροσωπευτικοτητος τους, κυριως με τις εκτιμητριες (iv) και (v).

Το προβλημα μας λοιπον ειναι η εκτιμηση της αγνωστης παραμετρου $\theta \in \Theta \subseteq \mathbb{R}^k$ της κατανομης $F(\cdot|\theta) \in \mathcal{F}_{\Theta}$.

Εστω δε ανεξαρτητο στοχαστικο δειγμα απο την $F(\cdot|\theta)$, δηλαδη, εστω $\underline{X} = (X_1, \dots, X_n)$ α.ι. $F(x|\theta)$, $x \in \mathbb{R}^m$, $\theta \in \Theta \subseteq \mathbb{R}^k$.

(3.1) <u>Ορισμος</u>: Καθε στατιστικη συναρτηση $d : \mathcal{X} \subseteq \mathbb{R}^{mn} \longrightarrow \mathcal{A} \subseteq \mathbb{R}^k$ (βλ.(2.5)) η οποια χρησιμοποιειται προς εκτιμηση της αγνωστης παραμετρου $\theta \in \Theta$, καλειται <u>εκτιμητρια</u> της θ.

Εστω λοιπον, οτι εχουμε επιλεξει μια συναρτηση απωλειας
$L : \Theta \times \mathcal{A} \ni (\theta, a) \longmapsto L(\theta, a) \in [0, +\infty)$,
για το προβλημα της εκτιμησης της παραμετρου θ, και αρα για καθε εκτιμητρια $d \in \mathcal{D}$ — συνολο στατιστικων συναρτησεων απο το δειγματικο χωρο $\mathcal{X}$ στο χωρο των αποφασεων $\mathcal{A}$ — εχουμε μια συγκεκριμενη συναρτηση κινδυνου
$R(\theta, d) := E_\theta\{L(\theta, d(\underline{X}))\}$, $\theta \in \Theta$.

Φυσικα, ιδανικα θα θελαμε να βρουμε μια εκτιμητρια $d^* \in \mathcal{D}$ τετοια ωστε

(3.2) $R(\theta, d^*) \leq R(\theta, d) \quad \forall \theta \in \Theta$ και $\forall d \in \mathcal{D}$.

Δυστυχως, τετοιες εκτιμητριες δεν υπαρχουν, διοτι $\forall \theta \in \Theta$ μπορουμε να ορισουμε μια συγκεκριμενη εκτιμητρια, την σταθερα $d_\theta(\underline{X}) = \theta$ — οποιο και αν ειναι το δειγμα, η οποια σ' αυτο το συγκεκριμενο θ θα εχη κινδυνο $R(\theta, d_\theta) = E_\theta L(\theta, d_\theta) = E_\theta L(\theta, \theta) = L(\theta, \theta) = 0$ ή την ελαχιστη τιμη της L, για καθε λογικη επιλογη συναρτησης απωλειας L. Για να ισχυη λοιπον η (3.2) $\forall \theta \in \Theta$, θα πρεπει η βελτιστη εκτιμητρια d^* να εχη κινδυνο $R(\theta, d^*) = 0 \ \forall \theta \in \Theta$, δηλαδη, ή η συναρτηση απωλειας ειναι η αχρηστη $L(\theta, d^*(\underline{x})) = 0 \ \forall \underline{x} \in \mathcal{X}$ και $\forall \theta \in \Theta$ ή ολη η μαζα πιθανοτητος της κατανομης μας συγκεντρωνεται στο σημειο θ (και αρα $x_i = \theta \ \forall i = 1, 2, \dots$) και η d^* ειναι μια απο τις συγκεκριμενες εκτιμητριες d_θ που προαναφεραμε.

Αρα, σε ολες τις ενδιαφερουσες περιπτωσεις η (3.2) ειναι ανεφικτη. Θα αναφερουμε τρεις τροπους αποφυγης αυτου του αδιεξοδου, με κοινο χαρακτηριστικο το οτι αποκλειουν τις τετριμμενες εκτιμητριες που προαναφεραμε.

<u>Ομοιομορφα αμεροληπτες εκτιμητριες ελαχιστης διασπορας</u> (ΟΑΕΕΔ):

Περιοριζομαστε στη συναρτηση απωλειας $L(\theta,a)=(a-\theta)^2$ (θα δουλεψουμε εδω μονο με $k=1$, η γενικευση για $k\geq 2$ ειναι συνηθως απλη). Τοτε, απο το (2.13),

$$R(\theta,d)=E_\theta[d(\underline{X})-\theta]^2=\mathcal{D}_\theta(d(\underline{X}))+[E_\theta d(\underline{X})-\theta]^2.$$

Αποκλειουμε δε τις τετριμμενες εκτιμητριες, με το να απαιτουμε την αμεροληψια των εκτιμητριων που θεωρουμε, δηλαδη, θεωρουμε μονο $d\in\mathcal{U}_{\Theta}:=\{d\in\mathcal{D} : E_\theta d(\underline{X})=\theta \ \forall\theta\in\Theta\}$.

Τοτε, εν γενει υπαρχουν $d^*\in\mathcal{U}_{\Theta}$ τετοια ωστε

(3.3) $R(\theta,d^*)=\mathcal{D}_\theta(d^*(\underline{X}))\leq\mathcal{D}_\theta(d(\underline{X}))=R(\theta,d)\ \forall\theta\in\Theta$ και $\forall d\in\mathcal{U}_{\Theta}$,

και μαλιστα θα δωσουμε αργοτερα μια γενικη μεθοδο κατασκευης ΟΑΕΕΔ εκτιμητριων (Θεωρημα Rao-Blackwell).

Το κοινο χαρακτηριστικο των επομενων δυο μεθοδων εκτιμησης ειναι οτι περιοριζουν την συγκριση των <u>συναρτησεων κινδυνων</u> $R(\theta,d_1)$, $R(\theta,d_2)$, $\theta\in\Theta$, δυο εκτιμητριων $d_1,d_2\in\mathcal{D}$ στην συγκριση δυο <u>χαρακτηριστικων τιμων</u> αυτων των συναρτησεων.

(3.4) <u>Παραδειγμα</u>: Εστω $X_1,\dots,X_n$ α.τ. $N(\theta,1)$, $\theta\in\Theta=\mathbb{R}$. Θεωρουμε ως εκτιμητριες της αγνωστης παραμετρου θ, τις $d_1(\underline{X})=\overline{X}_n=\frac{1}{n}\sum_{i=1}^n X_i$, $d_2(\underline{X})=\frac{1}{2}\overline{X}_n$, και θελουμε να τις συγκρινουμε με βαση το μεσο τετραγωνικο σφαλμα τους. Εχουμε: $R(\theta,d_1)=\mathcal{D}_\theta(\overline{X}_n)=1/n$, και

$$R(\theta,d_2)=\mathcal{D}_\theta(\tfrac{1}{2}\overline{X}_n)+[E_\theta(\tfrac{1}{2}\overline{X}_n)-\theta]^2=(1+n\theta^2)/4n,$$

παρατηρουμε δε οτι : $R(\theta,d_1) < R(\theta,d_2)$ για $|\theta| > (3/n)^{1/2}$

και $R(\theta,d_1) > R(\theta,d_2)$ για $|\theta| < (3/n)^{1/2}$,

δηλαδη καμμια δεν ειναι ομοιομορφα καλυτερη της αλλης.

Αυτου του ειδους τα προβληματα αναμενονται οταν ειναι να συγκρινη κανεις δυο συναρτησεις (του $\theta \in \Theta$).

<u>Εκτιμητριες Minimax</u>: Σ' αυτη την περιπτωση, συγκρινουμε τις συναρτησεις κινδυνων $R_d(\theta) \equiv R(\theta,d)$, $\theta \in \Theta$, με βαση τις μεγιστες τιμες τους μονο : $\sup_{\theta \in \Theta} R(\theta,d)$, $d \in \mathcal{D}$, χρησιμοποιουμε δε ως εκτιμητρια της θ, εκεινη την $d^* \in \mathcal{D}$ – αν υπαρχη – η οποια ελαχιστοποιει αυτη τη μεγιστη τιμη κινδυνου, δηλαδη, περιοριζομαστε σε εκεινες τις εκτιμητριες $d \in \mathcal{D}'$ τετοιες ωστε υπαρχει το $\max_{\theta \in \Theta} R(\theta,d)$, η minimax δε εκτιμητρια d^*, ειναι εκεινη για την οποια

$$(3.5) \quad \max_{\theta \in \Theta} R(\theta,d^*) = \min_{d \in \mathcal{D}'} \max_{\theta \in \Theta} R(\theta,d) .$$

Με βαση αυτο το κριτηριο, π.χ., η d_1 του (3.4) ειναι σαφως προτιμητεα της d_2 η οποια ουτε καν ανηκει στο $\mathcal{D}'$.

<u>Εκτιμητριες Bayes</u>: Σ' αυτη την περιπτωση, συγκρινουμε τις συναρτησεις κινδυνων $R_d(\theta) \equiv R(\theta,d)$, $\theta \in \Theta$, με βαση τις μεσες τιμες τους ως προς καποια συναρτηση βαρους $\pi(\theta)$, $\theta \in \Theta$, τετοια ωστε $\pi(\theta) \geq 0 \ \forall \theta \in \Theta$ και $\int_\Theta \pi(\theta)\,d\theta = 1$. Δηλαδη, μεταχειριζομαστε την παραμετρο θ σαν τυχαια μεταβλητη με κατανομη $\pi(\theta)$, $\theta \in \Theta$. Περιοριζομαστε δε, σε εκεινες τις εκτιμητριες $d \in \mathcal{D}''$ για τις οποιες ο <u>κινδυνος Bayes</u> <u>ως προς π</u> :

$$(3.6) \quad r(\pi,d) := \int_\Theta R(\theta,d)\,\pi(\theta)\,d\theta < +\infty ,$$

χρησιμοποιουμε δε ως εκτιμητρια εκεινη την $d_\pi \in \mathcal{D}''$ για την οποια:

(3.7) $$r(\pi, d_\pi^*) = \min_{d \in \mathcal{D}''} r(\pi, d) .$$

Φυσικα, αν αλλαξουμε την "a priori" κατανομη π της θ, θα αλλαξη και η εκτιμητρια του Bayes d_π.

(3.8) <u>Προταση</u>. Εστω $X_1, \dots, X_n$ α.ι. $F(x|\theta)$, $x \in \mathbb{R}$, $\theta \in \Theta \subseteq \mathbb{R}$, και εστω η συναρτηση βαρους $\pi(\theta)$, $\theta \in \Theta$ ή αν θελετε $\vartheta \sim \pi(\theta)$, $\theta \in \Theta$. Αν $L(\theta, a) = (a-\theta)^2$, $(\theta, a) \in \Theta \times \mathcal{A}$, τοτε η εκτιμητρια Bayes ειναι:

$$d_\pi(\underline{X}) = E(\vartheta | \underline{X}) = \begin{cases} \int_\Theta \theta \pi(\theta|\underline{x})\, d\theta & \text{αν } \vartheta \text{ απολ. συνεχης} \\ \sum_{\theta \in \Theta} \theta \pi(\theta | \underline{X}) & \text{αν } \vartheta \text{ διακριτη} \end{cases},$$

δηλαδη, η d_π ειναι η μεση τιμη της "a posteriori" κατανομης $\pi(\theta|\underline{X})$ της ϑ δεδομενου του δειγματος $\underline{X}$, εφοσον αυτη υπαρχει.

<u>Αποδ.</u> Ο κινδυνος Bayes μπορει να εκφρασθη ως εξης:

$$\begin{aligned} r(\pi, d) &= \int_\Theta R(\theta, d) \pi(\theta) d\theta = \int_\Theta E\left[(d(\underline{X}) - \vartheta)^2 \middle| \vartheta = \theta \right] \pi(\theta) d\theta \\ &= E(\vartheta - d(\underline{x}))^2 = E\{\vartheta^2 + d(\underline{x})^2 - 2\vartheta d(\underline{X})\} \\ &= E\{(\vartheta - d_\pi(\underline{X}))^2 - d_\pi(\underline{X})^2 + d(\underline{x})^2 + 2\vartheta d_\pi(\underline{X}) - 2\vartheta d(\underline{X})\} \\ &= r(\pi, d_\pi) - E(d_\pi(\underline{x})^2) + E(d(\underline{X}))^2 + 2E[\vartheta d_\pi(\underline{x})] - 2E[\vartheta d(\underline{x})] . \end{aligned}$$

Αλλα, $E[\vartheta d_\pi(\underline{X})] = E\{E[\vartheta d_\pi(\underline{x}) | \underline{X}]\} = E\{d_\pi(\underline{X}) E[\vartheta | \underline{X}]\}$
$= E\{d_\pi(\underline{X})^2\}$, οπου χρησιμοποιησαμε την (1.70 a). Με τον ιδιο δε τροπο βλεπουμε οτι $E[\vartheta d(\underline{x})] = E\{d_\pi(\underline{X}) d(\underline{X})\}$. Αρα,

$$\begin{aligned} r(\pi, d) &= r(\pi, d_\pi) - E(\cancel{d_\pi(\underline{x})^2}) + E(d(\underline{x})^2) + \cancel{2}E(d_\pi(\underline{x})^2) - 2E(d_\pi(\underline{x}) d(\underline{x})) \\ &= r(\pi, d_\pi) + E[d_\pi(\underline{x}) - d(\underline{X})]^2 \geqslant r(\pi, d_\pi) \quad \forall d \in \mathcal{D}'' . \end{aligned}$$

(3.9) <u>Παραδειγμα</u>: Εστω $X_1, \dots, X_n$ α.ι. Poisson(θ), $\theta \in \Theta = (0, +\infty)$. Θεωρουμε $\pi(\theta) = \frac{\lambda^\alpha}{\Gamma(\alpha)} \theta^{\alpha-1} e^{-\lambda\theta} 1(\theta > 0)$, δηλαδη, $\vartheta \sim \mathcal{G}(\alpha, \lambda)$. Δειξτε οτι η εκτιμητρια Bayes της αγνωστης παραμετρου θ ειναι η:

$$d_\pi(\underline{X}) = E(\vartheta|\underline{X}) = \int_\Theta \theta\pi(\theta|\underline{X})d\theta = \frac{\alpha+\Sigma_1^n X_i}{\lambda+n} = \frac{\lambda}{\lambda+n}E\vartheta + \left(1-\frac{\lambda}{\lambda+n}\right)\bar{X}_n$$

Απαντηση. $\underline{X}\,|\,\vartheta=\theta \sim p(\underline{x}|\theta) = \prod_{i=1}^n \left\{\frac{e^{-\theta}\theta^{x_i}}{x_i!}\right\} = e^{-n\theta}\theta^{n\bar{x}} \Big/ \prod_{i=1}^n (x_i!)$.

Αρα,

$$\vartheta\,|\,\underline{X}=\underline{x} \sim \pi(\theta|\underline{x}) = \frac{p(\underline{x}|\theta)\pi(\theta)}{\int_\Theta p(\underline{x}|\theta_1)\pi(\theta_1)d\theta_1} =$$

$$= \frac{e^{-n\theta}\theta^{n\bar{x}}\,\theta^{\alpha-1}e^{-\lambda\theta}}{\int_0^\infty e^{-(n+\lambda)\theta_1}\theta_1^{n\bar{x}+\alpha-1}d\theta_1} = \frac{e^{-(\lambda+n)\theta}\,\theta^{n\bar{x}+\alpha-1}(n+\lambda)^{n\bar{x}+\alpha-1}}{\Gamma(n\bar{x}+\alpha)},\ \theta>0,$$

δηλαδη, $\vartheta|\underline{X}=\underline{x} \sim \mathcal{G}(n\bar{x}+\alpha, \lambda+n)$ και αρα:

$$E(\vartheta|\underline{X}=\underline{x}) = \frac{n\bar{x}+\alpha}{\lambda+n} \Rightarrow E(\vartheta|\underline{X}) = \frac{n\bar{X}_n+\alpha}{\lambda+n}.$$

(3.10) Ασκηση. Εστω $X_1,\ldots,X_n$ α.ι. Weibull(α,θ), $\theta\in\Theta=(0,+\infty)$ και $\alpha>0$ γνωστο. Εστω δε οτι $\pi(\theta)=\lambda e^{-\lambda\theta}1(\theta>0)$, $\lambda>0$. Δειξτε οτι η εκτιμητρια Bayes ειναι η

$$d_\pi(\underline{X}) = \frac{n+1}{\lambda+\sum_{i=1}^n X_i^\alpha}.$$

(3.11) Ασκηση. Εστω $X_1,\ldots,X_n$ α.ι. $\mathcal{G}(\alpha,\theta)$, $\theta\in\Theta=(0,+\infty)$ και $\alpha>0$ γνωστο. Εστω δε οτι $\pi(\theta)=\frac{\lambda^\beta}{\Gamma(\beta)}\theta^{\beta-1}e^{-\lambda\theta}1(\theta>0)$. Βρητε την εκτιμητρια Bayes $d_\pi(\underline{X})$ της αγνωστης παραμετρου θ.

Σημειωνουμε εδω οτι μια εκτιμητρια minimax μπορει να ληφθη και ως εκτιμητρια Bayes για καποια συγκεκριμενη a priori κατανομη π_0, η οποια και καλειται "η λιγοτερο επιθυμητη". Φυσικα, ολες οι εκτιμητριες μπορουν να θεωρηθουν Bayes για καποια συναρτηση βαρους π, αλλα ποια π ακριβως και γιατι να χρησιμοποιηθη αυτη η συγκεκριμενη π, ειναι ερωτηματα που δεν μπορει εν γενει να απαντηση αυτη η θεωρια εκτιμησεως. Σημειωνουμε επισης, οτι η στρατηγικη minimax της (3.5), για την κατασκευη εκτιμητριων, θεωρειται υπερ το δεον συντηρητικη για τα προβληματα της στατιστικης. Η αξια της εδω ειναι περισσοτερο θεωρητικη.

Δεν θα μελετήσουμε εδώ περισσότερο τις εκτιμήτριες minimax και Bayes, αλλά θα μελετήσουμε με καποια λεπτομέρεια, μια άλλη μέθοδο εκτίμησης η οποία επίσης βασίζεται σε κριτήρια της θεωρίας αποφάσεων, και συγκεκριμένα τις ομοιόμορφα αμερόληπτες εκτιμήτριες ελαχίστης διασποράς (ΟΑΕΕΔ) της (3.3).

Γιαυτή τη μελέτη είναι απαραίτητο να εισάγουμε τώρα τις έννοιες της επάρκειας και πληρότητος στατιστικών συναρτήσεων ως προς μια οικογένεια κατανομών $\mathcal{F}_{\Theta}$. Οι ιδιότητες αυτές ορισμένων στατιστικών συναρτήσεων, χαρακτηρίζουν την ικανότητα αυτών των στατιστικών συναρτήσεων, να ξεχωρίζουν από το δείγμα την απαραίτητη και μόνο "πληροφορία" σχετική με την παράμετρο θ που μας ενδιαφέρει στο εκάστοτε στατιστικό πρόβλημα. Οι έννοιες αυτές θα μας φανούν χρήσιμες και αργότερα, στη μελέτη των εκτιμητριών μεγίστης πιθανοφανείας.

(3.12) Ορισμός: Έστω $\underline{X} = (X_1, \dots, X_n)$ στοχαστικό δείγμα από την κατανομή $F(\cdot|\theta)$, $\theta \in \Theta$. Μια στατιστική συνάρτηση $T(\underline{X})$ (βλ. (2.5)) λέγεται επαρκής για την παράμετρο θ (ή την οικογένεια κατανομών $\mathcal{F}_{\Theta}$) αν και μόνο αν η δεσμευμένη κατανομή των $\underline{X} \mid T(\underline{X}) = t$ δεν εξαρτάται από την παράμετρο θ.

Δηλαδή, εφόσον ξέρουμε την τιμή της επαρκούς στατιστικής συνάρτησης $T(\underline{X})$, το δείγμα $\underline{X}$ δεν περιέχει επι πλέον πληροφορία για την παράμετρο θ. Όλη η πληροφορία του δείγματος $\underline{X}(\omega) = \underline{x}$ σχετική με την θ, έχει συναχθή από την επαρκή στατιστική συνάρτηση T και περιέχεται στην τιμή της $T(\underline{x})$.

(3.13) <u>Παράδειγμα</u>. Έστω $X_1, \ldots, X_n$ α.ι. Poisson(θ), $\theta \in \Theta = (0,+\infty)$. Θα δείξουμε ότι η στατιστική συνάρτηση $T(\underline{X}) := \sum_{i=1}^n X_i$ είναι επαρκής για την θ.

$$P(X_1=x_1, \ldots, X_n = x_n \mid T(\underline{X})=t) = \begin{cases} 0 & \text{αν } \sum_{i=1}^n x_i \neq t \\ \dfrac{P(X_1=x_1,\ldots,X_n=x_n, T(\underline{X})=t)}{P(T(\underline{X})=t)} & \text{αν } \sum_1^n x_i = t \end{cases}$$

Άρα, εφόσον $T(\underline{X}) = \sum_{i=1}^n X_i \sim$ Poisson$(n\theta)$ (βλ.(1.54)), αν $\sum_1^n x_i = t$ έχουμε: $P(\underline{X}=\underline{x} \mid T(\underline{x})=t) = P(\underline{X}=\underline{x}, \sum_1^n X_i = t)/P(\sum_1^n X_i = t)$

$$= P\left(X_1=x_1, \ldots, X_{n-1}=x_{n-1}, X_n = t - \textstyle\sum_{i=1}^{n-1} x_i\right) / P(\textstyle\sum_{i=1}^n X_i = t)$$

$$= \left(\prod_{i=1}^{n-1} P(X_i = x_i)\right) P(X_n = t - \textstyle\sum_{i=1}^{n-1} x_i) / P(\textstyle\sum_{i=1}^n X_i = t)$$

$$= \frac{e^{-(n-1)\theta}\, \theta^{\sum_{i=1}^{n-1} x_i}}{\prod_{i=1}^{n}(x_i!)} \; \frac{e^{-\theta}\, \theta^{t-\sum_{i=1}^{n-1} x_i}}{(t - \sum_{i=1}^{n-1} x_i)!} \cdot \frac{t!}{e^{-n\theta}\,(n\theta)^t} =$$

$$= \frac{t!}{x_1! \cdots x_{n-1}!\,(t - \sum_{i=1}^{n-1} x_i)!} \; \frac{1}{n^t} = \binom{t}{x_1, \ldots, x_{n-1}, t - \sum_{i=1}^{n-1} x_i} \left(\frac{1}{n}\right)^t .$$

Δηλαδή, $P(\underline{X}=\underline{x} \mid \sum_{i=1}^n X_i = t)$ ανεξάρτητη της $\theta \in \Theta$.

Η επάρκεια είναι εν γένει δύσκολο να αποδειχθεί απ' ευθείας απ' τον ορισμό (3.12), ιδίως στις περιπτώσεις συνεχών κατανομών. Ευτυχώς, έχουμε το ακόλουθο ισοδύναμο κριτήριο επάρκειας το οποίο είναι πολύ εύκολο στη χρήση του.

(3.14) <u>Πρόταση</u>. (<u>Κριτήριο Επάρκειας</u>) Έστω $\underline{X} = (X_1, \ldots, X_n)$ στοχαστικό δείγμα από την ομαλή κατανομή $F(x \mid \theta)$, $x \in \mathbb{R}^m$, $\theta \in \Theta \subseteq \mathbb{R}^k$ (βλ.(2.2)). Μια στατιστική συνάρτηση $T : \mathcal{X} \subseteq \mathbb{R}^{mn} \longrightarrow I \subseteq \mathbb{R}^k$ είναι επαρκής για την θ αν και μόνο αν υπάρχουν συναρτήσεις: $g : I \times \Theta \ni (t,\theta) \longmapsto g(t,\theta) \in \mathbb{R}$ — συνάρτηση του $\underline{x} \in \mathcal{X}$ <u>μόνο</u> μέσω του $T(\underline{x}) \in I$ — και $h : \mathcal{X} \to \mathbb{R}$ — συνάρτηση του δείγματος $\underline{x}$, <u>ανεξάρτητη</u> της θ — τέτοιες ώστε: $\forall\, \underline{x} \in \mathcal{X} \subseteq \mathbb{R}^{mn}$ και $\forall\, \theta \in \Theta \subseteq \mathbb{R}^k$

$$f(\underline{x} \mid \theta) = g(T(\underline{x}), \theta)\, h(\underline{x}) \quad \text{αν } F(\cdot \mid \theta) \text{ απολ. συνεχής,}$$

$$\text{ή} \quad p(\underline{x} \mid \theta) = g(T(\underline{x}), \theta)\, h(\underline{x}) \quad \text{αν } F(\cdot \mid \theta) \text{ διακριτή.}$$

Αποδ. στην περίπτωση διακριτής $F(\cdot|\theta)$ (η απόδειξη αυτού του αποτελέσματος στη γενικότητά του χρειάζεται σχετικά προχωρημένα εργαλεία από τη Θεωρία Μέτρου, βλ., π.χ., Lehmann: Testing Statistical Hypotheses σελ. 47-50).

Έστω $\mathcal{X} \equiv \{\underline{x}_1, \underline{x}_2, \dots\}$, $I \equiv \{t_1, t_2, \dots\}$ με $t_i = T(\underline{x}_i)$, $i=1,2,\dots$ και άρα η διακριτή τ.μ. $T(\underline{X})$ έχει σ.μ.π.

$$p(t|\theta) = \sum_{\{\underline{x}\in\mathcal{X}:\, T(\underline{x})=t\}} p(\underline{x}|\theta) \quad \forall t \in I \text{ και } \forall \theta \in \Theta.$$

Έστω, πρώτα, ότι η T είναι επαρκής για τη θ. Τότε,

$$\begin{aligned} p(\underline{x}|\theta) &= P_\theta(\underline{X}=\underline{x}) = P_\theta(\underline{X}=\underline{x}, T(\underline{X})=T(\underline{x})) = \\ &= P_\theta(\underline{X}=\underline{x} \mid T(\underline{X})=T(\underline{x}))\, P_\theta(T(\underline{X})=T(\underline{x})) \\ &= h(\underline{x})\, g(T(\underline{x}),\theta), \quad \forall \underline{x}\in\mathcal{X} \text{ και } \forall\theta\in\Theta, \end{aligned}$$

όπου $h(\underline{x}) := P_\theta(\underline{X}=\underline{x} \mid T(\underline{X})=T(\underline{x}))$ ανεξάρτητη της θ, λόγω του ότι η T είναι επαρκής για τη θ, και $g(t,\theta) := P_\theta(T(\underline{X})=t)$.

Έστω, τώρα, ότι $p(\underline{x}|\theta) = g(T(\underline{x}),\theta)\, h(\underline{x})$ $\forall \underline{x}\in\mathcal{X}$ και $\forall\theta\in\Theta$, έχουμε ότι: $P_\theta(\underline{X}=\underline{x} \mid T(\underline{X})=t) = 0$ αν $T(\underline{x}) \neq t$,

και αν $T(\underline{x})=t$ $\quad = P_\theta(\underline{X}=\underline{x}, T(\underline{X})=t) / P_\theta(T(\underline{X})=t)$

$$= \frac{P_\theta(\underline{X}=\underline{x})}{P_\theta(T(\underline{X})=t)} = \frac{p(\underline{x}|\theta)}{\sum_{\{\underline{y}\in\mathcal{X}:\, T(\underline{y})=t\}} p(\underline{y}|\theta)} = \frac{g(T(\underline{x}),\theta)\, h(\underline{x})}{\sum_{\{\underline{y}\in\mathcal{X}:\, T(\underline{y})=t\}} g(T(\underline{y}),\theta)\, h(\underline{y})} =$$

$$= \frac{g(t,\theta)\, h(\underline{x})}{g(t,\theta) \sum_{\{\underline{y}\in\mathcal{X}:\, T(\underline{y})=t\}} h(\underline{y})} = \frac{h(\underline{x})}{\sum_{\{\underline{y}\in\mathcal{X}:\, T(\underline{y})=t\}} h(\underline{y})},$$

ανεξάρτητη της θ. Άρα $\forall \underline{x}\in\mathcal{X}$, $\forall t\in I$, $\forall\theta\in\Theta$ $P_\theta(\underline{X}=\underline{x} \mid T(\underline{X})=t)$ ανεξάρτητη της $\theta\in\Theta$.

(3.15) Παράδειγμα: Έστω $X_1, \dots, X_n$ α.ι. $\mathcal{D}(m,\theta)$, $\theta\in\Theta=[0,1]$, $m\in\mathbb{N}$ γνωστό. Δείξτε ότι η $T(\underline{X}) := \sum_{i=1}^n X_i$ είναι επαρκής για την παράμετρο θ.

<u>Απαντ.</u> $p(\underline{x}|\theta) = P_\theta(\underline{X}=\underline{x}) = \prod_{i=1}^{n} p(x_i|\theta) = \prod_{i=1}^{n} \binom{m}{x_i} \theta^{x_i}(1-\theta)^{m-x_i}$

$= \left\{\theta^{\Sigma_{i=1}^n x_i} (1-\theta)^{mn-\Sigma_{i=1}^n x_i}\right\} \cdot \left\{\prod_{i=1}^{n} \binom{m}{x_i}\right\} = g(\Sigma_{i=1}^n x_i, \theta) \cdot h(\underline{x})$

$\forall \underline{x} \in \mathcal{X}$, $\forall \theta \in \Theta$, και αρα $\Sigma_{i=1}^n X_i$ επαρκης για την θ.

(3.16) <u>Παραδειγμα</u>: Εστω $X_1, \ldots, X_n$ α.ι. Weibull$(\alpha=3, \lambda)$, $\lambda \in \Theta = (0,+\infty)$. Θα δειξουμε οτι η $T(\underline{X}) := \Sigma_{i=1}^n X_i^3$ ειναι επαρκης για την λ.

Εχουμε οτι: $f(\underline{x}|\lambda) = \prod_{i=1}^{n} f(x_i|\lambda) = \prod_{i=1}^{n}\{3\lambda x_i^2 e^{-\lambda x_i^3} 1(x_i>0)\}$

$= \left[(3\lambda)^n \exp\{-\lambda \Sigma_{i=1}^n x_i^3\}\right] \cdot \left[(\prod_{i=1}^{n} x_i)^2 1(x_{n1}>0)\right] =$

$= g(\Sigma_1^n x_i, \lambda) h(\underline{x})$, $\forall \underline{x} \in \mathcal{X}$, $\forall \lambda \in \Theta$.

(3.17) <u>Παραδειγμα</u>: Εστω $X_1, \ldots, X_n$ α.ι. $N(\mu,\sigma^2)$, $\underline{\theta} = (\mu,\sigma^2) \in \Theta :=$ $:= \mathbb{R} \times (0,+\infty) \subseteq \mathbb{R}^2$. Θα δειξουμε οτι $\underline{T}(\underline{X}) \equiv (T_1(\underline{X}), T_2(\underline{X})) :=$ $:= (\Sigma_{i=1}^n X_i, \Sigma_{i=1}^n X_i^2)$ ειναι επαρκης για τη $\underline{\theta} = (\mu, \sigma^2)$.

Εχουμε οτι: $f(\underline{x}|\underline{\theta}) = \prod_{i=1}^{n} f(x_i|\underline{\theta}) = \prod_{i=1}^{n}\left\{\frac{1}{\sigma}\phi\left(\frac{x_i-\mu}{\sigma}\right)\right\} =$

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

$= \sigma^{-n}(2\pi)^{-n/2} \exp\left\{-\frac{1}{2\sigma^2}\left(\Sigma_{i=1}^n x_i^2 + n\mu^2 - 2\mu\Sigma_{i=1}^n x_i\right)\right\}$

$= \sigma^{-n}(2\pi)^{-n/2} \exp\left\{-\frac{1}{2\sigma^2}\left(T_2(\underline{x}) + n\mu^2 - 2\mu T_1(\underline{x})\right)\right\}$

$= g(\underline{T}(\underline{x}), \underline{\theta}) \cdot h(\underline{x})$, με $h(\underline{x}) \equiv 1$.

(3.18) <u>Παραδειγμα</u>. Εστω $X_1, \ldots, X_n$ α.ι. $f(x|\theta)$, $\theta \in \Theta \subseteq \mathbb{R}$, οπου η $f(x|\theta)$ ειναι η πυκνοτητα ή σ.μ.π. της <u>μονοπαραμετρικης</u> <u>εκθετικης</u> <u>οικογενειας</u> κατανομων (βλ.(1.39) με $k=1$), δηλαδη,

$f(x|\theta) = \exp\{c(\theta)T(x) + d(\theta) + S(x)\} 1(x \in A)$,

με συναρτησεις $c, d : \Theta \subseteq \mathbb{R} \longrightarrow \mathbb{R}$ και συναρτησεις $T, S : \mathbb{R}^m \longrightarrow \mathbb{R}$, $A \subseteq \mathbb{R}^m$ ο φορεας της κατανομης.

Θα δειξουμε οτι $T^{(n)}(\underline{X}) := \Sigma_{i=1}^n T(x_i) : \mathbb{R}^m \to \mathbb{R}$ ειναι επαρκης για την $\theta \in \Theta \subseteq \mathbb{R}$.

<u>Απαντ.</u> $f(\underline{x}|\theta) = \prod_{i=1}^{n} f(x_i|\theta) = \exp\{c(\theta)\Sigma_{i=1}^{n} T(x_i) + nd(\theta) +$
$+ \Sigma_{i=1}^{n} S(x_i)\} \prod_{i=1}^{n} 1(x_i \in A) =$
$= \exp\{c^{(n)}(\theta)\, T^{(n)}(\underline{x}) + d^{(n)}(\theta) + S^{(n)}(\underline{x})\}\, 1(\underline{x} \in A^{(n)})$
$= \exp\{c^{(n)}(\theta)\, T^{(n)}(\underline{x}) + d^{(n)}(\theta)\} \cdot \exp\{S^{(n)}(\underline{x})\} 1(\underline{x} \in A^{(n)})$
και αρα, η κατανομη $f(\underline{x}|\theta)$ του δειγματος $\underline{X}$ ανηκει
και αυτη σε μια μονοπαραμετρικη εκθετικη οικογενεια με
$c^{(n)}(\theta) = c(\theta)$, $d^{(n)}(\theta) = nd(\theta)$, $A^{(n)} = A^n$,
$T^{(n)}(\underline{x}) = \Sigma_{i=1}^{n} T(x_i)$, $S^{(n)}(\underline{x}) = \Sigma_{i=1}^{n} S(x_i)$,
και η $T^{(n)}(\underline{x})$ ειναι επαρκης για την θ, με
$g(T^{(n)}(\underline{x}), \theta) = \exp\{c^{(n)}(\theta)\, T^{(n)}(\underline{x}) + d^{(n)}(\theta)\}$ και
$h(\underline{x}) = \exp\{S^{(n)}(\underline{x})\}\, 1(\underline{x} \in A^{(n)})$.

(3.19) <u>Παρατηρηση.</u> Η <u>μονοπαραμετρικη</u> <u>εκθετικη</u> <u>οικογενεια</u> <u>κατανομων</u>
$$f(x|\theta) = \exp\{c(\theta)T(x) + d(\theta) + S(x)\}\, 1(x \in A), \quad \theta \in \Theta \subseteq \mathbb{R},$$
δεχεται μια πιο βολικη και <u>φυσικη</u> <u>παραμετριση</u> με $\eta = c(\theta)$:

$$(3.20)\quad f(x|\eta) = \exp\{\eta T(x) + d_0(\eta) + S(x)\}\, 1(x \in A), \quad \eta \in H \equiv c(\Theta) \subseteq \mathbb{R},$$
$$\text{με}\quad d_0(\eta) := -\log\left[\int_A \exp\{\eta T(x) + S(x)\}\, dx\right]$$
$$= d(c^{-1}(\eta)) \quad \text{αν η } c(\cdot) \text{ ειναι } 1\text{-}1.$$

Μ' αυτη τη φυσικη παραμετριση, η ροπογεννητρια της $T(X)$ ειναι:
$$M_T(s) = E\{e^{-sT}\} = \int_A \exp\{-sT(x) + \eta T(x) + d_0(\eta) + S(x)\}\, dx$$
$$= \exp\{d_0(\eta)\} \int_A \exp\{(\eta - s)T(x) + S(x)\}\, dx \Rightarrow$$

$$(3.21)\quad M_T(s) = \exp\{d_0(\eta) - d_0(\eta - s)\}, \quad \text{εφοσον } \eta, \eta - s \in H,$$

και σ' αυτη την περιπτωση:
$$(3.22)\quad E(T(X)) = -M_T'(0) = -d_0'(\eta),$$
$$(3.23)\quad \mathcal{D}(T(X)) = M_T''(0) - (M_T'(0))^2 = -d_0''(\eta).$$

(3.24) <u>Ασκηση</u>. Εστω $X_1, \ldots, X_n$ α.ι. Weibull$(\alpha=2, \lambda=\frac{1}{\theta})$, $\theta \in \Theta = (0,+\infty)$.

(α) Δειξτε οτι η $T(\underline{X}) = \Sigma_{i=1}^n X_i^2$ ειναι επαρκης για τη θ.

(β) Δειξτε οτι $ET(\underline{X}) = n\theta$, $\mathcal{D}(T(\underline{X})) = n\theta^2$.

(3.25) <u>Ασκηση</u>. Εστω $X_1, \ldots, X_n$ α.ι. $\mathcal{AD}(\kappa, \theta)$, $\theta \in \Theta = [0,1]$.

Βρητε μια επαρκη στατιστικη συναρτηση για την θ.

(3.26) <u>Ασκηση</u>. Εστω $X_1, \ldots, X_n$ α.ι. $\mathcal{G}(\alpha, \lambda)$, $\lambda \in \Theta = (0,+\infty)$, $\alpha > 0$ γνωστο. Δειξτε οτι η $T(\underline{X}) := \Sigma_{i=1}^n X_i$ ειναι επαρκης για την λ.

(3.27) <u>Παραδειγμα</u>: Εστω $X_1, \ldots, X_n$ α.ι. $\mathcal{U}(0,\theta)$, $\theta \in \Theta = (0,+\infty)$

Θα δειξουμε οτι η $T(\underline{X}) := X_{nn}$ ειναι επαρκης για την θ.

$$f(\underline{x}|\theta) = \prod_{i=1}^n \left\{\frac{1}{\theta} 1(0 < x_i < \theta)\right\} = \theta^{-n} 1(x_{nn} < \theta) \cdot 1(x_{n1} > 0)$$
$$= g(T(\underline{x}), \theta) . h(\underline{x}) .$$

Παρατηρουμε οτι στα παραδειγματα (2.17), (2.21) η εκτιμητρια και ο ελεγχος που χρησιμοποιησαμε για την θ της $\mathcal{U}(0,\theta)$, βασιζονονταν στην επαρκη στατιστικη X_{nn} του προβληματος και μονο σ' αυτην.

(3.27') <u>Ασκηση</u>: Εστω $X_1, \ldots, X_n$ α.ι. $\mathcal{U}(\theta, 1)$, $\theta \in \Theta = (-\infty, 1)$.

Βρητε μια επαρκη στατιστικη για την θ.

(3.28) <u>Ασκηση</u>: Εστω $X_1, \ldots, X_n$ α.ι. $f(x|\theta) = e^{-(x-\theta)} 1(x > \theta)$, $\theta \in \Theta = \mathbb{R}$. Βρητε μια επαρκη στατιστικη για την θ.

(3.29) <u>Ασκηση</u>: Εστω $X_1, \ldots, X_n$ α.ι. $f(x|\underline{\theta}) = \lambda e^{-\lambda(x-\mu)} 1(x > \mu)$, $\underline{\theta} = (\lambda, \mu) \in \Theta = (0,+\infty) \times \mathbb{R} \subseteq \mathbb{R}^2$. Βρητε μια επαρκη στατιστικη συναρτηση $\underline{T} = (T_1, T_2)$ για την $\underline{\theta} = (\lambda, \mu)$.

Απο τον ορισμο (3.12) της επαρκειας, ειναι σαφες οτι ολη η "πληροφορια" που περιεχεται στο δειγμα και αφορα στην παραμετρο θ, περιεχεται επισης και σε καθε επαρκη για την θ στατιστικη συναρτηση του δειγματος, και μαλιστα σε συμπυκνωμενη μορφη. Για παραδειγμα, απο το (3.15) βλεπουμε οτι αν ο σκοπος μας ειναι η εκτιμηση του θ, στην περιπτωση της $\mathcal{D}(m,\theta)$, αρκει να σημειωσουμε την τιμη της επαρκους για το θ στατιστικης $T(\underline{X}(\omega)) = \sum_{i=1}^{n} X_i(\omega) = \sum_{i=1}^{n} x_i$, μπορουμε δε διαγραψουμε το ακριβες δειγμα $\underline{X}(\omega) = \underline{x}$. Παρομοια στην περιπτωση του (3.17), αν μας ενδιαφερει, π.χ., η εκτιμηση των $\underline{\theta} = (\mu, \sigma^2)$, αρκει να σημειωσουμε τους αριθμους $(\sum_1^n x_i, \sum_1^n x_i^2)$, μπορουμε δε να διαγραψουμε το ακριβες δειγμα $\underline{x}$ που μας τους εδωσε. Αυτη η συχνα τρομακτικη μειωση της διαστασης των δεδομενων απο n - το μεγεθος του δειγματος - σε $1, 2$ ή $k << n$, χωρις καμμια απωλεια στην πληροφορια σχετικη με την θ, συνιστα τη λειτουργικοτητα αυτων των στατιστικων συναρτησεων στην στατιστικη συμπερασματολογια.

Οι επαρκεις στατιστικες συναρτησεις δεν ειναι μοναδικες στην περιπτωση του καθε μοντελου.

(3.30) Ορισμος. Δυο στατιστικες συναρτησεις T_1, T_2 καλουνται ισοδυναμες αν και μονο αν $\forall\, \underline{x}, \underline{y} \in \mathcal{X}$

$$T_1(\underline{x}) = T_1(\underline{y}) \Leftrightarrow T_2(\underline{x}) = T_2(\underline{y}),$$

δηλαδη, ανν συντελουν στην αυτη ελαττωση των δεδομενων.

Ειναι δε προφανες οτι:

(3.31) Αν οι T_1, T_2 ειναι ισοδυναμες και η T_1 ειναι επαρκης για τη θ τοτε και η T_2 ειναι επαρκης για τη θ.

Για παραδειγμα, οι στατιστικες συναρτησεις: $(\Sigma_1^n x_i, \Sigma_1^n x_i^2)$ και $(\bar{X}_n, S_n^2)$ — βλ. (1.62) — ειναι ισοδυναμες (αν ξερουμε την τιμη της μιας μπορουμε να υπολογισουμε την αλλη), και αρα στην περιπτωση της $N(\mu, \sigma^2)$ — Παραδ. (3.17) — η $(\bar{X}_n, S_n^2)$ ειναι επισης επαρκης για την $\underline{\theta} = (\mu, \sigma^2)$.

Η επαρκεια δεν ειναι μονο βολικη ιδιοτητα, αλλα και απαραιτητη για την αποσυμπορηση της ασχετης προς την θ πληροφοριας η οποια περιεχεται στο δειγμα και ειναι δυνατον να αμβλυνη την οξυτητα της στατιστικης συμπερασματολογιας περι την παραμετρο θ. Οι επαρκεις στατιστικες που επιτυγχανουν την πληρεστερη αποσυμφορηση ασχετης προς την θ πληροφοριας, λεγονται επαρκεις και αναγκαιες στατιστικες συναρτησεις ή, επι το τεχνικοτερο:

(3.32) Ορισμος Μια επαρκης στατιστικη M λεγεται επαρκης & αναγκαια αν και μονο αν για καθε αλλη επαρκη στατιστικη S, υπαρχει συναρτηση h τετοια ωστε $M = h(S)$.

Για παραδειγμα, στην περιπτωση της $N(\mu, \sigma^2)$ — βλ. (3.17) — το ιδιο το δειγμα $\underline{X}(\omega) = \underline{x}$ ειναι επαρκης για την $\underline{\theta} = (\mu, \sigma^2)$, εφοσον $P(\underline{X} = \underline{y} \mid \underline{X} = \underline{x}) = 1(\underline{x} = \underline{y})$ ανεξαρτητη της θ, καθως επισης και η $(X_{n1}, \ldots, X_{ni}, \ldots, X_{nn})$ ειναι επαρκης για την θ, δεδομενου οτι η πιθανοτητα

$$P(\underline{X} = \underline{x} \mid X_{n1} = x_1^o, \ldots, X_{nn} = x_n^o) = \frac{1}{n!} \prod_{i=1}^{n} 1(x_{ni} = x_i^o),$$

ανεξαρτητη της θ. (Αυτες οι δυο ειναι επαρκεις για καθε παραμετρικο μοντελο) Ομως η $(\Sigma_1^n x_i, \Sigma_1^n x_i^2)$ ή η $(\bar{X}_n, S_n^2)$ ειναι επαρκεις και αναγκαιες στατιστικες για την $\underline{\theta} = (\mu, \sigma^2)$, στην περιπτωση της $N(\mu, \sigma^2)$.

Δεν θα ασχοληθουμε εδω περισσοτερο με την εννοια της επαρκους και αναγκαιας στατιστικης αν και ολες οι επαρκεις στατιστικες που θα χρησιμοποιησουμε σε οτι ακολουθει θα ειναι επισης και αναγκαιες.

Μια άλλη εξαιρετικά χρήσιμη έννοια στη στατιστική, η οποία συνήθως χρησιμοποιείται παράλληλα με την επάρκεια, είναι η έννοια της πληρότητος στατιστικών συναρτήσεων.

(3.32) Ορισμός. Μια στατιστική συνάρτηση $T(\underline{X})$ λέγεται πλήρης για την οικογένεια κατανομών $\mathcal{F}_\Theta$ ή αλλιώς για την παράμετρο θ, αν και μόνο αν:

$$E_\theta\{g(T)\} = 0 \;\; \forall \theta \in \Theta \iff g(t) = 0 \;\; \forall t \in I.$$

Παρατηρούμε ότι αν μια οικογένεια κατανομών $\mathcal{F}$ είναι μικρή, ίσως να είναι εφικτό να ευρεθή συνάρτηση $g \not\equiv 0$ τέτοια ώστε $E_F\{g(T)\} = 0 \;\; \forall F \in \mathcal{F}$. Όσο όμως μεγαλώνουμε την $\mathcal{F}$ τόσο αυτό γίνεται και πιο δύσκολο. Όταν η οικογένεια $\mathcal{F}$ έχει μεγαλώσει τόσο ώστε αυτό να είναι αδύνατο, τότε η $\mathcal{F}$ λέγεται πλήρης.

(3.33) Παράδειγμα: Έστω $X_1, \ldots, X_n$ α.ι. $\mathcal{U}(0,\theta)$, $\theta \in \Theta = (0,+\infty)$. Θα δείξουμε ότι η επαρκής για την θ στατιστική $T(\underline{X}) := X_{nn}$ — βλ.(3.27) — είναι και πλήρης:

Από την (1.57) ή την (2.17) έχουμε ότι:

$T(\underline{X}) = X_{nn} \sim n\theta^{-n} t^{n-1} 1(0<t<\theta)$, και άρα,

$E_\theta g(T) = 0 \;\; \forall \theta \in \Theta \iff \int_0^\theta g(t)\, n\theta^{-n} t^{n-1} dt = 0 \;\; \forall \theta > 0$

$\iff \psi(\theta) \equiv \int_0^\theta g(t) t^{n-1} dt = 0 \;\; \forall \theta > 0 \implies \psi'(\theta) = g(\theta)\theta^{n-1} = 0 \;\; \forall \theta > 0$

$\implies g(\theta) = 0 \;\; \forall \theta > 0 \implies g \equiv 0.$

(3.34) Παράδειγμα. Έστω $X_1, \ldots, X_n$ α.ι. $\mathcal{D}(m,\theta)$, $\theta \in \Theta = [0,1]$. Θα δείξουμε ότι η επαρκής για τη θ στατιστική $T(\underline{X}) := \Sigma_{i=1}^n X_i$ — βλ.(3.15) — είναι και πλήρης:

Από την (1.54a), $T \sim \mathcal{D}(mn,\theta)$ και άρα

$E_\theta g(T) = 0 \;\; \forall \theta \in \Theta \iff \sum_{k=0}^{mn} g(k)\binom{mn}{k}\theta^k(1-\theta)^{mn-k} = 0 \;\; \forall \theta \in \Theta$

$\iff \sum_{k=0}^{mn} g(k)\binom{mn}{k} x^k = 0 \;\; \forall x > 0$, όπου $x := \frac{\theta}{1-\theta} \in (0,+\infty)$.

Αρα για να ειναι το πολυωνυμο ταυτοτικα μηδεν, θρεπει οι συντελεστες του να ειναι μηδεν, δηλαδη,

$g(k)\binom{mn}{k}=0 \quad \forall\, k=0,1,\dots,mn \Rightarrow$

$\Rightarrow g(k)=0 \quad \forall\, k=0,1,\dots,mn$ για αυθαιρετο n, αρα $g\equiv 0$.

(3.35) <u>Παραδειγμα</u>. Εστω $X_1,\dots,X_n$ α.ι. Poisson(θ), $\theta\in\Theta=(0,+\infty)$. Θα δειξουμε οτι η επαρκης για την θ στατιστικη συναρτηση $T(\underline{X}):=\Sigma_{i=1}^{n}X_i$ — βλ.(3.13) — ειναι και πληρης:

Απο την (1.54γ), εχουμε οτι $T\sim$ Poisson$(n\theta)$ και αρα:

$E_\theta g(T)=0 \quad \forall\theta\in\Theta \Leftrightarrow \Sigma_{k=0}^{\infty} g(k)\,\dfrac{e^{-n\theta}(n\theta)^k}{k!}=0 \quad \forall\theta>0$

$\Leftrightarrow \Sigma_{k=0}^{\infty} g(k)\dfrac{\theta^k}{k!}=0 \quad \forall\theta>0$. Αρα, εφοσον η δυναμοσειρα αυτη ειναι ταυτοτικα μηδεν σε μια περιοχη του μηδενος, επεται οτι ολοι οι συντελεστες της δυναμοσειρας ειναι μηδεν, δηλαδη, $g(k)=0 \quad \forall\, k\in\mathbb{N}_0 \Rightarrow g\equiv 0$.

(3.36) <u>Παραδειγμα</u>. Εστω $X_1,\dots,X_n$ α.ι. $\mathcal{G}(\alpha,\theta)$, $\theta\in\Theta=(0,+\infty)$, $\alpha>0$ γνωστο. Θα δειξουμε οτι η επαρκης για την θ στατιστικη $T(\underline{X})=\Sigma_{i=1}^{n}X_i$ — βλ.(3.26) — ειναι και πληρης:

Απο την (1.54ε) εχουμε οτι $T\sim\mathcal{G}(n\alpha,\theta)$ και αρα

$E_\theta g(T)=0 \quad \forall\theta\in\Theta \Leftrightarrow \int_0^\infty g(t)\dfrac{\theta^{n\alpha}}{\Gamma(n\alpha)}t^{n\alpha-1}e^{-\theta t}\,dt=0$

$\forall\theta>0 \Leftrightarrow \int_0^\infty e^{-\theta t}g(t)t^{n\alpha-1}dt=0 \quad \forall\theta>0 \Rightarrow$

$\Rightarrow g(t)t^{n\alpha-1}=0 \quad \forall t>0 \Rightarrow g\equiv 0$, βασει της Προτασεως (3.37) που παραδεχουμε τωρα χωρις αποδειξη.

(3.37) <u>Προταση</u>: Εστω συναρτηση $f:\mathbb{R}\to\mathbb{R}$ της οποιας υπαρχει ο μετασχηματισμος Laplace:

$\hat{f}(s):=\int_{\mathbb{R}} e^{-sx}f(x)\,dx$, $s\in I\subseteq\mathbb{R}$, καποιο διαστημα.

Τοτε, $\hat{f}(s)=0 \quad \forall s\in I \Leftrightarrow f(x)=0 \quad \forall x\in\mathbb{R}$.

Δεδομενου οτι η ταυτοτικα μηδεν συναρτηση εχει μετασχημασχημα Laplace ταυτοτικα μηδεν, η αποδειξη ειναι αμεση εφοσον δειχθη το μονοσημαντο του μετασχηματισμου Laplace, το οποιο ομως δεν θα δειξουμε εδω.

Η προταση ισχυει και για πολυδιαστατες συναρτησεις.

(3.38) <u>Παραδειγμα</u>: Εστω $X_1, \dots, X_n$ α.ι. $N(\mu, \sigma^2)$, $\underline{\theta} = (\mu, \sigma^2) \in \mathbb{R} \times (0, +\infty)$. Θα δειξουμε οτι η επαρκης για την $\underline{\theta}$ στατιστικη $\underline{T} = (\bar{X}_n, S_n^2)$ — βλ. (3.17) και (3.31) — ειναι και πληρης:

Απο το (1.62) εχουμε οτι $\bar{X}_n \sim N(\mu, \sigma^2/n)$ και $S_n^2 \sim \mathcal{G}\left(\frac{n-1}{2}, \frac{n-1}{2\sigma^2}\right)$ και οι $\bar{X}_n$, S_n^2 ειναι ανεξαρτητες.

Αρα,

$$E_{\underline{\theta}}\, g(T_1, T_2) = \int_{-\infty}^{+\infty}\left\{\int_0^{\infty} g(t_1, t_2) \sqrt{\frac{n}{2\pi\sigma^2}}\, e^{-\frac{n}{2}\frac{(t_1-\mu)^2}{\sigma^2}} \left(\frac{n-1}{2\sigma^2}\right)^{\frac{n-1}{2}} \frac{1}{\Gamma\left(\frac{n-1}{2}\right)} \cdot\right.$$

$$\left.\cdot\, t_2^{\frac{n-1}{2}-1} e^{-\frac{n-1}{2\sigma^2} t_2}\, dt_2\right\} dt_1 = 0 \quad \forall\, \underline{\theta} \in \Theta \iff \forall\, \mu \in \mathbb{R},\ \sigma^2 > 0$$

$$\int_{-\infty}^{+\infty}\left\{\int_0^{\infty} g(t_1, t_2)\, t_2^{\frac{n-3}{2}} \exp\left\{-\frac{n}{2\sigma^2}(t_1-\mu)^2 - \frac{n-1}{2\sigma^2} t_2\right\} dt_2\right\} dt_1 = 0 \iff$$

$$\iff \int_{-\infty}^{+\infty}\int_0^{\infty} \exp\left\{+\frac{n\mu}{\sigma^2} t_1 - \frac{n-1}{2\sigma^2} t_2\right\} g(t_1, t_2)\, t_2^{\frac{n-3}{2}} e^{-\frac{n}{2\sigma^2} t_1^2}\, dt_1 dt_2 = 0$$

$\forall\, \mu \in \mathbb{R}, \sigma^2 > 0$. Αρα, απο τη διδιαστατη μορφη της (3.37), $g(t_1, t_2)\, t_2^{(n-3)/2} \exp\{-\frac{n}{2\sigma^2} t_1^2\}$ $\forall\, t_1 \in \mathbb{R}, t_2 > 0 \Rightarrow g \equiv 0$.

(3.39) <u>Σημ.</u>: Στην περιπτωση της εκθετικης οικογενειας κατανομων — βλ. (1.39) και (3.18) — η επαρκης στατιστικη $\underline{T}(\underline{X}) = (T_1(\underline{X}), \dots, T_k(\underline{X}))$ για την θ ειναι και πληρης εφοσον η εικονα $c(\Theta)$ του Θ μεσω της $c(\theta) = (c_1(\theta), \dots, c_k(\theta))$ περιεχει ενα ανοικτο παραλληλογραμμο διαστασεως k.

(3.40) <u>Ασκηση</u>. Εστω $X_1, \dots, X_n$ α.ι. με κοινη κατανομη μια απο τις παρακατω. Βρητε μια επαρκη και πληρη στατιστικη για τη θ αν:
(α) $X_i \sim \mathcal{AD}(k, \theta)$, (β) $X_i \sim \mathcal{U}(\theta, 1)$, $\theta < 1$,
(γ) $X_i \sim N(\theta, 1)$, (δ) $X_i \sim N(0, \theta)$, (ε) $X_i \sim e^{-(x-\theta)} 1(x > \theta)$, $\theta \in \mathbb{R}$.

3.1. ΟΜΟΙΟΜΟΡΦΑ ΑΜΕΡΟΛΗΠΤΕΣ ΕΚΤΙΜΗΤΡΙΕΣ ΕΛΑΧΙΣΤΗΣ ΔΙΑΣΠΟΡΑΣ (ΟΑΕΕΔ)

Θα μελετησουμε τωρα, καθως ευτενεστερα, τον κυριο για μας αντιπροσωπο εκτιμητριων που βασιζονται σε μεθοδους της θεωριας αποφασεων.

Εστω λοιπον $X_1, \ldots, X_n$ α.ι. $f(\cdot|\theta)$ ή $p(\cdot|\theta)$, $\theta \in \Theta$. Μας ενδιαφερει να κατασκευασουμε εκτιμητριες $d(\underline{X})$ για την παραμετρικη συναρτηση $q(\theta)$, τετοιες ωστε

$E_\theta d(\underline{X}) = q(\theta) \ \forall \theta \in \Theta$, δηλαδη, η $d(\underline{X})$ να ειναι αμεροληπτη για την $q(\theta)$ — $d \in \mathcal{U}^q_\Theta$, και

$$R(\theta, d) = \mathcal{D}_\theta(d(\underline{X})) = \min_{d' \in \mathcal{U}^q_\Theta} \mathcal{D}_\theta(d'(\underline{X})) \quad \forall \theta \in \Theta.$$

Τις εκτιμητριες αυτες καλουμε ομοιομορφα αμεροληπτες εκτιμητριες ελαχιστης διασπορας (ΟΑΕΕΔ), βλ. (3.3), και τις κατασκευαζουμε με βαση τα δυο θεωρηματα που ακολουθουν.

(3.41) Θεωρημα (Rao–Blackwell) Εστω εκτιμητρια $d(\underline{X})$ της $q(\theta)$, με $E_\theta|d(\underline{X})| < +\infty \ \forall \theta \in \Theta$. Εστω επισης επαρκης για την θ στατιστικη συναρτηση $T(\underline{X})$. Οριζουμε τη στατιστικη συναρτηση — ανεξαρτητη της θ λογω της επαρκειας της $T(\underline{X})$ (βλ. 3.12) —

(3.42) $\quad d^*(\underline{X}) := E_\theta[d(\underline{X}) \,|\, T(\underline{X})] \equiv \mu(T(\underline{X}))$,

οπου $\mu(t) := E_\theta[d(\underline{X}) \,|\, T(\underline{X}) = t]$. Τοτε,

(3.43) $\quad E_\theta[d^*(\underline{X}) - q(\theta)]^2 \leq E_\theta[d(\underline{X}) - q(\theta)]^2 \quad \forall \theta \in \Theta$.

Επι πλεον, αν $\mathcal{D}_\theta(d(\underline{X})) < +\infty \ \forall \theta \in \Theta$, τοτε η (3.43) ισχυει με ισοτητα αν και μονο αν $P_\theta(d^*(\underline{X}) = d(\underline{X})) = 1 \ \forall \theta \in \Theta$.

Αποδ. Απο την (1.70 α) $E_\theta d^*(\underline{X}) = E_\theta d(\underline{X}) \quad \forall \theta \in \Theta$, και αρα απο την (2.13), η (3.43) ειναι ισοδυναμη της

(3.44) $\quad \mathcal{D}_\theta(d^*(\underline{X})) \leq \mathcal{D}_\theta(d(\underline{X})) \quad \forall \theta \in \Theta$.

Θα δειξουμε λοιπον την (3.44) χρησιμοποιωντας την (1.70β):

$$\mathcal{D}_\theta(d^*(\underline{X})) = \mathcal{D}_\theta\{E_\theta[d(\underline{X}) \,|\, T(\underline{X})]\} =$$
$$= \mathcal{D}_\theta(d(\underline{X})) - E_\theta\{\mathcal{D}_\theta[d(\underline{X}) \,|\, T(\underline{X})]\} \quad \forall \theta \in \Theta,$$

και αρα: $\forall\, \theta \in \Theta$,

$$\mathcal{D}_\theta(d^*(\underline{X})) - \mathcal{D}_\theta(d(\underline{X})) = -E_\theta\{\mathcal{D}_\theta[d(\underline{X}) \mid T(\underline{X})]\} =$$
$$= -E_\theta\{E_\theta\{[d(\underline{X}) - E_\theta[d(\underline{X}) \mid T(\underline{X})]]^2 \mid T(\underline{X})\}\} =$$
$$= -E_\theta[d(\underline{X}) - E_\theta[d(\underline{X}) \mid T(\underline{X})]]^2 = -E_\theta[d(\underline{X}) - d^*(\underline{X})]^2 \le 0,$$

με ισοτητα αν και μονο αν $E_\theta[d(\underline{X}) - d^*(\underline{X})]^2 = 0 \ \forall \theta \in \Theta$
$\iff P_\theta(d(\underline{X}) = d^*(\underline{X})) = 1 \ \forall \theta \in \Theta$ – βλ. (1.27γ). □

(3.45) <u>Θεωρημα</u>. (Lehmann–Scheffé). Εστω $d(\underline{X})$ αμεροληπτη εκτιμητρια της $q(\theta)$, δηλαδη, $E_\theta d(\underline{X}) = q(\theta) \ \forall \theta \in \Theta$. Εστω επισης επαρκης <u>και</u> πληρης για την θ στατιστικη $T(\underline{X})$. Τοτε, η $d^*(\underline{X})$ της (3.42), ειναι μια ΟΑΕΕΔ για την $q(\theta)$. Επι πλεον, αν $\mathcal{D}_\theta(d^*(\underline{X})) < +\infty \ \forall \theta \in \Theta$, τοτε η $d^*(\underline{X})$ ειναι η <u>μοναδικη</u> ΟΑΕΕΔ για την $q(\theta)$.

<u>Αποδ.</u> Απο την (1.70a) εχουμε οτι $E_\theta d^*(\underline{X}) = E_\theta\{E_\theta[d(\underline{X}) \mid T(\underline{X})]\} = E_\theta d(\underline{X}) = q(\theta) \ \forall \theta \in \Theta$ και αρα η $d^*(\underline{X})$ ειναι επισης αμεροληπτη εκτιμητρια της $q(\theta)$. Απο δε το Θεωρημα (3.41) εχουμε οτι $\mathcal{D}_\theta(d^*(\underline{X})) \le \mathcal{D}_\theta(d(\underline{X})) \ \forall \theta \in \Theta$. Αν λοιπον δειξουμε οτι <u>η d^* δεν εξαρταται απο την d</u> – εφοσον η d ειναι αμεροληπτη – τοτε η αμεροληπτη d^* θα εχει διασπορα μικροτερη απο οποιαδηποτε αλλη αμεροληπτη εκτιμητρια της $q(\theta)$ και αυτο για καθε $\theta \in \Theta$, δηλαδη, η d^* θα ειναι μια ΟΑΕΕΔ. Το οτι η d^* δεν εξαρταται απο την συγκεκριμενη αμεροληπτη d, μεσω της οποιας κατασκευαζεται – βλ. (3.42), επεται απο την πληροτητα της T, ως εξης:

Εστω λοιπον δυο οποιεσδηποτε αμεροληπτες εκτιμητριες $d_1(\underline{X})$, $d_2(\underline{X})$ της $q(\theta)$ και οι αντιστοιχες τους – μεσω της (3.42) – $d_1^*(\underline{X}) = \mu_1(T(\underline{X}))$ και $d_2^*(\underline{X}) = \mu_2(T(\underline{X}))$. Θετουμε $g := \mu_1 - \mu_2$ και εχουμε οτι $\forall \theta \in \Theta$,

$$E_\theta g(T) = E_\theta\{d_1^*(X) - d_2^*(X)\} = E_\theta d_1^*(\underline{X}) - E_\theta d_2^*(\underline{X}) =$$
$$= q(\theta) - q(\theta) = 0,$$

και αρα απο την πληροτητα της $T(\underline{X})$, $g \equiv 0$ και αρα $\mu_1 \equiv \mu_2$ και αρα $d_1^* = d_2^*$. Αρα η d^* ειναι μια ΟΑΕΕΔ της $q(\theta)$.

Εστω τωρα οτι $\mathcal{D}_\theta(d^*(\underline{X})) < +\infty \ \forall \theta \in \Theta$, εστω δε οτι υπαρχει και αλλη ΟΑΕΕΔ της $q(\theta)$, η $d'(\underline{X})$ διαφορη της $d^*(\underline{X})$. Τοτε αναγκαστικα $\mathcal{D}_\theta(d'(\underline{X})) < +\infty \ \forall \theta \in \Theta$, επισης, και αρα απο το Θεωρημα (3.41), εκτος αν $d^* = d'$,

$$E_\theta\left[E_\theta(d'(\underline{X})\,|\,T(\underline{X})) - q(\theta)\right]^2 < E_\theta\left[d'(\underline{X}) - q(\theta)\right]^2 \quad \forall \theta \in \Theta .$$

Αλλα, απ' οτι αποδειξαμε παραπανω $E_\theta[d'(\underline{X})\,|\,T(\underline{X})] = d^*(\underline{X})$, και αρα,

$$\mathcal{D}_\theta(d^*(\underline{X})) = E_\theta[d^*(\underline{X}) - q(\theta)]^2 < E_\theta[d'(\underline{X}) - q(\theta)]^2 = \mathcal{D}_\theta(d'(\underline{X})) \quad \forall \theta \in \Theta,$$

εκτος αν $d^* = d'$. Αρα η d^* ειναι η μοναδικη ΟΑΕΕΔ της $q(\theta)$. Π

(3.46) <u>Σχμ.</u> Τα Θεωρηματα (3.41) και (3.45) χρησιμοποιουνται, στην πραξη, ως εξης:

Εστω $X_1, \ldots, X_n$ α.ι. $F(\cdot\,|\,\theta)$, $\theta \in \Theta$ και εστω οτι υπαρχει επαρκης και πληρης στατιστικη συναρτηση $T(\underline{X})$ για την θ. Εστω δε οτι υπαρχει αμεροληπτη εκτιμητρια $d(\underline{X})$ της $q(\theta)$. Μας ενδιαφερει να κατασκευασουμε μια ΟΑΕΕΔ για την $q(\theta)$.

Η κατασκευη μιας ΟΑΕΕΔ $d^*(\underline{X})$ για την $q(\theta)$, γινεται μεσω της (3.42) ως εξης:

κατασκευαζουμε την συναρτηση $\mu(t) := E_\theta(d(\underline{X})\,|\,T(\underline{X}) = t)$, η οποια ειναι ανεξαρτητη της θ διοτι η T ειναι επαρκης για την θ – βλ. (3.12) – και επισης, βασει του (3.45), δεν εξαρταται απο την συγκεκριμενη αμεροληπτη $d(\underline{X})$, εφοσον η στατιστικη T και πληρης. Τοτε, η εκτιμητρια

(3.47) $d^*(\underline{X}) := \mu(T(\underline{X})) \ \left(\equiv E_\theta(d(\underline{X})\,|\,T(\underline{X}))\right)$

ειναι ΟΑΕΕΔ. Αν επισης η $\mathcal{D}_\theta(d^*(\underline{X})) < +\infty \ \forall \theta \in \Theta$, τοτε η d^* ειναι η μοναδικη ΟΑΕΕΔ της $q(\theta)$.

Παρατηρουμε επισης, οτι αν τυχη να ξερουμε $\bar{\eta}$ να μπορουμε να κατασκευασουμε εκ των προτερων αμεροληπτη εκτιμητρια $d(\underline{X})$ της $q(\theta)$, τετοια ωστε $d(\underline{X}) = h(T(\underline{X}))$, τοτε $d^*(\underline{X}) := E_\theta[d(\underline{X}) \mid T(\underline{X})] = E_\theta[h(T(\underline{X})) \mid T(\underline{X})] =$

$$= h(T(\underline{X})) = d(\underline{X}),$$

και αρα η $d(\underline{X})$ ειναι ηδη ΟΑΕΕΔ και ειναι η μοναδικη αν επι πλεον $\mathcal{D}_\theta(d(\underline{X})) < +\infty \ \forall \theta \in \Theta$.

(3.47) <u>Παραδειγμα</u>: Εστω $X_1, \ldots, X_n$ α.ι. $\mathcal{P}oisson(\theta)$, $\theta \in \Theta = (0, +\infty)$. Απο τις (3.13) και (3.35) εχουμε οτι $\exists\, T(\underline{X}) := \sum_{i=1}^n X_i$ επαρκης και πληρης για την θ. Βρητε ΟΑΕΕΔ για την:

(α) $q(\theta) = \theta$,

(β) $q(\theta) = e^{-\theta}$,

(γ) $q(\theta) = \theta^\nu$, $\nu \in \mathbb{N}$.

<u>Αποδ.</u> (α) Εφοσον $EX_1 = \theta = q(\theta)$, η $d(\underline{X}) := \overline{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$ ειναι μια αμεροληπτη εκτιμητρια της $q(\theta) = \theta$. Αλλα,

$d(\underline{X}) = \overline{X}_n = \frac{1}{n} T(\underline{X})$ και αρα απο την (3.46) $d^* = d$,

και αρα η <u>$\overline{X}_n$ ειναι ΟΑΕΕΔ</u>. Επισης, εφοσον $\mathcal{D}_\theta(\overline{X}_n) = \frac{1}{n}\mathcal{D}_\theta(X_1) = \frac{\theta}{n} < +\infty \ \forall \theta \in \Theta$, η εκτιμητρια $\overline{X}_n$ ειναι η μοναδικη ΟΑΕΕΔ για την $q(\theta) = \theta$.

(β) Εχουμε οτι $q(\theta) = e^{-\theta} = P_\theta(X_1 = 0) = E_\theta\{1(X_1 = 0)\} \ \forall \theta \in \Theta$, και αρα η $d(\underline{X}) = 1(X_1 = 0)$ ειναι αμεροληπτη για την $q(\theta)$. Αρα μια ΟΑΕΕΔ της $q(\theta) = e^{-\theta}$ ειναι η

$d^*(\underline{X}) := \mu(T(\underline{X}))$,

οπου $\mu(t) := E_\theta[d(\underline{X}) \mid T(\underline{X}) = t] = E_\theta[1(X_1 = 0) \mid T(\underline{X}) = t] =$

$= P_\theta(X_1 = 0 \mid \Sigma_1^n X_i = t) = P_\theta(X_1 = 0, \Sigma_1^n X_i = t) / P(\Sigma_1^n X_i = t)$

$= P(X_1 = 0, \Sigma_{i=2}^n X_i = t) / P(\Sigma_{i=1}^n X_i = t)$

$= P(X_1 = 0)\, P(\Sigma_{i=2}^n X_i = t) / P(\Sigma_{i=1}^n X_i = t)$.

Αλλα, απο την (1.54γ), $\Sigma_{i=2}^n X_i \sim \mathcal{P}oisson((n-1)\theta)$ και $\Sigma_{i=1}^n X_i \sim \mathcal{P}oisson(n\theta)$. Αρα,

$$\mu(t) = \frac{e^{-\theta} \cdot e^{-(n-1)\theta}[(n-1)\theta]^t / t!}{e^{-n\theta}(n\theta)^t / t!} = \left(\frac{n-1}{n}\right)^t = \left(1 - \frac{1}{n}\right)^t.$$

Αρα, $d^*(\underline{X}) = \mu(T(\underline{X})) = \left(1-\frac{1}{n}\right)^{T(X)} = \left(1-\frac{1}{n}\right)^{n\bar{X}_n}$,

και μια ΟΑΕΕΔ.

Επισης, εφοσον, $0 \leq d^*(\underline{X}) < 1 \Rightarrow E_\theta[d^*(\underline{X})]^2 < +\infty \ \forall \theta \in \Theta$

$\Rightarrow \mathcal{D}_\theta(d^*(\underline{X})) < +\infty \ \forall \theta \in \Theta$,

και αρα η $d^*(\underline{X})$ ειναι η μοναδικη ΟΑΕΕΔ.

(γ) Πρεπει κατ' αρχας να κατασκευασουμε καποια αμεροληπτη εκτιμητρια $d(\underline{X})$ της $q(\theta) = \theta^\nu$. Αν μαλιστα $d(\underline{X}) = h(T(\underline{X}))$ τοτε η d θα ειναι αμεσως και ΟΑΕΕΔ.

Θα βασισουμε λοιπον την κατασκευη της d απευθειας στην $T \equiv T(\underline{X}) = \Sigma_{i=1}^n X_i \sim \mathcal{P}oisson(n\theta)$. Ζητουμε λοιπον, καποια συναρτηση $h(\cdot)$ τετοια ωστε:

$$\forall \theta \in \Theta,\ \theta^\nu = E_\theta h(T) = \Sigma_{k=0}^\infty h(k)\,\frac{e^{-n\theta}(n\theta)^k}{k!}$$

$$\Leftrightarrow \forall \theta \in \Theta,\ \Sigma_{k=0}^\infty \frac{h(k)(n\theta)^k}{k!} = \theta^\nu e^{n\theta}$$

$$\Leftrightarrow \forall \theta \in \Theta,\ \Sigma_{k=0}^\infty \frac{h(k)(n\theta)^k}{k!} = \theta^\nu \Sigma_{l=0}^\infty \frac{(n\theta)^l}{l!}$$

$$\Leftrightarrow \forall \theta \in \Theta,\ \Sigma_{k=0}^\infty \frac{h(k)(n\theta)^k}{k!} = \Sigma_{l=0}^\infty \frac{(n\theta)^{l+\nu}}{n^\nu\, l!}$$

$$\Leftrightarrow \forall \theta \in \Theta,\ \Sigma_{k=0}^\infty \frac{h(k)(n\theta)^k}{k!} = \Sigma_{k=\nu}^\infty \frac{(n\theta)^k}{n^\nu (k-\nu)!}$$

Για να ειναι λοιπον ταυτοτικα ισες αυτες οι δυο δυναμοσειρες πρεπει να ειναι ισοι οι συντελεστες τους, δηλαδη, $\forall k \in \mathbb{N}_0$

$$h(k) = \begin{cases} 0 & \text{αν } k < \nu \\ \dfrac{k!}{n^\nu (k-\nu)!} & \text{αν } k \geq \nu \end{cases} = \begin{cases} 0 & \text{αν } k < \nu \\ \dfrac{(k-\nu+1)\cdots k}{n^\nu} & \text{αν } k \geq \nu. \end{cases}$$

Αρα, η εκτιμητρια:

$$d(\underline{X}) := h(T(\underline{X})) = \begin{cases} 0 & \text{αν } T(\underline{X}) < \nu \\ \dfrac{(T(\underline{X})-\nu+1)(T(\underline{X})-\nu+2)\cdots(T(\underline{X})-1)\,T(\underline{X})}{n^\nu} & \text{αν } T(\underline{X}) \geq \nu \end{cases}$$

$$= \begin{cases} 0 & \text{αν } \bar{X}_n < (\nu/n) \\ \left(\bar{X}_n - \frac{\nu-1}{n}\right)\left(\bar{X}_n - \frac{\nu-2}{n}\right)\cdots\left(\bar{X}_n - \frac{1}{n}\right)\bar{X}_n & \text{αν } \bar{X}_n \geq (\nu/n) \end{cases},$$

$$= \begin{cases} 0 & \text{αν } \overline{X}_n < \frac{\nu}{n} \\ \prod_{m=0}^{\nu-1} (\overline{X}_n - \frac{m}{n}) & \text{αν } \overline{X}_n \geq \frac{\nu}{n} \end{cases} = \prod_{m=0}^{\nu-1} (\overline{X}_n - \tfrac{m}{n})_+ \ ,$$

ειναι μια ΟΑΕΕΔ της $q(\theta) = \theta^\nu$,

οπου $(x)_+ := x\,\mathbb{1}(x>0)$.

Επισης, εφοσον $0 < d(\underline{X}) < (\overline{X}_n)^\nu = n^{-\nu} T^\nu$ και $\mathcal{D}_\theta(T^\nu) < +\infty \ \forall \theta \in \Theta$ – εφοσον $T \sim \text{Poisson}(n\theta)$, επεται οτι $\mathcal{D}_\theta(d(\underline{X})) < +\infty \ \forall \theta \in \Theta$, και αρα η εν λογω εκτιμητρια ειναι η μοναδικη ΟΑΕΕΔ της $q(\theta) = \theta^\nu$.

(3.48) Παραδειγμα. Εστω $X_1, \dots, X_n$ α.ι. $N(\mu, \sigma^2)$, $\underline{\theta} = (\mu, \sigma^2) \in \Theta \equiv \mathbb{R} \times (0, +\infty)$. Απο την (3.38) εχουμε οτι η στατιστικη $\underline{T}(\underline{X}) \equiv (T_1(\underline{X}), T_2(\underline{X})) := (\overline{X}_n, S_n^2)$ ειναι επαρκης και πληρης για την $\underline{\theta} = (\mu, \sigma^2)$.

Μας ενδιαφερει να κατασκευασουμε μια ΟΑΕΕΔ για την $q(\underline{\theta}) = \sigma^\nu$, $\nu \in \mathbb{Z}$ με $\nu > -(n-1)$.

Απο την (1.62) εχουμε οτι η

$$\frac{(n-1) S_n^2}{\sigma^2} = \frac{1}{\sigma^2} \sum_{i=1}^n (X_i - \overline{X}_n)^2 \sim \chi^2_{n-1} = \mathcal{G}\left(\frac{n-1}{2}, \frac{1}{2}\right),$$

και αρα $S_n^2 \sim \mathcal{G}\left(\frac{n-1}{2}, \frac{n-1}{2\sigma^2}\right)$ και αρα, απο την (1.38η),

$$E_{\underline{\theta}} S_n^\nu = \frac{\Gamma\left(\frac{n+\nu-1}{2}\right) 2^{\nu/2}}{\Gamma\left(\frac{n-1}{2}\right) (n-1)^{\nu/2}} \sigma^\nu \qquad \forall \theta \in \Theta .$$

Αρα, η $d(\underline{X}) := c(n,\nu) S_n^\nu$ ειναι αμεροληπτη για την $q(\underline{\theta})$, οπου $c(n,\nu) := \left(\frac{n-1}{2}\right)^{\nu/2} \Gamma\left(\frac{n-1}{2}\right) / \Gamma\left(\frac{n+\nu-1}{2}\right)$.

Αρα η $d^*(\underline{X}) := E_\theta[d(\underline{X}) \mid (\overline{X}_n, S_n^2)] = d(\underline{X}) = c(n,\nu) S_n^\nu$ ειναι μια ΟΑΕΕΔ για την $q(\underline{\theta}) = \sigma^\nu$ και μαλιστα ειναι μοναδικη διοτι $\mathcal{D}_{\underline{\theta}}(d^*(\underline{X})) < +\infty \ \forall \underline{\theta} \in \Theta$, δεδομενου οτι η $\mathcal{G}(\alpha, \lambda)$ εχει ολες τις ροπες της θετικης ταξεως.

(3.49) <u>Παράδειγμα</u>. Έστω $X_1, \dots, X_n$ α.ι. Bernoulli(p), $p \in \Theta = [0,1]$. Από τις (3.15) και (3.34) η στατιστική $T(\underline{X}) := \sum_{i=1}^n X_i$ είναι επαρκής και πλήρης για την p. Θέλουμε να κατασκευάσουμε την μοναδική ΟΑΕΕΔ για την $q(p) = p(1-p)$. (Η απόδειξη που θα δώσουμε δεν είναι η ταχύτερη, αλλά η μέθοδος της είναι χρήσιμη στην άσκηση (3.50) παρακάτω. Δώστε μια κατά πολύ ταχύτερη απόδειξη, σαν άσκηση.)

<u>Απάντηση</u>. Ψάχνουμε για μια συνάρτηση $h(\cdot)$, τέτοια ώστε η $d^*(\underline{X}) = h(T(\underline{X}))$ να είναι αμερόληπτη και άρα ΟΑΕΕΔ (γιατί;) το ότι είναι και μοναδική θα το δούμε μετά.

$$E_p(d^*(\underline{X})) = E_p h(T) = q(p) = p(1-p) \quad \forall p \in \Theta \iff$$

$$\iff \sum_{k=0}^n h(k) \binom{n}{k} p^k (1-p)^{n-k} = p(1-p) \quad \forall p \in \Theta \iff$$

$$\iff \sum_{k=0}^n h(k) \binom{n}{k} x^k = x(1+x)^{n-2} \quad \forall x = \frac{p}{1-p} \in (0,+\infty)$$

$$\iff \sum_{k=0}^n h(k) \binom{n}{k} x^k = \sum_{k=1}^{n-1} \binom{n-2}{k-1} x^k \quad \forall x > 0$$

$$\iff h(k) = \begin{cases} 0 & \text{αν } k \leq 0 \text{ ή } k \geq n \\ \dfrac{\binom{n-2}{k-1}}{\binom{n}{k}} & \text{αν } k = 1, 2, \dots, n-1 \end{cases}$$

$$= \begin{cases} 0 & \text{αν } k \leq 0 \text{ ή } k \geq n \\ \dfrac{k(n-k)}{n(n-1)} & \text{αν } k = 1, 2, \dots, n-1 . \end{cases}$$

Άρα, $d^*(\underline{X}) = h(T) = \dfrac{T(n-T)}{n(n-1)} = \dfrac{n}{n-1} \overline{X}_n (1 - \overline{X}_n)$ είναι η μοναδική ΟΑΕΕΔ για την $q(p) = p(1-p)$, εφόσον $\mathcal{D}_p(d^*(\underline{X})) < +\infty \ \forall p \in \Theta$, εφόσον η $\mathcal{B}(n,p)$ έχει τέταρτη ροπή.

(3.50) <u>Άσκηση</u>. Έστω $X_1, \dots, X_n$ α.ι. Bernoulli(p), $p \in \Theta = [0,1]$. Κατασκευάστε την μοναδική ΟΑΕΕΔ για την $q(p) = p^\nu$, $\nu \in \mathbb{N}$.

(<u>Υπόδ.</u> $d^*(\underline{X}) = \prod_{m=0}^{\nu-1} \dfrac{(T-m)_+}{n-m}$, όπου T η επαρκής και πλήρης στατιστική της p.)

(3.51) <u>Ασκηση</u>. Εστω $X_1, \dots, X_n$ α.ι. $N(0, \sigma^2)$, $\sigma^2 \in \Theta = (0, +\infty)$.
Κατασκευαστε την μοναδικη ΟΑΕΕΔ για την $q(\sigma^2) = \sigma^\nu$, $\nu \in \mathbb{N}$.
(<u>Υποδ.</u> Χρησιμοποιειστε την (3.408)

(3.52) <u>Ασκηση</u>. Εστω $X_1, \dots, X_n$ α.ι. $\mathcal{E}(\lambda)$, $\lambda \in \Theta = (0, +\infty)$.
(α) Βρητε την μοναδικη ΟΑΕΕΔ της λ.
(β) Μας ενδιαφερει τωρα η εκτιμηση της συναρτησης επιβιωσης στη θεση $s_0 > 0$, δηλαδη, η εκτιμηση της $q(\lambda) = P_\lambda(X_1 > s_0) = e^{-\lambda s_0}$. Βρητε την μοναδικη ΟΑΕΕΔ της $q(\lambda)$.
<u>Υποδ.</u> $\frac{X_1}{\sum_1^n X_i} \sim \mathrm{Beta}(1, n-1)$ ανεξ. της $\sum_1^n X_i \sim \Gamma(n, \lambda)$.

(3.53) <u>Ασκηση</u>. Εστω $X_1, \dots, X_n$ α.ι. $p(x|\theta) = \frac{1(x \in A_\theta)}{\theta}$,
$A_\theta = \{1, 2, \dots, \theta\}$, $\theta \in \mathbb{N}$, δηλαδη, οι X_i ειναι διακριτες ομοιομορφες.
(α) Δειξτε οτι $T(\underline{X}) = X_{nn}$ ειναι επαρκης και πληρης για την θ.
(β) Κατασκευαστε τη μοναδικη ΟΑΕΕΔ $d^*(\underline{X})$ για την $q(\theta) = \theta$.
(<u>Υποδ.</u> (i) $T \sim p(t|\theta) = \frac{t^n - (t-1)^n}{\theta^n} 1(t \in A_\theta)$

$$(ii)\ d^*(\underline{x}) = T + \frac{(T-1)^n}{T^n - (T-1)^n} = \frac{X_{nn}^{n+1} - (X_{nn}-1)^{n+1}}{X_{nn}^n - (X_{nn}-1)^n}.\ )$$

(3.54) <u>Ασκηση</u>. Εστω $X_1, \dots, X_n$ α.ι. $N(\mu, \sigma^2)$. Δειξτε οτι:
(α) αν $\sigma = \sigma_0$ -γνωστο- και $\mu \in \Theta = \mathbb{R}$ η $\bar{X}_n$ ειναι η μοναδικη ΟΑΕΕΔ για την μ.
(β) αν $\mu = \mu_0$ -γνωστο- και $\sigma^2 \in \Theta = (0, +\infty)$ η $\frac{1}{n}\sum_{i=1}^n (X_i - \mu_0)^2$ ειναι η μοναδικη ΟΑΕΕΔ για την σ^2.

(3.55) <u>Ασκηση</u>. Εστω $X_1, \dots, X_n$ α.ι. $N(\mu, 1)$, $\mu \in \Theta = \mathbb{R}$.
Βρητε τη μοναδικη ΟΑΕΕΔ για την $q(\mu) = P_\mu(X_1 \geq 0) = \Phi(\mu)$.
(<u>Υποδ.</u> Η δ.τ.μ. $(X_1, \bar{X}_n)$ ακολουθει μια διδιαστατη κανονικη κατανομη.
Βλεπε, επισης, την (1.79).)

(3.56) Σημ. Τα κυρια μειονεκτηματα των ΟΑΕΕΔ ειναι τα εξης:

(α) ΟΑΕΕΔ μπορει να μην υπαρχουν σε συγκεκριμενα προβληματα.

(β) Ο υπολογισμος μιας ΟΑΕΕΔ μπορει να ειναι δυσκολος και ο υπολογισμος της αποδοσης τους, π.χ., της διασπορας τους, δυσκολοτερος.

(γ) Μια ΟΑΕΕΔ μπορει να ειναι ομοιομορφα χειροτερη μιας μη αμεροληπτης εκτιμητριας, ακομη και με κριτηριο το μεσο τετραγωνικο σφαλμα. Για παραδειγμα,

απο το (3.48) με $\nu=2$, βλεπουμε οτι η μοναδικη ΟΑΕΕΔ της $q(\underline{\theta}) = q(\mu,\sigma^2) = \sigma^2$ ειναι η

$S_n^2 := \frac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X}_n)^2$,

με $(n-1)S_n^2/\sigma^2 \sim \chi^2_{n-1}$ και αρα

$$R(\theta, S_n^2) = \mathcal{D}_\theta(S_n^2) = \frac{\sigma^4}{(n-1)^2}\mathcal{D}_\theta\left(\frac{(n-1)}{\sigma^2}S_n^2\right) = \frac{\sigma^4}{(n-1)^2}\frac{\frac{n-1}{2}}{1/4} =$$
$$= \frac{2\sigma^4}{n-1}, \quad \underline{\theta} = (\mu,\sigma^2) \in \Theta = \mathbb{R}\times(0,+\infty)$$

Ομως, η μη αμεροληπτη εκτιμητρια (μεγιστης πιθανοφανειας),

(3.57) $\quad \hat{\sigma}_n^2 := \frac{1}{n}\sum_{i=1}^n (X_i - \overline{X}_n)^2$,

με μεροληψια $b_\theta(\hat{\sigma}_n^2) = E_\theta(\hat{\sigma}_n^2) - \sigma^2 = \frac{n-1}{n}\sigma^2 - \sigma^2 = -\frac{\sigma^2}{n}$,

εχει διασπορα $\mathcal{D}_\theta(\hat{\sigma}_n^2) = \left(\frac{n-1}{n}\right)^2 \mathcal{D}_\theta(S_n^2) = \frac{2\sigma^4(n-1)}{n^2}$,

και αρα μεσο τετραγωνικο σφαλμα – βλ.(2.13),

$$R(\theta, \hat{\sigma}_n^2) = \mathcal{D}_\theta(\hat{\sigma}_n^2) + b_\theta^2(\hat{\sigma}_n^2) = \frac{2\sigma^4(n-1)}{n^2} + \frac{\sigma^4}{n^2} = \frac{(2n-1)\sigma^4}{n^2}.$$

Εχουμε δε οτι, $\forall \theta \in \Theta$,

$$R(\theta, \hat{\sigma}_n^2) < R(\theta, S_n^2) \Leftrightarrow \frac{(2n-1)\sigma^4}{n^2} < \frac{2\sigma^4}{n-1} \Leftrightarrow$$
$$\Leftrightarrow (2n-1)(n-1) < 2n^2 \Leftrightarrow 2n^2 - 3n + 1 < 2n^2 \Leftrightarrow n > 1/3$$

Δηλαδη, $\forall n \in \mathbb{N}$ και $\forall \underline{\theta} = (\mu,\sigma^2) \in \Theta = \mathbb{R}\times(0,+\infty)$

$$R(\underline{\theta}, \hat{\sigma}_n^2) < R(\underline{\theta}, S_n^2).$$

Η μικρη μεροληψια $b_\theta(\hat{\sigma}_n^2) = -\frac{\sigma^2}{n}$ της $\hat{\sigma}_n^2$ ομως, την ειχε αποκλεισει απο το συνολο $\mathcal{U}_\Theta$ οπου ψαχναμε την ΟΑΕΕΔ.

3.2. ΕΚΤΙΜΗΤΡΙΕΣ ΤΗΣ ΜΕΘΟΔΟΥ ΤΩΝ ΡΟΠΩΝ.

Η μέθοδος των ροπών είναι μια γρήγορη μέθοδος κατασκευής εκτιμητριών, όχι πάντα "καλών", οι οποίες μπορούν να χρησιμοποιηθούν σαν προκαταρκτικές εκτιμήτριες σε αριθμητικούς αλγορίθμους βελτιστοποίησης που οδηγούν, π.χ., σε καθοια εκτιμήτρια μεγίστης πιθανοφάνειας που δεν μπορεί να υπολογισθεί σε κλειστή μορφή. Οι εκτιμήτριες αυτές δεν είναι εν γένει μοναδικές.

Έστω $X_1, X_2, \ldots, X_n$ στοχαστικό δείγμα από την F_θ, $\theta \in \Theta$, και έστω ότι ενδιαφερόμαστε για την εκτίμηση της $q(\theta)$ την οποία μπορούμε να εκφράσουμε σαν μια συνεχή συνάρτηση g των m πρώτων ροπών

$$(3.58)\quad \mu_k(\theta) \equiv T_k(F_\theta) := \int_{\mathbb{R}} x^k \, dF_\theta(x) = \begin{cases} \sum_x x^k p(x|\theta) & \text{αν } F_\theta \text{ διακριτή} \\ \int_{\mathbb{R}} x^k f(x|\theta)dx & \text{αν } F_\theta \text{ συνεχής}, \end{cases}$$

$k = 1, \ldots, m$, εφόσον αυτές υπάρχουν. Δηλαδή,

$$q(\theta) = g(\mu_1(\theta), \ldots, \mu_m(\theta)), \quad \theta \in \Theta .$$

Η <u>εκτιμήτρια της μεθόδου των ροπών</u> ορίζεται τότε ως:

$$(3.59)\quad \tilde{q} \equiv d_p(\underline{X}) := g(\hat{\mu}_1, \ldots, \hat{\mu}_m)$$

$$= g(T_1(\mathbb{F}_n), \ldots, T_m(\mathbb{F}_n)),$$

όπου, $\hat{\mu}_k \equiv T_k(\mathbb{F}_n) = \int_{\mathbb{R}} x^k \, d\mathbb{F}_n(x) = \frac{1}{n}\sum_{i=1}^n X_i^k$,

$k = 1, 2, \ldots, m$, είναι οι <u>εμπειρικές ροπές</u>, και

$$(3.60)\quad \mathbb{F}_n(x) := \frac{1}{n}\sum_{i=1}^n 1(X_i \le x), \quad x \in \mathbb{R},$$

είναι η <u>εμπειρική συνάρτηση κατανομής</u> – βλ. (1.89).

(3.61) <u>Παραδειγμα</u>. Εστω $X_1, \dots, X_n$ α.ι. $\mathcal{G}(\alpha,\lambda)$, $\theta=(\alpha,\lambda)\in\Theta \equiv$ $\equiv (0,+\infty)\times(0,+\infty)$. Βρητε τις εκτιμητριες των ροπων:

(α) της παραμετρου λ,

(β) της παραμετρου α.

<u>Απαντηση</u>. (α) $q_1(\theta) = \lambda = \dfrac{E_\theta(X)}{\mathcal{D}_\theta(X)} = \dfrac{E_\theta X}{E_\theta X^2 - (E_\theta X)^2} = \dfrac{\mu_1(\theta)}{\mu_2(\theta)-\mu_1^2(\theta)}$,

και αρα $\tilde{q}_1 = \tilde{\lambda}_n = \dfrac{\hat{\mu}_1}{\hat{\mu}_2 - (\hat{\mu}_1)^2} = \dfrac{\frac{1}{n}\Sigma_{i=1}^n X_i}{\frac{1}{n}\Sigma_{i=1}^n X_i - (\frac{1}{n}\Sigma_{i=1}^n X_i)^2}$

$$= \frac{\overline{X}_n}{\frac{1}{n}\Sigma_{i=1}^n X_i - \overline{X}_n^{\,2}} = \frac{\overline{X}_n}{\frac{1}{n}\Sigma_{i=1}^n (X_i-\overline{X}_n)^2} = \frac{\overline{X}_n}{\hat{\sigma}_n^2}$$

(β) $q_2(\theta) = \alpha = \dfrac{(E_\theta X)^2}{\mathcal{D}_\theta(X)} = \dfrac{\mu_1^2(\theta)}{\mu_2(\theta)-\mu_1^2(\theta)}$,

και αρα, $\tilde{q}_2 = \tilde{\alpha}_n = \dfrac{\overline{X}_n^{\,2}}{\hat{\sigma}_n^2}$.

Η κατασκευη μιας πιο "σωστης" εκτιμητριας για την παραμετρο α ειναι πολυ πιο δυσκολη.

(3.62) <u>Παραδειγμα</u>. Εστω $X_1, \dots, X_n$ α.ι. διακριτες ομοιομορφες στο $A_\theta := \{1,2,\dots,\theta\}$, $\theta\in\Theta\equiv\mathbb{N}$, δηλαδη,

$p(x/\theta) = \theta^{-1}\, 1(x\in A_\theta)$. Βρητε την εκτιμητρια της μεθοδου των ροπων για την παραμετρο $\theta\in\mathbb{N}$.

<u>Απαντηση</u>. Εχουμε οτι $E_\theta X = \Sigma_{x=1}^{\theta} x\, p(x|\theta) = \Sigma_{x=1}^{\theta} x \frac{1}{\theta} =$

$= \theta^{-1}\Sigma_{x=1}^{\theta} x = \frac{1}{\theta}\,\frac{\theta(\theta+1)}{2} = \frac{\theta+1}{2} \Rightarrow$

$\Rightarrow \theta = 2E_\theta X - 1 = 2\mu_1(\theta) - 1$

Αρα η εκτιμητρια της μεθοδου των ροπων ειναι η

$\tilde{\theta}_n = 2\overline{X}_n - 1$

Παρατηρησε οτι μπορει καλλιστα να συμβη το "παραλογο":

$2\overline{X}_n - 1 = \tilde{\theta}_n < X_{nn}$ ή και αλλων παρατηρησεων, ενω ολες τους ειναι φυσικα μικροτερες ή ισες του θ. Εν τουτοις για n "μεγαλο" οι εκτιμητριες της μεθοδου των ροπων πλησιαζουν την αληθη τιμη της παραμετρου (ειναι δηλαδη συνεπεις — βλ. 3.), π.χ., απο τον Νομο των μεγαλων αριθμων — βλ. 1.86 — εχουμε οτι:

$$\tilde{\theta}_n = 2\bar{X}_n - 1 \xrightarrow[n\to\infty]{P} 2E_\theta X - 1 = 2\cdot\frac{\theta+1}{2} - 1 = \theta \ \forall \theta \in \Theta.$$

(3.63) <u>Ασκηση</u>. Εστω $X_1, \dots, X_n$ α.ι. $N(\mu, \sigma^2)$, $\underline{\theta} = (\mu, \sigma^2) \in \Theta = \mathbb{R} \times (0, +\infty)$. Βρητε τις εκτιμητριες της μεθοδου των ροπων $\tilde{\mu}_n$ και $\tilde{\sigma}_n^2$ των μ, σ^2 και δειξτε οτι $\tilde{\mu}_n \xrightarrow{P} \mu$ και $\tilde{\sigma}_n^2 \xrightarrow{P} \sigma^2$ καθως το $n \to +\infty$.

(3.64) <u>Ασκηση</u>. Εστω $X_1, \dots, X_n$ α.ι. $\mathcal{B}\text{ητα}(\alpha, \beta)$, $\alpha, \beta > 0$. Βρητε τις εκτιμητριες της μεθοδου των ροπων για τις α, β.

(3.65) <u>Ασκηση</u>. Εστω $X_1, \dots, X_n$ α.ι. $\mathcal{D}(m, p)$, $\theta = (m, p) \in \Theta = \mathbb{N} \times [0, 1]$. Βρητε τις εκτιμητριες της μεθοδου των ροπων για τις παραμετρους m, p.

(3.66) <u>Σημ.</u> Η μεθοδος των ροπων γενικευεται πολυ απλα στην περιπτωση πολυδιαστατων τυχαιων μεταβλητων. Για παραδειγμα, εστω $(X_1, Y_1), \dots, (X_n, Y_n)$ α.ι. $F(x, y|\theta)$, $\theta \in \Theta$, $(x, y) \in \mathbb{R}^2$. Η εκτιμητρια των ροπων της

$$q(\theta) = g\left(\mu_i^x(\theta), \mu_j^y(\theta), \mu_{ij}^{xy}(\theta), \ i, j = 1, 2, \dots\right)$$

οριζεται η

$$\tilde{q}(\theta) = g\left(\hat{\mu}_i^x, \hat{\mu}_j^y, \hat{\mu}_{ij}^{xy}, \ i, j = 1, 2, \dots\right),$$

οπου,

$$\hat{\mu}_k^x := \frac{1}{n}\Sigma_{i=1}^n X_i^k, \quad \hat{\mu}_k^y := \frac{1}{n}\Sigma_{i=1}^n Y_i^k, \quad \hat{\mu}_{kl}^{xy} := \frac{1}{n}\Sigma_{i=1}^n X_i^k Y_i^l.$$

(3.67) <u>Ασκηση</u>. Εστω (X_i, Y_i), $i = 1, \dots, n$ α.ι. δ.τ.μ. με $\rho := \dfrac{\mathrm{Cov}(X_1, Y_1)}{\sqrt{\mathcal{D}(X_1)\mathcal{D}(Y_1)}}$.
Δειξτε οτι $\tilde{\rho} = \dfrac{\Sigma_{i=1}^n (X_i - \bar{X}_n)(Y_i - \bar{Y}_n)}{\sqrt{\Sigma_1^n (X_i - \bar{X}_n)^2 \Sigma_1^n (Y_i - \bar{Y}_n)^2}}$.

3.3. ΕΚΤΙΜΗΤΡΙΕΣ ΜΕΓΙΣΤΗΣ ΠΙΘΑΝΟΦΑΝΕΙΑΣ.

Εστω $\underline{X}=(X_1,\ldots,X_n)$ α.ε. $p(x|\theta)$, $\theta\in\Theta$, και εστω οτι το αποτελεσμα ω_0 ενος πειραματος εδωσε το δειγμα $\underline{X}(\omega_0)=\underline{x}^0$, βασει του οποιου θελουμε να εκτιμησουμε την αληθη τιμη θ_0 της αγνωστης παραμετρου $\theta\in\Theta$, δηλαδη, την αληθη κατανομη $p(\cdot|\theta_0)$ η οποια δημιουργησε το δειγμα. Αν η τιμη της αγνωστης παραμετρου ηταν θ, τοτε, η πιθανοτητα να παρατηρηθει ακριβως το συγκεκριμμενο αυτο δειγμα $\underline{x}^0$ ειναι

$$(3.68)\qquad P_\theta(\underline{X}=\underline{x}^0)=\prod_{i=1}^{n}p(x_i^0|\theta).$$

Δεδομενου λοιπον, του οτι απο ολα τα στοιχεια του δειγματικου χωρου $\mathcal{X}$, η αγνωστη $p(\cdot|\theta_0)$, δηλαδη, η θ_0, εξελεξε το $\underline{x}^0$, ειναι διαισθητικα ικανοποιητικη εκτιμητρια της αγνωστης θ_0, εκεινο το $\hat{\theta}\in\Theta$ υπο το οποιο μεγιστοποιειται η πιθανοτητα $P_\theta(\underline{X}=\underline{x}^0)$ παρατηρησης του συγκεκριμμενου δειγματος $\underline{x}^0$ το οποιο πραγματι συνεβη, δηλαδη, το $\hat{\theta}\in\Theta$ ειναι τετοιο ωστε

$$(3.69)\qquad P_{\hat{\theta}}(\underline{X}=\underline{x}^0)=\max_{\theta\in\Theta}P_\theta(\underline{X}=\underline{x}^0),$$

αν φυσικα ενα τετοιο μεγιστο υπαρχει.

Στην περιπτωση που το δειγμα προερχεται απο καποια απολυτα συνεχη κατανομη $f(\cdot|\theta)$, $\theta\in\Theta$, η (3.68) μπορει να αντικατασταθη απο την

$$(3.70)\qquad \begin{aligned} P_\theta(x_i^0-\varepsilon<X_i\le x_i^0+\varepsilon,\ i=1,\ldots,n)&=\\ =\prod_{i=1}^{n}\left[F(x_i^0+\varepsilon|\theta)-F(x_i^0-\varepsilon|\theta)\right]&\\ \approx(2\varepsilon)^n\prod_{i=1}^{n}f(x_i^0|\theta),\qquad \theta\in\Theta,\quad &\forall\varepsilon>0\text{ "μικρο"},\end{aligned}$$

και οι προηγουμενοι συλλογισμοι μας ισχυουν και τωρα, τουλαχιστον κατα προσεγγιση, οδηγουν δε τωρα σε εκεινη την εκτιμητρια $\hat{\theta}$ της θ_0 για την οποια ισχυει οτι

$$(3.71)\qquad f(\underline{x}^0|\hat{\theta})=\max_{\theta\in\Theta}f(\underline{x}^0|\theta),$$

εφοσον φυσικα ενα τετοιο μεγιστο υπαρχει.

Οι συλλογισμοί που οδήγησαν στην εκτιμήτρια $\hat{\theta}$ της (3.69) ή της (3.71), γίνονται αθλουστεροι (αλλα και εξηρτημένοι απο την Θεωρια Bayes) αν δεχθη κανεις μια "a priori" κατανομη $\pi(\theta)$, $\theta \in \Theta$, για την αγνωστη παραμετρο, η οποια υπο την επιδραση των δεδομενων $\underline{X} = \underline{x}^o$ τροποποιειται (βλ. (1.72) – (1.74)) στην "a posteriori" κατανομη:

$$(3.72) \quad \pi(\theta \mid \underline{x}^o) = \frac{f(\underline{x}^o \mid \theta)\,\pi(\theta)}{\int_{\Theta} f(\underline{x}^o \mid \theta)\,\pi(\theta)\,d\theta}, \quad \theta \in \Theta,$$

και αρα η $\hat{\theta}$ της (3,69) ή της (3,71), μεγιστοποιει την "a posteriori" πιθανοτητα ή πυκνοτητα της θ, δεδομενου του δειγματος $\underline{x}^o$, δηλαδη,

$$(3,73) \quad \pi(\hat{\theta} \mid \underline{x}^o) = \max_{\theta \in \Theta} \pi(\theta \mid \underline{x}^o),$$

αν φυσικα η $\pi(\theta|\underline{x}^o)$ εχει κορυφη $\hat{\theta}$.

(3.73) Ορισμος. Εστω $X_1, \dots, X_n$ α.ι. $p(\cdot|\theta)$ ή $f(\cdot|\theta)$, $\theta \in \Theta \subseteq \mathbb{R}^k$. Η στοχαστικη συναρτηση $L : \Theta \to \mathbb{R}$, με

$$L_n(\theta) \equiv L_n(\theta \mid \underline{X}) := \begin{cases} \prod_{i=1}^{n} p(X_i \mid \theta) & \text{αν οι } X_i \text{ ειναι διακριτες} \\ \prod_{i=1}^{n} f(X_i \mid \theta) & \text{αν οι } X_i \text{ ειναι απολ. συνεχεις,} \end{cases}$$

καλειται <u>στοχαστικη συναρτηση πιθανοφανειας</u>, ή απλως συναρτηση πιθανοφανειας της παραμετρου $\theta \in \Theta$.

Παρατηρουμε οτι $\forall\, \underline{x} \in \mathcal{X}$ η πιθανοφανεια της θ στο $\underline{x}$, δηλαδη, η $L(\theta \mid \underline{X} = \underline{x})$, ειναι ακριβως η συναρτηση μαζας πιθανοτητος ή η πυκνοτητα του στοχαστικου δειγματος $\underline{X}$ στο σημειο $\underline{x}$, υπολογισμενες με παραμετρο το θ. Βλεπουμε ομως την πιθανοφανεια σαν συναρτηση της θ – για δεδομενο δειγμα – και οχι σαν συναρτηση του δειγματος.

(3.74) <u>Κατασκευη των Εκτιμητριων Μεγιστης Πιθανοφανειας</u> (ε.μ.π.):

Εστω οτι $\forall \theta \in \Theta \subseteq \mathbb{R}^k$ με πιθανοτητα (F_θ) 1 υπαρχει μεγιστο

$$\max_{\theta \in \Theta} L(\theta) \equiv \max_{\theta \in \Theta} L(\theta / \underline{X}) .$$

Αυτο συμβαινει π.χ., αν η $L(\theta)$ ειναι κοιλη συναρτηση της $\theta \in \Theta$.

Τοτε, καθε στατιστικη $\hat{\theta}_n \equiv \hat{\theta}(\underline{X}) : \Omega \ni \omega \longmapsto \hat{\theta}(\underline{X}(\omega)) = \hat{\theta}(\underline{x}) \in \Theta$

τετοια ωστε

$$(3.75) \quad L(\hat{\theta}_n) = \max_{\theta \in \Theta} L(\theta) ,$$

καλειται <u>εκτιμητρια μεγιστης πιθανοφανειας</u> (εμπ) της παραμετρου $\theta \in \Theta$. Μια εμπ —αν υπαρχη— δεν ειναι απαραιτητα μοναδικη.

Εστω $k = \dim \Theta = 1$, και εστω οτι $\exists\ L'(\theta)$, τοτε εφοσον υπαρχει η εμπ $\hat{\theta}_n$, πρεπει $L'(\hat{\theta}_n) = 0$, εχουμε λοιπον να λυσουμε την εξισωση

$$(3.76) \quad L'(\theta) = 0 ,$$

εστω δε θ^* μια λυση της (3.76), και εστω επιπλεον οτι

$$(3.77) \quad \exists\ L''(\theta^*) < 0 ,$$

τοτε $\hat{\theta}_n := \theta^*$ ειναι μια (οχι απαραιτητα μοναδικη) ε.μ.π.

Αν γενικα $k \in \mathbb{N}$, οι (3.76) και (3.77) πρεπει να αντικατασταθουν αντιστοιχα απο τις:

$$(3.78) \quad \nabla L(\underline{\theta}) \equiv \left(\frac{\partial}{\partial \theta_1} L(\underline{\theta}), \ldots, \frac{\partial}{\partial \theta_k} L(\underline{\theta}) \right) = \underline{0}$$

και μια λυση $\underline{\theta}^*$ του συστηματος αυτου των εξισωσεων με Hessian

$$(3.79) \quad \underline{\underline{H}}'(\underline{\theta}^*) := \left[\frac{\partial^2}{\partial \theta_i \partial \theta_j} L(\underline{\theta}) \right]_{\underline{\theta} = \underline{\theta}^*} < 0 ,$$

ειναι μια εμπ $\underline{\hat{\theta}}_n = \underline{\theta}^*$, οπου

$\underline{\underline{H}}' < 0$ σημαινει οτι ο $k \times k$ πινακας $\underline{\underline{H}}'(\underline{\theta}^*)$ ειναι αρνητικα ορισμενος, δηλαδη, $\forall \underline{\alpha} \in \mathbb{R}^k \setminus \{\underline{0}\}$ $\underline{\alpha}^T \underline{\underline{H}}' \underline{\alpha} < 0$.

Επειδη για ανεξαρτητα δειγματα η πιθανοφανεια ειναι ενα γινομενο – βλ. (3.73) – ειναι ευκολοτερο και ισοδυναμο να μεγιστοποιηθη ο λογαριθμος της πιθανοφανειας

(3.80) $$\ell_n(\theta) \equiv \ell_n(\theta \mid \underline{X}) := \log L_n(\theta) \equiv \log L_n(\theta \mid \underline{X})$$
$$= \begin{cases} \sum_{i=1}^n \log p(X_i \mid \theta) & \text{αν οι } X_i \text{ ειναι διακριτες,} \\ \sum_{i=1}^n \log f(X_i \mid \theta) & \text{αν οι } X_i \text{ ειναι απ. συνεχεις.} \end{cases}$$

Εφοσον η συναρτηση $\log(\cdot)$ ειναι αυξουσα τα σημεια μεγιστου της $L(\cdot)$, αν υπαρχουν, ειναι τα ιδια με τα σημεια μεγιστου της $\ell(\cdot)$.

Ετσι, ψαχνουμε την ε.μ.π. $\hat{\underline{\theta}}_n$ ως λυσεις $\underline{\theta}^*$ του συστηματος:

(3.81) $$\nabla \ell_n(\underline{\theta}) = \sum_{i=1}^n \nabla \log f(X_i \mid \underline{\theta}) = \underline{0} \quad ,$$

οι εξισωσεις του οποιου καλουνται εξισωσεις μεγιστης πιθανοφανειας, οι οποιες καθιστουν την Hessian αρνητικα ορισμενη:

(3.82) $$H(\underline{\theta}^*) := \left[\frac{\partial^2}{\partial \theta_i \partial \theta_j} \ell_n(\underline{\theta}) \right]$$
$$= \left[\sum_{k=1}^n \frac{\partial^2}{\partial \theta_i \partial \theta_j} \log f(X_k \mid \underline{\theta}) \right]_{\underline{\theta} = \underline{\theta}^*} < 0 \ .$$

Φυσικα αν οι X_i ειναι διακριτες η f αντικαθισταται απο την p ως (3.81) και (3.82).

Εφοσον λοιπον οι σχετικες παραγωγοι υπαρχουν θα βρισκουμε τις ε.μ.π. $\hat{\theta}_n$ μεσω των (3.81) και (3.82). Αν η (3.81) δεν διδει την $\hat{\theta}_n$ σε κλειστη μορφη θα την λυνουμε με αριθμητικες μεθοδους – βλ. (3.110).

(3.83) <u>Παραδειγμα</u>: Εστω $X_1, \dots, X_n$ α.ι. $N(\mu,\sigma^2)$, $\underline{\theta}=(\mu,\sigma^2)\in$
$\in \Theta \equiv \mathbb{R}\times(0,+\infty)$. Δειξτε οτι:

(α) η ε.μ.π. της μ ειναι η $\hat{\mu}_n=\overline{X}_n$.

(β) αν $\mu=\mu_0$ <u>γνωστο</u>, τοτε η ε.μ.π. της σ^2 ειναι η $\hat{\sigma}_0^2=\frac{1}{n}\Sigma_1^n(X_i-\mu_0)^2$.

(γ) η ε.μ.π. της $\underline{\theta}=(\mu,\sigma^2)$, ειναι η $\underline{\hat{\theta}}_n=(\hat{\mu}_n,\hat{\sigma}_n^2)$
οπου $\hat{\mu}_n=\overline{X}_n$ και $\hat{\sigma}_n^2=\frac{1}{n}\Sigma_{i=1}^n(X_i-\overline{X}_n)^2$ (οι οποιες
και ειναι στοχαστικα ανεξαρτητες και $\overline{X}_n\sim N(\mu,\frac{\sigma^2}{n})$,
$n\hat{\sigma}_n^2/\sigma^2\sim\chi^2_{n-1}$ — βλ. (1.62)) (επισης, βλ. 3.57).

<u>Απαντηση</u>. (α) Η πιθανοφανεια της $\underline{\theta}$ ειναι η:

$$L(\underline{\theta})=L(\underline{\theta}|\underline{X})=\prod_{i=1}^n\left\{\frac{1}{\sigma}\phi\left(\frac{X_i-\mu}{\sigma}\right)\right\}=$$

$$=(2\pi\sigma^2)^{-n/2}\exp\left\{-\frac{1}{2\sigma^2}\Sigma_{i=1}^n(X_i-\mu)^2\right\}\Rightarrow$$

(3.84) $$\ell(\underline{\theta})\equiv\ell(\underline{\theta}|\underline{X})=\log L(\underline{\theta})=$$
$$=-\frac{1}{2\sigma^2}\Sigma_{i=1}^n(X_i-\mu)^2-\frac{n}{2}\log(2\pi\sigma^2),$$

η οποια εχει δευτερη παραγωγο ως προς μ και αρα απο τις (3.81) και (3.82) η ε.μ.π. της μ ειναι η λυση της εξισωσης: $\frac{\partial}{\partial\mu}\ell(\mu,\sigma^2)=0 \Leftrightarrow$

$$\Leftrightarrow -\frac{1}{2\sigma^2}\Sigma_{i=1}^n 2(X_i-\mu)(-1)=0 \Leftrightarrow \Sigma_{i=1}^n(X_i-\mu)=0$$
$$\Leftrightarrow \Sigma_{i=1}^n X_i-n\mu=0 \Leftrightarrow \mu=\overline{X}_n,$$

δηλαδη $\forall\sigma^2>0$, $\hat{\mu}_n=\overline{X}_n$,

εφοσον, $\forall\sigma^2>0$,
$$\frac{\partial^2}{\partial\mu^2}\ell(\mu,\sigma^2)\Big|_{\mu=\hat{\mu}_n}<0$$ το οποιο ισχυει διοτι,

$$\frac{\partial^2}{\partial\mu^2}\ell(\mu,\sigma^2)=\frac{\partial}{\partial\mu}\left\{\frac{1}{\sigma^2}\Sigma_{i=1}^n(X_i-\mu)\right\}=-\frac{n}{\sigma^2}<0.$$

(β) Εφοσον η $\mu=\mu_0$ ειναι γνωστη απο την (3.84) εχουμε
$$\ell(\sigma^2)=-\frac{1}{2\sigma^2}\Sigma_{i=1}^n(X_i-\mu_0)^2-\frac{n}{2}\log(2\pi\sigma^2),$$

και αρα η εμπ της σ^2 ειναι η λυση της εξισωσης

$$l'(\sigma^2) = 0 \iff \frac{1}{2(\sigma^2)^2}\sum_{i=1}^n (X_i-\mu_0)^2 - \frac{n}{2}\frac{2\pi}{2\pi\sigma^2} = 0$$

$$\iff \sum_{i=1}^n (X_i-\mu_0)^2 - n\sigma^2 = 0 \iff \sigma^2 = \frac{1}{n}\sum_{i=1}^n (X_i-\mu_0)^2$$

Αρα, η εμπ της σ^2 ειναι η $\hat{\sigma}_0^2 := \frac{1}{n}\sum_{i=1}^n (X_i-\mu_0)^2$,

εφοσον,

$$l''(\sigma^2) = \frac{d}{d\sigma^2}\left(\frac{n\hat{\sigma}_0^2}{2(\sigma^2)^2} - \frac{n}{2\sigma^2}\right) = -\frac{n\hat{\sigma}_0^2}{(\sigma^2)^3} + \frac{n}{2(\sigma^2)^2}$$

και αρα $l''(\hat{\sigma}_0^2) = -\frac{n\hat{\sigma}_0^2}{(\hat{\sigma}_0^2)^3} + \frac{n}{2(\hat{\sigma}_0^2)^2} = -\frac{n}{2(\hat{\sigma}_0^2)^2} < 0$.

(γ) Η πιθανοφανεια της $\underline{\theta} = (\mu, \sigma^2)$ ειναι η (3.84) και αρα, απο την (3.81) οι ε.μ.π. των μ, σ^2 ειναι οι λυσεις των εξισωσεων:

$$\left.\begin{aligned}\frac{\partial}{\partial\mu} l(\mu,\sigma^2) &= 0\\ \frac{\partial}{\partial\sigma^2} l(\mu,\sigma^2) &= 0\end{aligned}\right\} \iff \left\{\begin{aligned} n(\overline{X}_n-\mu) &= 0\\ \sum_{i=1}^n (X_i-\mu)^2 - n\sigma^2 &= 0\end{aligned}\right\} \iff$$

$$\iff \left\{\begin{aligned}\mu &= \overline{X}_n\\ \sigma^2 &= \frac{1}{n}\sum_{i=1}^n (X_i-\mu)^2\end{aligned}\right\} \iff \left\{\begin{aligned}\mu &= \overline{X}_n\\ \sigma^2 &= \frac{1}{n}\sum_{i=1}^n (X_i-\overline{X}_n)^2\end{aligned}\right.,$$

δηλαδη, οι εμπ των μ, σ^2 ειναι οι

$\hat{\mu}_n := \overline{X}_n$ και $\hat{\sigma}_n^2 := \frac{1}{n}\sum_{i=1}^n (X_i-\overline{X}_n)^2$,

εφοσον η Hessian ειναι αρνητικα ορισμενη:

$$H(\hat{\mu}_n, \hat{\sigma}_n^2) = \begin{pmatrix}\frac{\partial^2}{\partial\mu^2} l(\underline{\theta}) & \frac{\partial^2}{\partial\sigma^2\partial\mu} l(\underline{\theta})\\ \frac{\partial^2}{\partial\mu\partial\sigma^2} l(\underline{\theta}) & \frac{\partial^2}{\partial(\sigma^2)^2} l(\underline{\theta})\end{pmatrix}_{\underline{\theta}=\hat{\underline{\theta}}} =$$

$$= \begin{pmatrix}-\frac{n}{\sigma^2} & -\frac{1}{(\sigma^2)^2}\sum_{i=1}^n (X_i-\mu)\\ -\frac{1}{(\sigma^2)^2}\sum_{i=1}^n (X_i-\mu) & -\frac{1}{(\sigma^2)^3}\sum_{i=1}^n (X_i-\mu)^2 + \frac{n}{2(\sigma^2)^2}\end{pmatrix}_{\underline{\theta}=\hat{\underline{\theta}}} =$$

$$= \begin{pmatrix}-\frac{n}{\hat{\sigma}_n^2} & 0\\ 0 & -\frac{n}{2(\hat{\sigma}_n^2)^2}\end{pmatrix} = -\frac{n}{2(\hat{\sigma}_n^2)^2}\begin{pmatrix}2\hat{\sigma}_n^2 & 0\\ 0 & 1\end{pmatrix},$$

και αρα $\forall \underline{\alpha} \in \mathbb{R}^2 \setminus \{\underline{0}\}$, $\underline{\alpha}^T H(\hat{\underline{\theta}}_n)\, \underline{\alpha} = -\frac{n}{2(\hat{\sigma}_n^2)^2}\left(2\hat{\sigma}_n^2 \alpha_1^2 + \alpha_2^2\right) < 0$.

Θα μελετησουμε τωρα δυο περιπτωσεις κατασκευης ε.μ.π., οπου δεν εχουμε τη δυνατοτητα να παραγωγισουμε την συναρτηση πιθανοφανειας της παραμετρου.

(3.85) Παραδειγμα. Εστω $X_1, \dots, X_n$ α.ι. $\mathcal{U}(0,\theta)$, $\theta \in \Theta \equiv (0,+\infty)$. Δειξτε οτι η ε.μ.π. της θ ειναι η $\hat{\theta}_n = X_{nn}$.

Απαντηση. $L(\theta) \equiv L(\theta/\underline{X}) = \prod_{i=1}^{n} f(X_i|\theta) = \theta^{-n} \prod_{i=1}^{n} 1(0 \le X_i \le \theta)$

$= \frac{1}{\theta^n} 1(X_{nn} \le \theta)\, 1(X_{n1} \ge 0) = 1(X_{n1} \ge 0)\, \frac{1(\theta \ge X_{nn})}{\theta^n}$,

η οποια ειναι μια φθινουσα συναρτηση του θ, εφοσον $\theta \ge X_{nn}$. Αρα, η $L(\theta)$ μεγιστοποιειται για την μικροτερη τιμη του θ, τετοια ωστε $\theta \ge X_{nn}$

αρα, $L(\theta) \le L(X_{nn}) = X_{nn}^{-n} 1(X_{n1} > 0)$ $\forall \theta \in \Theta$,

δηλαδη, $\hat{\theta}_n = X_{nn}$.

(Σημ. Αν $f(x|\theta) = \frac{1}{\theta} 1(0 < x < \theta)$ αντι για $\frac{1}{\theta} 1(0 \le x \le \theta)$, τοτε $\hat{\theta}_n = X_{nn}^+$. Αλλα $P(X_{nn}^+ = X_{nn}) = 1$, και θετουμε $\hat{\theta}_n := X_{nn}$.)

(3.86) Ασκηση. Εστω $X_1, \dots, X_n$ α.ι. $\mathcal{U}(0,\theta)$, $\theta > 0$. Βρητε την μοναδικη ΟΑΕΕΔ της θ και δειξτε οτι εχει ομοιομορφα μικροτερο μεσο τετραγωνικο σφαλμα απο την ε.μ.π. της θ.

(3.87) Παραδειγμα. Εστω $X_1, \dots, X_n$ α.ι. Laplace(θ, λ), $\theta \in \mathbb{R}$, $\lambda > 0$, δηλαδη,

$f(x|\theta,\lambda) = \frac{\lambda}{2} \exp\{-\lambda |x-\theta|\}$, $x \in \mathbb{R}$.

Δειξτε οτι η ε.μ.π της θ ειναι η $\hat{\theta}_n = \hat{X}_n \equiv \underset{1 \le i \le n}{\text{med}}\, X_i$ — βλ.(2.8).

Απαντ. Η log-πιθανοφανεια της θ ειναι

$$\ell(\theta) \equiv \ell(\theta|\underline{x}) = -\lambda \sum_{i=1}^{n} |X_i - \theta| + n\log(\lambda/2),$$

της οποιας το σημειο μεγιστου $\hat{\theta}_n$ ειναι το ιδιο με το σημειο ελαχιστου της ακολουθης συναρτησης:

$$\rho(\underline{x},\theta) = \sum_{i=1}^{n} |X_i - \theta| = \sum_{i=1}^{n} |X_{ni} - \theta| .$$

Θα δειξουμε λοιπον οτι

$$\rho(\underline{x},\theta) \geq \rho(\underline{x},\hat{x}_n) = \sum_{i=1}^{[\frac{n}{2}]} \left(X_{n(n+1-i)} - X_{ni} \right) :$$

Εστω $\theta \in [X_{n(i_0-1)}, X_{ni_0}]$, $i_0 = 2,3,\dots,[\frac{n}{2}]$ (Για $i_0 > [\frac{n}{2}]$, οι συλλογισμοι και τα αποτελεσματα ειναι τα ιδια. Επισης ειναι προφανες οτι $\hat{\theta}_n \notin [X_{n1}, X_{nn}]^C$.), τοτε,

$$\rho(\underline{x},\theta) \geq \rho(\underline{x},\hat{x}_n) \iff \sum_{i=i_0}^{n}(X_{ni}-\theta) - \sum_{i=1}^{i_0-1}(X_{ni}-\theta) \geq \rho(\underline{x},\hat{x}_n)$$

$$\iff \sum_{i=i_0}^{n} X_{ni} - \sum_{i=1}^{i_0-1} X_{ni} - [n-2(i_0-1)]\theta \geq \rho(\underline{x},\hat{x}_n)$$

$$\iff \sum_{i=1}^{n-(i_0-1)} X_{n(n+1-i)} - \sum_{i=1}^{i_0-1} X_{ni} - [n-2(i_0-1)]\theta \geq \sum_{i=1}^{[n/2]} X_{n(n+1-i)} - \sum_{i=1}^{[n/2]} X_{ni}$$

$$\iff \sum_{i=[\frac{n}{2}]+1}^{n-(i_0-1)} X_{n(n+1-i)} \geq [n-2(i_0-1)]\theta - \sum_{i=i_0}^{[\frac{n}{2}]} X_{ni}$$

$$\iff \sum_{i=i_0}^{n-[\frac{n}{2}]} X_{ni} + \sum_{i=i_0}^{[\frac{n}{2}]} X_{ni} \geq [n-2(i_0-1)]\theta$$

$$\iff 2\sum_{i=i_0}^{[\frac{n}{2}]} X_{ni} + X_{n([\frac{n}{2}]+1)} \mathbb{1}(n \text{ περιττος}) \geq [n-2(i_0-1)]\theta,$$

το οποιο αληθευει διοτι $X_{ni} \geq \theta \ \forall i \geq i_0$.

Αρα $\forall \lambda > 0$,

$$\max_{\theta\in\mathbb{R}} \ell(\theta) = -\lambda \min_{\theta\in\mathbb{R}} \rho(\underline{x},\theta) + n\log(\lambda/2)$$

$$= -\lambda\, \rho(\underline{x},\hat{x}_n) + n\log(\lambda/2),$$

και αρα $\hat{\theta}_n = \hat{x}_n$.

(3.91) Προταση. Εστω οτι υπαρχει ε.μ.π. $\hat{\theta} \in \Theta$ της $\theta \in \Theta$, δηλ., $L(\hat{\theta}) = \max\{L(\theta) : \theta \in \Theta\}$. Εστω, επισης, συναρτηση $q : \Theta \ni \theta \longmapsto q(\theta) \in \eta \equiv q(\Theta)$. Τοτε, υπαρχει ε.μ.π. $\widehat{q(\theta)} \in \eta$ της $q(\theta) \in \eta$ και $\widehat{q(\theta)} = q(\hat{\theta})$.

Αποδ. Αν η q ειναι 1:1, τοτε, η πιθανοφανεια της $\eta (= q(\theta))$ οριζεται φυσιολογικα ως $L^*(\eta) := L(q^{-1}(\eta))$, $\eta \in \eta$, και εχουμε οτι:

$L^*(q(\hat{\theta})) = L(\hat{\theta}) = \max\{L(\theta) : \theta \in \Theta\} = \max\{L^*(q(\theta)) : \theta \in \Theta\} =$
$\quad = \max\{L^*(\eta) : \eta \in \eta\}$.

Αν, ομως, η συναρτηση q δεν ειναι απαραιτητα αμφιμονοσημαντη, τοτε, εχουμε καποιο προβλημα στον ορισμο της πιθανοφανειας $L^*(\eta)$ της $\eta \in \eta$. Ο πιο φυσιολογικος – και απαιτητικος – ορισμος της πιθανοφανειας της η ειναι ο εξης: $\quad L^*(\eta) := \sup\{L(\vartheta) : \vartheta \in q^{-1}(\{\eta\})\}$, $\eta \in \eta$.

Προφανως, $L^*(\eta) \leq L(\hat{\theta}) \quad \forall\, \eta \in \eta$,

και αρα $\quad L^*(q(\hat{\theta})) \leq L(\hat{\theta})$.

Αλλα, $L^*(q(\hat{\theta})) \equiv \sup\{L(\vartheta) : \vartheta \in q^{-1}(\{q(\hat{\theta})\})\} \geq L(\hat{\theta})$, διοτι, $\hat{\theta} \in q^{-1}(\{q(\hat{\theta})\})$,

και αρα $\quad L^*(q(\hat{\theta})) = L(\hat{\theta})$. Εχουμε, λοιπον, οτι:

$$L^*(\eta) \leq L(\hat{\theta}) = L^*(q(\hat{\theta})) \quad \forall\, \eta \in \eta,$$

και αρα,

$$\sup\{L^*(\eta) : \eta \in \eta\} \leq L^*(q(\hat{\theta})).$$

Αλλα, $q(\hat{\theta}) \in \eta$, και αρα:

$$\exists \max\{L^*(\eta) : \eta \in \eta\} = L^*(q(\hat{\theta})) \quad (= L(\hat{\theta})),$$

δηλαδη, $\widehat{q(\theta)} = q(\hat{\theta})$.

Το αποτελεσμα αυτο ειναι εξαιρετικα χρησιμο στις εφαρμογες, και καθιστα την μεθοδο μεγιστης πιθανοφανειας πρακτικοτατη. Για παραδειγμα, αν μας ενδιαφερη η εκτιμηση της παραμετρου λ της $\mathcal{G}(\alpha_0, \lambda)$ ή της παραμετρου θ της $\mathcal{G}(\alpha_0, \lambda = \frac{1}{\theta})$, οι ε.μ.π. των λ & θ συνδεονται με τη σχεση $\hat{\lambda}_n = \frac{1}{\hat{\theta}_n}$, και αρα αρκει να βρουμε την μια απο τις δυο, π.χ., $\hat{\theta}_n = \bar{X}_n/\alpha_0$. Επισης, αν μας ενδιαφερη, π.χ., η εκτιμηση της τυπικης αποκλισης της $N(\mu,\sigma^2)$ ή της Bernoulli(p) με ε.μ.π., αρκει να βρουμε τις εμπ. των σ^2 και p αντιστοιχως, ετσι:

$$\hat{\sigma}_n = (\hat{\sigma}_n^2)^{1/2} = \sqrt{\frac{1}{n}\sum_{i=1}^n (X_i-\bar{X}_n)^2}\ , \quad \text{και},$$

$$\widehat{\sqrt{p(1-p)}} = \sqrt{\hat{p}_n(1-\hat{p}_n)} = \sqrt{\bar{X}_n(1-\bar{X}_n)}\ ,$$

ειναι οι ε.μ.π των σ και $\sqrt{p(1-p)}$ αντιστοιχως.

Η χρησιμοτητα, ομως, της μεθοδου φαινεται κυριως στην εκτιμηση πιθανοτητων: αν υπαρχη $\hat{\theta}$ ε.μ.π. της θ, τοτε, υπαρχει ε.μ.π. της $P_\theta(X\in A)$ και $\widehat{P_\theta(X\in A_\theta)} = P_{\hat{\theta}}(X\in A_{\hat{\theta}})$.

<u>Παραδειγμα:</u> Εστω $X_1,\dots,X_n$ α.ι. $f(x\mid\theta,\tau) = \frac{1}{2\tau}\exp\{-|\frac{x-\theta}{\tau}|\}, x\in\mathbb{R}$. Απ' το παραδειγμα (3.87), εχουμε οτι οι ε.μ.π. των θ,τ ειναι οι $\hat{\theta}_n = \hat{X}_n$, $\hat{\tau}_n = \frac{1}{n}\sum_{i=1}^n |X_i-\hat{X}_n|$. Αρα, υπαρχει, ε.μ.π. της

$$P_{(\theta,\tau)}(X\le 2\theta+\tau) =$$
$$= \tfrac{1}{2}\exp\{\tfrac{\theta}{\tau}+1\}\,\mathbb{1}(\theta\le-\tau) + \{1-\tfrac{1}{2}\exp\{-(\tfrac{\theta}{\tau}+1)\}\}\mathbb{1}(\theta>-\tau)$$

και ειναι η: $P_{(\hat{\theta}_n,\hat{\tau}_n)}(X\le 2\hat{\theta}_n+\hat{\tau}_n)$.

Η δυναμη της μεθοδου, ενισχυεται ακομη περισσοτερο οταν συνδιασθη με τα Θεωρηματα 1.94 και 1.95. Αν, δηλαδη, η $\hat{\theta}_n$ ειναι συνεπης εκτιμητρια της θ, τοτε και η $q(\hat{\theta}_n)$ ειναι συνεπης εκτιμητρια της $q(\theta)$, εφοσον η q ειναι συνεχης, και, επισης, αν $\sqrt{n}(\hat{\theta}_n-\theta) \xrightarrow{d} N(0,\sigma^2(\theta))$, τοτε και

$$\sqrt{n}(q(\hat{\theta}_n)-q(\theta)) \xrightarrow{d} N(0,[q'(\theta)]^2\sigma^2(\theta)),$$ εφοσον η q παραγωγιζεται.

Στα αρνητικα των ε.μ.π., συμπεριλαμβανεται το οτι η ιδιοτητα της αμεροληψιας της $\hat{\theta}_n$ (αν υφισταται) δεν διατηρειται, εν γενει, για την $q(\hat{\theta}_n)$. Διατηρειται, ομως, εν γενει, η ιδιοτητα της ασυμπτωτικης αμεροληψιας. Για το οτι οι ε.μ.π. ειναι, εν γενει, συνεπεις και ασυμπτωτικα κανονικες (και μαλιστα με την ελαχιστη δυνατη διασπορα), βλ. σελ. 108, 109.

(3.88) <u>Παρατηρηση</u>: Εστω $X_1, \ldots, X_n$ α.ι. $f(x|\eta) = \exp\{\eta T(x) + d_0(\eta) + S(x)\}$, $x \in A$, απο την μονοπαραμετρικη εκθετικη οικογενεια, δηλαδη, με τη φυσικη της παραμετριση – βλ. (3.18) και (3.19). Θα δειξουμε οτι, σ' αυτη την περιπτωση, οι εξισωσεις μεγιστης πιθανοφανειας (3.81) παιρνουν την μορφη:

(3.89) $$E_\eta\{T(X_1)\} = \frac{1}{n}\sum_{i=1}^n T(X_i), \quad \eta \in \mathring{H},$$

και διδουν την μοναδικη εμπ της $\eta \in H \equiv c(\Theta)$. (<u>Υποθ.</u> A ανεξαρτητο της η.)

Στην (3.18) ειδαμε οτι η κατανομη του δειγματος $\underline{X}$ ανηκει επισης σε μια μονοπαραμετρικη εκθετικη οικογενεια, και συγκεκριμμενα

$$f(\underline{x}|\eta) = \exp\{\eta \Sigma_{i=1}^n T(x_i) + n d_0(\eta) + \Sigma_{i=1}^n S(x_i)\} \, 1(\underline{x} \in A^n), \quad \eta \in H.$$

Αρα, η log-πιθανοφανεια της η ειναι η εξης:

$$\ell(\eta) = \ell(\eta|\underline{X}) = \eta \Sigma_{i=1}^n T(X_i) + n d_0(\eta) + \Sigma_1^n S(X_i), \quad \eta \in H, \ \underline{X} \in A^n.$$

Τοτε, απο την (3.22),

$$\ell'(\eta) = \Sigma_{i=1}^n T(X_i) + n d_0'(\eta) = \Sigma_{i=1}^n T(X_i) - n E_\eta\{T(X_1)\}, \quad \eta \in \mathring{H}$$

και αρα η (3.81) ειναι ισοδυναμη με την (3.89).

Το οτι η (3.89) διδει την μοναδικη εμπ της η επεται ειτου οτι:

$$\ell''(\eta) = n d_0''(\eta) = -n \mathcal{D}_\eta(T(X_1)) < 0,$$

βασει της (3.23). (Σημ. Η διαφορισιμοτητα της $d_0(\cdot)$ επεται ειτου ορισμου της – βλ. (3.19).)

(3.90) <u>Παραδειγμα</u>. Εστω $X_1, \ldots, X_n$ α.ι. Poisson(θ), $\theta \in \Theta = (0, +\infty)$.

Δειξτε οτι: (α) η εμπ της $\eta = \log\theta$ ειναι η $\hat{\eta} = \log \overline{X}_n$.

(β) η εμπ της θ ειναι η $\hat{\theta} = \overline{X}_n$ και αρα,

(γ) $\hat{\eta} = \widehat{\log\theta} = \log\hat{\theta}$.

<u>Απαντηση</u>. (α) Η $\mathcal{P}(\theta)$ ανηκει στην μονοπαραμετρικη εκθετικη οικογενεια:

$$p(x|\theta) = \frac{e^{-\theta}\theta^x}{x!} 1(x \in \mathbb{N}_0) = \exp\{(\log\theta)x - \theta - \log(x!)\} 1(x \in \mathbb{N}_0),$$

δηλαδη,

$$p(x|\eta) = \exp\{\eta x + (-e^\eta) + [-\log(x!)]\} 1(x \in \mathbb{N}_0).$$

Αρα, απο την (3.88), η εμπ της η διδεται απο την (3.89), δηλαδη, την

$$E_\eta(X) = \overline{X}_n \Leftrightarrow \theta = \overline{X}_n \Leftrightarrow e^\eta = \overline{X}_n \Leftrightarrow \eta = \log \overline{X}_n.$$

(β) Η log-πιθανοφάνεια της θ είναι η $\ell(\theta) = -n\theta + n\bar{X}_n \log\theta + \Sigma_1^n \log(X_i!)$
$\Rightarrow \ell'(\theta) = -n + \frac{n\bar{X}_n}{\theta} = 0 \Leftrightarrow \theta = \bar{X}_n$,
και $\ell''(\theta) = -\frac{n\bar{X}_n}{\theta^2} < 0$. Άρα η εμπ της θ είναι η $\hat{\theta} = \bar{X}_n$.
(γ) Έχουμε λοιπόν ότι, $\widehat{q(\theta)} = q(\hat{\theta})$, με $q(\theta) = \log\theta = \eta$.

Αυτή τελευταία σχέση του (3.90γ) είναι μια ιδιότητα των εμπ η οποία ισχύει πολύ γενικότερα — βλ. (3.91) — και αυξάνει την δημοτικότητα των εμπ. Άλλες εκτιμήτριες, π.χ., οι ΟΑΕΕΔ δεν έχουν αυτή την ιδιότητα.

(3.91) <u>Πρόταση</u>. Έστω $\underline{X}$ στοχαστικό δείγμα από την $F(\cdot|\theta)$, $\theta \in \Theta$, και έστω ότι υπάρχει εμπ $\hat{\theta}$ για την θ. Έστω δε αμφιμονοσήμαντη συνάρτηση $q: \Theta \longrightarrow Q \equiv q(\Theta)$. Η εμπ της $\eta = q(\theta)$ υπάρχει και $\hat{\eta} \equiv \widehat{q(\theta)} = q(\hat{\theta})$.

<u>Αποδ.</u> Έστω $L(\theta) \equiv L(\theta|\underline{X})$ η πιθανοφάνεια της $\theta \in \Theta$. Τότε η πιθανοφάνεια της $\eta = q(\theta)$ είναι $L^*(\eta) = L(q^{-1}(\eta))$, $\eta \in Q$, και άρα

$$\exists \max_{\eta\in Q} L^*(\eta) = \max_{\eta\in Q} L(q^{-1}(\eta)) = \max_{\theta\in\Theta} L(\theta),$$

και επίσης έχουμε ότι:

$$L^*(q(\hat{\theta})) = L(q^{-1}(q(\hat{\theta}))) = L(\hat{\theta}) = \max_{\theta\in\Theta} L(\theta) = \max_{\eta\in Q} L^*(\eta),$$

δηλαδή $q(\hat{\theta}) = \hat{\eta} \equiv \widehat{q(\theta)}$.

(3.92) <u>Παράδειγμα</u>. Έστω $X_1, \ldots, X_n$ α.ι. παρατηρήσεις της διάρκειας ζωής ενός τύπου λυχνιών, των οποίων η διάρκεια ζωής, σε μήνες, X, ακολουθεί την Εκθετική(λ), $\lambda > 0$. Βρείτε την εμπ της πιθανότητος

$q(\lambda) := P_\lambda(X > 5) = e^{-5\lambda}$.

<u>Απάντηση</u>. Από την (3.91), $\widehat{q(\lambda)} = q(\hat{\lambda})$ και άρα αρκεί να υπολογίσουμε την εμπ $\hat{\lambda}$ της λ. Έχουμε λοιπόν,

$\ell(\lambda) = \ell(\lambda|\underline{X}) = n\log\lambda - n\lambda\bar{X}_n \Rightarrow 0 = \ell'(\lambda) = \frac{n}{\lambda} - n\bar{X}_n \Leftrightarrow$

$\Leftrightarrow \lambda = \frac{1}{\bar{X}_n}$, και επειδη $l''(\lambda) = -n/\lambda^2 < 0$, $\hat{\lambda} = \frac{1}{\bar{X}_n}$ και αρα $\widehat{q(\lambda)} = q(\hat{\lambda}) = \exp\{-5/\bar{X}_n\}$.

(3.93) <u>Παραδειγμα</u>: Εστω $X_1, \dots, X_n$ α.ι. παρατηρησεις των εισοδηματων των κατοικων μιας περιοχης, τα οποια υποθετουμε οτι ακολουθουν καποια $N(\mu, \sigma^2)$, $\underline{\theta} = (\mu, \sigma^2) \in \Theta \equiv \mathbb{R} \times (0, +\infty)$ κατανομη. Θελουμε να εκτιμησουμε την πιθανοτητα $q(\underline{\theta}) = P_{\underline{\theta}}(X > x_0)$, για καποιο συγκεκριμενο εισοδηματικο επιπεδο $x_0 \in \mathbb{R}$.

<u>Απαντ.</u> Απο την (3.91), η εμπ της $q(\underline{\theta}) = P_{\underline{\theta}}(X > x_0) =$
$= P_{\underline{\theta}}\left(\frac{X-\mu}{\sigma} > \frac{x_0 - \mu}{\sigma}\right) = 1 - \Phi\left(\frac{x_0 - \mu}{\sigma}\right)$ ειναι η
$\widehat{q(\underline{\theta})} = q(\hat{\underline{\theta}}) = q(\hat{\mu}, \hat{\sigma}^2) = q(\bar{X}_n, \hat{\sigma}_n^2) = 1 - \Phi\left(\frac{x_0 - \bar{X}_n}{\hat{\sigma}_n}\right)$,

εφοσον, απο το (3.83γ), $\hat{\mu} = \bar{X}_n$, $\hat{\sigma}^2 = \hat{\sigma}_n^2$ και αρα απο την (3.91) παλι $\hat{\sigma} = \sqrt{\hat{\sigma}_n^2} = \sqrt{\frac{1}{n}\sum_1^n (X_i - \bar{X}_n)^2} =: \hat{\sigma}_n$.

(3.94) <u>Παρατηρηση</u>: Εστω $X_1, \dots, X_n$ α.ι. $f(x|\theta) = \exp\{c(\theta)T(x) + d(\theta) + S(x)\}$, $x \in A$ −ανεξαρτητο της θ− και $\theta \in \Theta$, εχουμε, δηλαδη, δειγμα απο την μονοπαραμετρικη εκθετικη οικογενεια του (3.18). Θα δειξουμε οτι αν η $c(\cdot)$ ειναι αμφιμονοσημαντη, οι εξισωσεις μεγιστης πιθανοφανειας (3.81) παιρνουν την μορφη:

(3.95) $E_\theta\{T(X_1)\} = \frac{1}{n}\sum_{i=1}^n T(X_i)$, $\theta \in \Theta : c(\theta) \in \mathring{H}$,

και διδουν την μοναδικη ε.μ.π. της $\theta \in \Theta$.

Απο την (3.88) η μοναδικη εμπ $\hat{\eta}$ της $\eta = c(\theta)$ διδεται απο την (3.89), δηλαδη, την $E_{\hat{\eta}}\{T(X_1)\} = \frac{1}{n}\sum_1^n T(X_i)$
$\Leftrightarrow E_{\widehat{c(\theta)}}\{T(X_1)\} = \frac{1}{n}\sum_{i=1}^n T(X_i) \Leftrightarrow E_{c(\hat{\theta})}\{T(X_1)\} = \frac{1}{n}\sum_1^n T(X_i)$,
δηλαδη, την $E_{\hat{\theta}}\{T(X_1)\} = \frac{1}{n}\sum_1^n T(X_1)$.

Η μοναδικοτητα της $\hat{\theta}$ επεται απο την μοναδικοτητα της $\hat{\eta}$ και το οτι, απο το αμφιμονοσημαντο της $c(\cdot)$ και την (3.91),
$\hat{\theta} = \widehat{c^{-1}(\eta)} = c^{-1}(\hat{\eta})$.

(3.96) <u>Παράδειγμα</u>. Εστω $X_1,\dots,X_n$ α.ι. $\mathcal{E}(\lambda=\frac{1}{\theta})$, $\theta\in\Theta=(0,+\infty)$. Δείξτε οτι η εμπ της θ ειναι η $\hat{\theta}=\bar{X}_n$.

<u>Απαντηση</u>. Η εμπ της θ μπορει φυσικα να ευρεθη με χρηση των (3.81) και (3.82), απ' ευθειας, οπως στο (3.92). Επειδη ομως η $\mathcal{E}(\lambda)$ ανηκει στη εκθετικη οικογενεια:

$$f(x|\theta)=\exp\{(-\tfrac{1}{\theta})x+(-\log\theta)\}\,1(x>0),$$

εχουμε την εναλλακτικη λυση της (3.95). Δηλαδη, η μοναδικη εμπ $\hat{\theta}$ της θ ειναι η λυση της εξισωσης:

$$E_\theta\{T(X_1)\}=\tfrac{1}{n}\Sigma_1^n T(X_i), \quad \text{δηλαδη, της εξισωσης,}$$

$$E_\theta(X_1)=\bar{X}_n \Leftrightarrow (1/\theta)^{-1}=\bar{X}_n\Leftrightarrow \theta=\bar{X}_n,$$

και αρα $\hat{\theta}=\bar{X}_n$.

(3.97) <u>Παρατηρηση</u>. Εστω στοχαστικο δειγμα $\underline{X}$ απο την $F(\cdot|\theta)$, $\theta\in\Theta$, και συναρτηση $q:\Theta\to Q\equiv q(\Theta)$. Στην (3.91) δειξαμε οτι αν η εμπ $\hat{\theta}$ της θ υπαρχει και η q ειναι αμφιμονοσημαντη, τοτε

(3.98) $\widehat{q(\theta)}=q(\hat{\theta})$.

Η (3.98) ισχυει ακομη και αν η q δεν ειναι αμφιμονοσημαντη.

Αν δουμε προσεκτικα την αποδειξη της (3.91), παρατηρουμε οτι το αμφιμονοσημαντο της $q(\cdot)$ μας χρειασθηκε μονο για την εκφραση της πιθανοφανειας $L(\theta)$ μεσω της $\eta=q(\theta)$, αλλα, δηλαδη, $L^*(\eta)=L(q^{-1}(\eta))$.

Ομως, εστω οτι η $g(\theta)=(q(\theta),q_1(\theta),\dots,q_k(\theta))\in G=g(\Theta)$, ειναι αμφιμονοσημαντη, για καποιες (συμπληρωματικες) συναρτησεις $q_1,\dots,q_k$, τοτε απο την (3.91), $\widehat{g(\theta)}=g(\hat{\theta})=(q(\hat{\theta}),q_1(\hat{\theta}),\dots,q_k(\hat{\theta}))$ και αρα $\widehat{q(\theta)}=q(\hat{\theta})$.

(3.99) <u>Παράδειγμα</u>. Εστω $X_1,\dots,X_n$ α.ι. $N(\mu,\sigma^2)$, $\mu\in\mathbb{R}$; $\sigma^2>0$. Απο την (3.83a) $\hat{\mu}=\bar{X}_n$ και αρα απο την (3.97) η εμπ της $q(\mu)=\mu^2$ ειναι η $\hat{q}(\mu)=\widehat{(\mu^2)}=(\hat{\mu})^2=\bar{X}_n^2$.

(3.100) <u>Παραδειγμα</u>. Εστω $X_1, \dots, X_n$ Bernoulli(p), $p \in \Theta = [0,1]$. Θελουμε να εκτιμησουμε τη διασπορα $q(p) = p(1-p)$ της κατανομης αυτης.

Απο την (3.97) η εμπ της $q(p)$ ειναι η $\widehat{q(p)} = q(\hat{p}) = \hat{p}(1-\hat{p})$ αρκει λοιπον να βρουμε την εμπ της p, η οποια απο την (3.101ε) ειναι $\hat{p} = \bar{X}_n$, αρα $\widehat{q(p)} = \bar{X}_n(1-\bar{X}_n)$ — βλ. (3.49) για σύγκριση με την ΟΑΕΕΔ της $q(p)$.

(3.101) <u>Ασκηση</u>. Εστω $X_1, \dots, X_n$ α.ι. $F(\cdot|\theta)$, $\theta \in \Theta$. Στη περιπτωση που η $F(\cdot|\theta)$ ειναι η σ.κ. της:

(α) $\mathcal{U}_\Delta(\theta)$, βρητε την ε.μ.π. της $\theta \in \Theta = \mathbb{N}$

(β) $\mathcal{U}(\theta, 1)$, βρητε την εμπ της $\theta \in \Theta = (-\infty, 1)$.

(γ) $\mathcal{U}(\theta_1, \theta_2)$, βρητε τις εμπ των $-\infty < \theta_1 < \theta_2 < +\infty$.

(δ) Εκθετικης (λ, θ) με πυκνοτητα $\lambda \exp\{-\lambda(x-\theta)\} 1(x>\theta)$, $(\lambda,\theta) \in \Theta = (0,+\infty) \times \mathbb{R}$, βρητε τις εμπ των λ, θ.

(ε) $\mathcal{D}(n,p)$ με n γνωστο, βρητε την εμπ της $p \in \Theta = [0,1]$.

(ζ) $\mathcal{AD}(\kappa,p)$ με κ γνωστο, βρητε την εμπ της $p \in \Theta = [0,1]$.

(η) $\mathcal{G}(\alpha, \frac{1}{\theta})$ με α γνωστο, βρητε την εμπ της $\theta \in \Theta = (0,+\infty)$.

(θ) $\mathcal{W}(\alpha, \lambda)$ με α γνωστο, βρητε την εμπ της $\lambda \in \Theta = (0,+\infty)$.

(3.102) <u>Ασκηση</u>. Εστω $X_1, \dots, X_n$ α.ι. Bernoulli(p), $p \in \Theta = [0,1]$. Μας ενδιαφερει η εκτιμηση του $q(p) = \frac{p}{1-p}$.

(α) Δειξτε οτι δεν υπαρχει αμεροληπτη εκτιμητρια του $q(p)$.

(β) Βρητε την ε.μ.π. της $q(p)$.

(3.103) <u>Ασκηση</u>. Εστω $X_1, \dots, X_n$ α.ι. $\mathcal{W}(\alpha, \lambda)$, α γνωστο, $\lambda > 0$. Εστω $h(x|\lambda) = -\frac{d}{dx} \log[1 - F(x|\lambda)]$, $x > 0$ η συναρτηση του ρυθμου κινδυνου της κατανομης. Βρητε την ε.μ.π. της $h(x_0|\lambda)$, για καθοιο σταθερο $x_0 > 0$.

(3.104) <u>Ασκηση.</u> Εστω $X_1,\dots,X_n$ α.ι. $N(\mu,\sigma^2)$, $\mu\in\mathbb{R}$, $\sigma^2>0$.
Βρητε την ε.μ.π. της $q(\mu,\sigma^2) = P_{(\mu,\sigma^2)}(-2 < X < 1)$.

(3.105) <u>Ασκηση.</u> Βρητε την ε.μ.π. της διασπορας:
(α) της $\mathcal{U}_\Delta(\theta)$, $\theta\in\mathbb{N}$.
(β) της $\mathcal{AD}(\kappa,p)$, με κ γνωστο, $p\in[0,1]$.
(γ) της $\mathcal{P}(\lambda)$, $\lambda>0$.
(δ) της $\mathcal{U}(\theta_1,\theta_2)$, $-\infty<\theta_1<\theta_2<+\infty$
(ε) της $\mathcal{G}(\alpha,\lambda)$, α γνωστο, $\lambda>0$.
(ζ) της $W(\alpha,\lambda)$, α γνωστο, $\lambda>0$.

(3.106) <u>Ασκηση.</u> Εστω $X_1,\dots,X_n$ α.ι. $\mathcal{G}(\alpha,\lambda)$, α γνωστο, $\lambda>0$, $n\alpha>2$.
(α) Βρητε την μοναδικη εμπ $\hat{\lambda}$ της λ.
(β) Βρητε την μοναδικη ΟΑΕΕΔ $\hat{\hat{\lambda}}$ της λ.
(γ) Συγκρινατε τις $\hat{\lambda}$, $\hat{\hat{\lambda}}$ ως προς το μεσο τετραγωνικο σφαλμα τους.

(3.107) <u>Ασκηση.</u> Εστω $X_1,\dots,X_n$ α.ι. Poisson (θ), $\theta>0$.
Στην (3.47) ειδαμε οτι η μοναδικη ΟΑΕΕΔ της $q(\theta)=e^{-\theta}$
ειναι η $\hat{\hat{q}}(\theta) = \left(1-\frac{1}{n}\right)^{n\bar{X}_n}$.
(α) Δειξτε οτι η μοναδικη εμπ της $q(\theta)$ ειναι η $\hat{q}(\theta)=e^{-\bar{X}_n}$.
(β) Συγκρινατε τις $\hat{q}$, $\hat{\hat{q}}$ ως προς το μεσο τετραγωνικο σφαλμα τους. (Ξεχαστε τους $O(n^{-2})$ ορους).

(3.108) <u>Ασκηση.</u> Εστω $X_1,\dots,X_n$ α.ι. Bernoulli(p), $p\in[0,1]$.
Ειδαμε οτι η εμπ της $q(p)=p(1-p)$ ειναι
η $\hat{q}(p)=\bar{X}_n(1-\bar{X}_n)$ και η ΟΑΕΕΔ της $q(p)$ ειναι
η $\hat{\hat{q}}(p)=\frac{n}{n-1}\bar{X}_n(1-\bar{X}_n)$.
Συγκρινατε τις $\hat{q}$, $\hat{\hat{q}}$ ως προς το μεσο τετραγωνικο σφαλμα τους.

Απο το Θεωρημα (3.41) των Rao-Blackwell, βλεπουμε οτι αν υπαρχει επαρκης στατιστικη συναρτηση $T(\underline{X})$ για την παραμετρο $\theta \in \Theta$, τοτε η ΟΑΕΕΔ της θ ειναι συναρτηση της $T(\underline{X})$ μονο - διοτι αν δεν ηταν θα μπορουσαμε να την βελτιωσαμε μεσω της (3.42). Αυτη την επιθυμητη ιδιοτητα των ΟΑΕΕΔ συμμεριζονται και οι ε.μ-π., διοτι αν υπαρχη επαρκης T για την θ τοτε απο κριτηριο επαρκειας - βλ. (3.14) - εχουμε οτι:

$$L(\theta) = L(\theta|\underline{x}) = f(\underline{x}|\theta) = g(T(\underline{x}),\theta)\, h(\underline{x}),$$

οπου η $h(\cdot)$ ειναι ανεξαρτητη της θ και η $g(\cdot,\cdot)$ εξαρταται απο το δειγμα μονο μεσω της T. Εχουμε λοιπον:

$$\sup_{\theta\in\Theta} L(\theta) = h(\underline{x}) \sup_{\theta\in\Theta} g(T(\underline{x}),\theta),$$

και αρα το σημειο μεγιστου της $L(\theta)$, αν υπαρχη, ταυτιζεται με το σημειο μεγιστου της $g(T(\underline{x}),\theta)$ η οποια εξαρταται απο το δειγμα μονο μεσω της T. Αρα η εμπ $\hat{\theta}$, αν υπαρχη, ειναι συναρτηση μονο της T. Εχουμε αποδειξει λοιπον οτι:

(3.109) <u>Προταση</u>: Εστω στοχαστικο δειγμα $\underline{X}$ απο την F_θ, $\theta\in\Theta$ και εστω οτι υπαρχει επαρκης για την θ στατιστικη συναρτηση $T(\underline{X})$. Τοτε η <u>ΟΑΕΕΔ</u> και η <u>εμπ</u> της θ, αν υπαρχουν, ειναι συναρτησεις του δειγματος μονο μεσω της $T(\underline{X})$.

(3.110) <u>Παρατηρηση</u>. Εστω $X_1,\dots,X_n$ a.i. F_θ, $\theta\in\Theta$ και εστω οτι υπαρχει εμπ $\hat{\theta}_n$ της θ. Συχνα, οι εξισωσεις μεγιστης πιθανοφανειας (3.81) δεν μπορουν να λυθουν ακριβως. Σ' αυτη την περιπτωση, για τον υπολογισμο - ή μαλλον την προσεγγιση - της $\hat{\theta}_n$, καταφευγουμε σε αλγοριθμικες μεθοδους της αριθμητικης αναλυσης. Συνηθως χρησιμοποιειται (και αρκει) το <u>πρωτο βημα της μεθοδου Newton-Raphson</u>, συγκεκριμμενα προσεγγιζουμε την εμπ $\underline{\hat{\theta}}_n$ με την $\underline{\hat{\hat{\theta}}}_n$, η οποια οριζεται ως εξης:

(3.111) $$\hat{\hat{\underline{\theta}}}_n := \underline{\theta}_n^0 - \left[H(\underline{\theta}_n^0)\right]^{-1} \nabla \ell_n(\underline{\theta}_n^0) ,$$

οπου η $\underline{\theta}_n^0$ ειναι καθοια <u>προκαταρκτικη</u> <u>εκτιμητρια</u> της $\underline{\theta}$, π.χ., μια εκτιμητρια της μεθοδου ροπων, η οποια να ειναι <u>συνεπης</u> – βλ.(3.   ).

Στη μια διασταση η (3.111) παιρνει την μορφη,

(3.112) $$\hat{\hat{\theta}}_n := \theta_n^0 - \frac{\ell_n'(\theta_n^0)}{\ell_n''(\theta_n^0)}$$

$$= \theta_n^0 - \left\{ \frac{\sum_{i=1}^n \frac{\partial}{\partial\theta} \log f(X_i|\theta)}{\sum_{i=1}^n \frac{\partial^2}{\partial\theta^2} \log f(X_i|\theta)} \right\}_{\theta=\theta_n^0} .$$

Η χρηση της (3.112) – και ακριβως αναλογα της (3.111) – μπορει να δικαιολογηθη μεσω του Θεωρηματος της Μεσης Τιμης, ως εξης:

Εστω οτι η log-πιθανοφανεια ειναι δυο φορες παραγωγισιμη (σε μια περιοχη της εμπ), τοτε απ' το θεωρημα της μεσης τιμης,

$$\ell_n'(\hat{\theta}_n) - \ell_n'(\theta_n^0) = \ell_n''(\bar{\theta}_n^0)\,(\hat{\theta}_n - \theta_n^0) ,$$

για καποιο $\bar{\theta}_n^0$ μεταξυ των θ_n^0 και $\hat{\theta}_n$. Αλλα, απο την (3.81), $\ell'(\hat{\theta}_n) = 0$ και αρα,

$$\hat{\theta}_n = \theta_n^0 - \frac{\ell_n'(\theta_n^0)}{\ell_n''(\bar{\theta}_n^0)} ,$$

αντικαθιστωντας λοιπον σ' αυτον τον τυπο την αγνωστη $\bar{\theta}_n^0$ με την θ_n^0 παιρνουμε την $\hat{\hat{\theta}}_n$ της (3.112), ως μια προσεγγιση της εμπ $\hat{\theta}_n$. Μαλιστα,

(3.113) $$|\hat{\theta}_n - \hat{\hat{\theta}}_n| = \frac{|\ell_n'(\theta_n^0)|}{|\ell_n''(\theta_n^0)|\,|\ell_n''(\bar{\theta}_n^0)|}\,|\ell_n''(\theta_n^0) - \ell_n''(\bar{\theta}_n^0)| ,$$

με $\theta_n^0 \wedge \hat{\theta}_n < \bar{\theta}_n^0 < \theta_n^0 \vee \hat{\theta}_n$. Εφοσον λοιπον η θ_n^0 εχει επιλεγη συνεπης και οι εμπ ειναι εν γενει συνεπεις – βλ. (3.   ) – οι $\theta_n^0, \hat{\theta}_n, \bar{\theta}_n^0$ ειναι "κοντα" στην αληθη τιμη της θ και αρα και μεταξυ τους. Αρα, $\ell_n'(\theta_n^0) \approx \ell_n'(\hat{\theta}_n) = 0$ και επισης $\ell_n''(\theta_n^0) - \ell_n''(\bar{\theta}_n^0) \approx 0$ και αρα οι

$\hat{\theta}_n$, $\hat{\hat{\theta}}_n$ πρεπει να ειναι "πολυ κοντα" η μια στην αλλη. Μπορουμε, λοιπον, να περιμενουμε οτι η προσεγγιση Newton-Raphson $\hat{\hat{\theta}}_n$ της εμπ $\hat{\theta}_n$ ειναι "πολυ καλη".

(3.114) <u>Παραδειγμα</u>. Εστω $X_1, \dots, X_n$ α.ι. Cauchy$(\theta, 1)$, $\theta \in \Theta \equiv \mathbb{R}$.
Η log-πιθανοφανεια της θ ειναι:

$$l_n(\theta) = -\Sigma_{i=1}^{n} \log[1+(X_i-\theta)^2] - n\log\pi \quad \Rightarrow$$

$$\Rightarrow l_n'(\theta) = 2\Sigma_{i=1}^{n} \frac{X_i-\theta}{1+(X_i-\theta)^2}, \quad l_n''(\theta) = -2\Sigma_{i=1}^{n} \frac{1-(X_i-\theta)^2}{[1+(X_i-\theta)^2]^2}.$$

Η εξισωση μεγιστης πιθανοφανειας $l'(\theta)=0$ δεν λυνεται ακριβως για $n \geq 3$. Προσεγγιζουμε λοιπον την εμπ $\hat{\theta}_n$ της θ με την (3.112), δηλαδη, την

$$\hat{\hat{\theta}}_n := \theta_n^0 - \frac{l_n'(\theta_n^0)}{l_n''(\theta_n^0)},$$

με $\theta_n^0 := \hat{X}_n$, την μεση του δειγματος, η οποια ειναι εν γενει μια συνεπης εκτιμητρια της μεσης θ της κατανομης.
Παρατηρουμε οτι στην περιπτωση της Cauchy δεν υπαρχουν εκτιμητριες της μεθοδου των ροπων, διοτι δεν υπαρχουν οι ροπες $E|X|^k$, για $k \geq 1$, της κατανομης αυτης.

(3.115) <u>Ασκηση</u>. Εστω $X_1, \dots, X_n$ α.ι. Cauchy(α, β), $\alpha \in \mathbb{R}$, $\beta > 0$.
(α) Υποθεσατε οτι $\alpha = 0$ και βρητε την N-R προσεγγιση $\hat{\hat{\beta}}_n$ της $\hat{\beta}_n$.
(β) Βρητε την N-R προσεγγιση $(\hat{\hat{\alpha}}_n, \hat{\hat{\beta}}_n)$ της $(\hat{\alpha}_n, \hat{\beta}_n)$.
(<u>Υποδ.</u> Χρησιμοποιειστε: $\alpha_n^0 := \hat{X}_n$, $\beta_n^0 := \underset{1 \leq i \leq n}{\text{med}} |X_i - \hat{X}_n|$.)

(3.116) <u>Ασκηση</u>. Εστω $X_1, \dots, X_n$ α.ι. Logistic(α, β), $\alpha \in \mathbb{R}$, $\beta > 0$, δηλαδη, $f(x|\alpha,\beta) = \{\beta[1+\exp\{-(x-\alpha)/\beta\}]\}^{-2} \exp\{-(x-\alpha)/\beta\}$, $x \in \mathbb{R}$.
(α) Υποθεσατε οτι $\beta = 1$ και βρητε την N-R προσεγγιση $\hat{\hat{\alpha}}_n$ της $\hat{\alpha}_n$.
(β) Υποθεσατε οτι $\alpha = 0$ και βρητε την N-R προσεγγιση $\hat{\hat{\beta}}_n$ της $\hat{\beta}_n$.
(γ) Βρητε την N-R προσεγγιση $(\hat{\hat{\alpha}}_n, \hat{\hat{\beta}}_n)$ της $(\hat{\alpha}_n, \hat{\beta}_n)$.

3.4. ΚΡΙΤΗΡΙΑ ΑΠΟΔΟΣΗΣ ΕΚΤΙΜΗΤΡΙΩΝ

Εστω $\underline{X}_n \equiv (X_1, \ldots, X_n)$ α.ι. $F(\cdot|\theta)$, $\theta \in \Theta$ και εστω $d_n(\underline{X}_n)$ μια εκτιμητρια της $q(\theta)$. Αυτο το οποιο μπορουμε να περιμενουμε — και το απαιτουμε — απο την τυχαια μεταβλητη $d_n(\underline{X}_n)$ ειναι παιρνη τιμες που να ειναι κοντα στην αγνωστη $q(\theta)$, με μεγαλη πιθανοτητα — ως προς $F(\cdot|\theta)$ — και αυτο για "λογικα" μεγεθη n του δειγματος. Αυτο ειναι που, στην ουσια, θεωρουμε "καλη αποδοση" μιας εκτιμητριας και σε οσα ακολουθουν θα κανουμε αυτη την εννοια ακριβεστερη, μεσω διαφορων κριτηριων καλης αποδοσης εκτιμητριων.

Ενα βασικο μεσο ελεγχου της αποδοσης μιας εκτιμητριας — το οποιο εχουμε ηδη γνωρισει — ειναι η συναρτηση κινδυνου $R(\theta, d_n)$, $\theta \in \Theta$, της εκτιμητριας d_n, και ειδικοτερα το μεσο τετραγωνικο σφαλμα

$$(3.116) \quad R(\theta, d_n) \equiv \text{ΜΤΣ}(\theta, d_n) := E_\theta\left[d_n(\underline{X}_n) - q(\theta)\right]^2, \ \theta \in \Theta.$$

Μικρο ΜΤΣ, για "λογικο" n, σημαινει οτι με "μεγαλη" πιθανοτητα η $d_n(\underline{X}_n)$ ειναι κοντα στην $q(\theta)$, οπως συναγεται απο την ακολουθη αμεση συνεπαγωγη της ανισοτητας του Chebyshev:

$$(3.117) \quad P_\theta\left(|d_n(\underline{X}_n) - q(\theta)| > \varepsilon \right) \leq \frac{\text{ΜΤΣ}(\theta, d_n)}{\varepsilon^2} =$$

$$= \varepsilon^{-2} \mathcal{D}_\theta(d_n(\underline{X}_n)) \left\{ 1 + \left[\frac{E_\theta[d_n(\underline{X}_n)] - q(\theta)}{\sqrt{\mathcal{D}_\theta(d_n(\underline{X}_n))}} \right]^2 \right\}$$

$$= \varepsilon^{-2} \mathcal{D}_\theta(d_n)\left[1 + \bar{b}(\theta, d_n)^2\right], \quad \forall \varepsilon > 0,$$

οπου καναμε χρηση του (2.13) και με:

$$(3.118) \quad \bar{b}(\theta, d_n) := \frac{E_\theta[d_n(\underline{X}_n)] - q(\theta)}{\sqrt{\mathcal{D}_\theta[d_n(\underline{X}_n)]}}, \quad \theta \in \Theta,$$

συμβολιζουμε τη συναρτηση <u>τυποποιημενου σφαλματος</u> της εκτιμητριας $d_n(\underline{X}_n)$ της $q(\theta)$, $\theta \in \Theta$.

Οι ΟΑΕΕΔ, μαλιστα, κατασκευαστηκαν ακριβως ετσι ωστε να ελαχιστοποιουν το ΜΤΣ μεταξυ ολων των αμεροληπτων εκτιμητριων, δηλαδη, αυτων για τις οποιες $\bar{b}(\theta, d_n) = 0$. Η συνθηκη της αμεροληψιας επεβληθη σ' αυτες τις εκτιμητριες, και για να αφαλειφθη το συστηματικο σφαλμα $b(\theta, d_n) := E_\theta[d_n(\underline{X}_n)] - q(\theta)$, της μεροληψιας, αλλα και για τον καθαρα τεχνικο λογο της υπαρξης μη τετριμμενου ελαχιστου της συναρτησης κινδυνου $R(\theta, d_n) = \text{ΜΤΣ}(\theta, d_n)$, $\theta \in \Theta$, με $d_n \in \mathcal{D}$. Κατ' αυτον τον τροπο αποκλειστηκαν, π.χ., οι τετριμμενες εκτιμητριες που περιγραφονται στην παραγραφο που ακολουθει την (3.2), αλλα μαζυ μ' αυτες αποκλειστηκαν και αλλες, "ελαφρα" μεροληπτες, εκτιμητριες, οι οποιες μαλιστα ειχαν ομοιομορφα μικροτερο ΜΤΣ απο τις αντιστοιχες ΟΑΕΕΔ, βλ. (3.56). Για να αποφευγουμε στο εξης τον αποκλεισμο τετοιων καλων – ως προς το ΜΤΣ – εκτιμητριων θα απαιτουμε απο τις εκτιμητριες, οχι τοσο το να ειναι αμεροληπτες, αλλα να ειναι <u>ασυμπτωτικα αμεροληπτες</u>, δηλαδη,

$$(3.119)\quad \bar{b}(\theta, d_n) := \frac{E_\theta[d_n(\underline{X}_n)] - q(\theta)}{\sqrt{\mathcal{D}_\theta(d_n(\underline{X}_n))}} \xrightarrow[n\to\infty]{} 0 \quad \forall \theta \in \Theta.$$

Επισης, θα απαιτουμε απο τη διασπορα των εκτιμητριων μας να τεινη στο μηδεν, οπως το $n \to \infty$. Μαλιστα, καλουμε μια εκτιμητρια d_n <u>ακριβη</u> αν

$$(3.120)\quad n\, \mathcal{D}_\theta[d_n(\underline{X}_n)] \xrightarrow[n\to\infty]{} c^2(\theta) < +\infty \quad \forall \theta \in \Theta.$$

Ακριβεις εκτιμητριες υπαρχουν συνηθως και αρα δεν χρειαζεται να απαιτησουμε τιποτα λιγοτερο απο τις εκτιμητριες μας.

Απο τις (3.117), (3.119) και (3.120), βλεπουμε οτι για μια ακριβη και ασυμπτωτικα αμεροληπτη εκτιμητρια d_n,

$$n^{1-\delta} P_\theta(|d_n(\underline{X}_n) - q(\theta)| > \varepsilon) \xrightarrow[n\to\infty]{} 0, \quad \forall \varepsilon > 0 \text{ και } \forall \delta > 0.$$

Δηλαδη, γρηγορα – για "λογικο" μεγεθος δειγματος n – η εκτιμητρια d_n της $q(\theta)$ βρισκεται κοντα της με μεγαλη πιθανοτητα.

(3.121) <u>Ορισμος</u>. Μια εκτιμητρια d_n της $q(\theta)$, $\theta \in \Theta$, καλειται <u>συνεπης</u> αν και μονο αν,

$$d_n(\underline{X}_n) \xrightarrow[n\to\infty]{P} q(\theta) \quad \forall\, \theta \in \Theta .$$

Σημειωνουμε οτι μια ακριβης, ασυμπτωτικα αμεροληπτη εκτιμητρια d_n της $q(\theta)$ ειναι συνεπης και μαλιστα ξερουμε και την <u>ταχυτητα συγκλισης</u> ($n^{1/2}$) της d_n στην αληθη $q(\theta)$:

(3.122) $$n^{1/2-\delta}\left[d_n(\underline{X}_n) - q(\theta)\right] \xrightarrow[n\to\infty]{P} 0 \quad , \quad \forall \delta>0, \ \forall\, \theta \in \Theta$$

διοτι, $\forall \varepsilon>0$, $P_\theta\left(n^{1/2-\delta}\,|d_n(\underline{X}_n) - q(\theta)| > \varepsilon\right) =$

$$= P_\theta\left(|d_n(\underline{X}_n) - q(\theta)| > \frac{\varepsilon}{n^{1/2-\delta}}\right) \leq \frac{1}{\varepsilon^2 n^{2\delta}}\, n\mathcal{D}_\theta(d_n)\left[1+\bar{b}(\theta,d_n)^2\right]$$

$\xrightarrow[n\to\infty]{} 0 \quad \forall\, \theta\in\Theta$, οπου κοναμε χρηση της (3.117).

(3.123) <u>Παραδειγμα</u>. Εστω $\underline{X}_n = (X_1,\dots,X_n)$ α.ι. $N(\mu,\sigma^2)$, $\underline{\theta} \equiv (\mu,\sigma^2) \in \Theta \equiv$ $\equiv \mathbb{R}\times(0,+\infty) \subseteq \mathbb{R}^2$. Θα δειξουμε οτι:

(α) Η εμπ και ΟΑΕΕΔ $d_n(\underline{X}_n) = \bar{X}_n$ της $q(\underline{\theta}) = \mu$ ειναι αμεροληπτη, ακριβης και αρα συνεπης (ισχυει μαλιστα η (3.122)).

(β) Η εμπ $d_n(\underline{X}_n) = \hat{\sigma}_n^2$ — βλ. (3.83) — της $q(\underline{\theta}) = \sigma^2$ ειναι ασυμπτωτικα αμεροληπτη, ακριβης και αρα συνεπης (ισχυει μαλιστα η (3.122)).

(γ) Η ΟΑΕΕΔ $d_n(\underline{X}_n) = S_n^2$ — βλ. (3.48) — της $q(\underline{\theta}) = \sigma^2$ ειναι αμεροληπτη, ακριβης και αρα συνεπης (ισχυει μαλιστα η (3.122)).

Εχουμε:

(α) $E_{\underline{\theta}}[d_n(\underline{X}_n)] = E_{\underline{\theta}}\bar{X}_n = E_{\underline{\theta}}\left\{\frac{1}{n}\Sigma_{i=1}^n X_i\right\} = \frac{1}{n}\Sigma_{i=1}^n E_{\underline{\theta}}X_i = \frac{1}{n}\Sigma_{i=1}^n \mu = \mu$

$\mathcal{D}_{\underline{\theta}}[d_n(\underline{X}_n)] = \mathcal{D}_{\underline{\theta}}(\bar{X}_n) = \mathcal{D}_{\underline{\theta}}\left(\frac{1}{n}\Sigma_{i=1}^n X_i\right) = \frac{1}{n^2}\Sigma_{i=1}^n \mathcal{D}_{\underline{\theta}}(X_i) =$

$$= \frac{1}{n^2}\Sigma_{i=1}^n \sigma^2 = \sigma^2/n \quad \Rightarrow$$

$$\Rightarrow n\mathcal{D}_{\underline{\theta}}(\bar{X}_n) = \sigma^2 \equiv c^2(\underline{\theta}) \quad \forall\, \underline{\theta}\in\Theta .$$

Αρα, $P_{\underline{\theta}}(|\bar{X}_n - \mu| > \varepsilon) \leq \varepsilon^{-2} E_{\underline{\theta}}|\bar{X}_n-\mu|^2 = \varepsilon^{-2}\mathcal{D}_{\underline{\theta}}(\bar{X}_n) = \frac{\sigma^2}{n\varepsilon^2}$

$\xrightarrow[n\to\infty]{} 0 \quad \forall \varepsilon>0$ και $\forall\, \underline{\theta}\in\Theta$,

δηλαδη, $\bar{X}_n \xrightarrow[n\to\infty]{P} \mu = q(\underline{\theta}) \quad \forall\, \underline{\theta}\in\Theta$.

(β) $n\hat{\sigma}_n^2 = \Sigma_{i=1}^n (X_i - \bar{X}_n)^2 \sim \sigma^2\chi^2_{n-1} \quad \Rightarrow$

$$\Rightarrow \begin{cases} E_{\underline{\theta}}\{n\hat{\sigma}_n^2\} = \sigma^2(n-1) & \Rightarrow b(\underline{\theta},\hat{\sigma}_n^2) = -\dfrac{\sigma^2}{n}, \\ \mathcal{D}_{\underline{\theta}}\{n\hat{\sigma}_n^2\} = 2\sigma^4(n-1) & \Rightarrow \mathcal{D}_{\underline{\theta}}(\hat{\sigma}_n^2) = \dfrac{2(n-1)\sigma^4}{n^2}. \end{cases}$$

Αρα,

$$n\mathcal{D}_{\underline{\theta}}(\hat{\sigma}_n^2) = 2\sigma^4\left(1-\frac{1}{n}\right) \xrightarrow[n\to\infty]{} 2\sigma^4 \equiv c^2(\underline{\theta}) \quad \forall \underline{\theta}\in\Theta,$$

και,

$$\bar{b}(\theta,\hat{\sigma}_n^2) = \frac{-\sigma^2/n}{\sqrt{2(n-1)}\,\sigma^2/n} = -\frac{1}{\sqrt{2(n-1)}} \xrightarrow[n\to\infty]{} 0 \quad \forall \theta\in\Theta,$$

είναι δηλαδή, η $\hat{\sigma}_n^2$ ακριβής και ασυμπτωτικά αμερόληπτη.

(γ) Η S_n^2 είναι φυσικά αμερόληπτη — είναι η ΟΑΕΕΔ — και εφόσον $(n-1)S_n^2 \sim \Sigma_{i=1}^n (X_i-\bar{X}_n)^2 \sim \sigma^2\chi^2_{n-1}$, έχουμε

$$\mathcal{D}_{\underline{\theta}}((n-1)S_n^2) = 2\sigma^4(n-1) \Rightarrow \mathcal{D}_{\underline{\theta}}(S_n^2) = \frac{2\sigma^4}{n-1} \Rightarrow$$

$$\Rightarrow n\mathcal{D}_{\underline{\theta}}(S_n^2) = 2\sigma^4\left(1+\frac{1}{n-1}\right) \xrightarrow[n\to\infty]{} 2\sigma^4 \equiv c^2(\underline{\theta}) \ \forall \underline{\theta}\in\Theta,$$

είναι δηλαδή η S_n^2 ακριβής.

(3.124) <u>Παράδειγμα</u>. Έστω $\underline{X}_n = (X_1,\dots,X_n)$ α.ι. $\mathcal{U}(0,\theta)$, $\theta\in\Theta=(0,+\infty)$. Από την (3.86) (βλ. επίσης (3.27), (3.33) και (2.17α)) έχουμε ότι η ΟΑΕΕΔ της θ είναι η $d_n(\underline{X}_n) := (1+\frac{1}{n})X_{nn}$ και από την (3.85) ότι η εμΠ της θ είναι η $\hat{\theta}_n \equiv \hat{\theta}_n(\underline{X}_n) := X_{nn}$.

Θα δείξουμε ότι:

(α) οι d_n, $\hat{\theta}_n$ είναι ακριβείς,

(β) η d_n είναι φυσικά αμερόληπτη, αλλά η $\hat{\theta}_n$ <u>δεν</u> είναι ούτε ασυμπτωτικά αμερόληπτη. Είναι όμως και οι δύο <u>συνεπείς</u>.

<u>Έχουμε:</u>

(α) Από την (2.17β) έχουμε ότι $\mathcal{D}_\theta(\hat{\theta}_n) = \dfrac{n\theta^2}{(n+1)^2(n+2)} \Rightarrow$

$$\Rightarrow n\mathcal{D}_\theta(\hat{\theta}_n) = \frac{\theta^2}{n+2}\left(1-\frac{1}{n+1}\right)^2 \xrightarrow[n\to\infty]{} 0 \equiv c^2(\theta) \quad \forall\theta\in\Theta.$$

Επίσης,

$$\mathcal{D}_\theta(d_n) = \left(1+\frac{1}{n}\right)^2\mathcal{D}_\theta(X_{nn}) = \frac{(n+1)^2}{n^2}\,\frac{n\theta^2}{(n+1)^2(n+2)} = \frac{\theta^2}{n(n+2)} \Rightarrow$$

$$\Rightarrow n\mathcal{D}_\theta(d_n) = \frac{\theta^2}{n+2} \xrightarrow[n\to\infty]{} 0 \equiv c^2(\theta) \ \forall\theta\in\Theta.$$

(β) $b(\theta,\hat{\theta}_n) = -\dfrac{\theta}{n+1} \xrightarrow[n\to\infty]{} 0 \quad \forall\theta\in\Theta$, <u>αλλά</u>

$$\bar{b}(\theta,\hat{\theta}_n)=\frac{b(\theta,\hat{\theta}_n)}{\sqrt{\mathcal{D}_\theta(\hat{\theta}_n)}}=\frac{-\frac{\theta}{n+1}}{\frac{\theta}{n+1}\sqrt{\frac{n}{n+2}}}=-\sqrt{1+\frac{2}{n}}\xrightarrow[n\to\infty]{}-1\neq 0$$

$\forall\theta\in\Theta$ και αρα η $\hat{\theta}_n=X_{nn}$ δεν ειναι ουτε αμεροληπτη ουτε ασυμπτωτικα αμεροληπτη.

Για τη συνεπεια των d_n, $\hat{\theta}_n$, θα δειξουμε πρωτα οτι:

(3.125) $n(\theta-X_{nn})\xrightarrow[n\to\infty]{d}\mathcal{E}(\lambda=\frac{1}{\theta})\quad\forall\theta\in\Theta$:

αφο την (2.17) εχουμε οτι $X_{nn}\sim f(t|\theta)=n\theta^{-n}t^{n-1}1(0<t<\theta)$ και αρα $F(t|\theta)=\left(\frac{t\wedge\theta}{\theta}\right)^n\Rightarrow\forall y>0$

$$P_\theta(n(\theta-X_{nn})\le y)=P(X_{nn}\geq\theta-\frac{y}{n})=1-F(\theta-\frac{y}{n}|\theta)=$$
$$=1-\left(1-\frac{y/\theta}{n}\right)^n\xrightarrow[n\to\infty]{}1-e^{-y/\theta}=P_\theta(Y\le y),$$

οπου $Y\sim\mathcal{E}(\lambda=\frac{1}{\theta})$, ισχυει δηλαδη η (3.125).

Αρα, αφο το Θεωρημα (1.93) του Slutsky, εχουμε οτι:

$$n^{1-\delta}(X_{nn}-\theta)=-\frac{1}{n^\delta}\,n(X_{nn}-\theta)\xrightarrow[n\to\infty]{d}0,$$

και αρα αφο την (1.84δ),

(3.126) $n^{1-\delta}(X_{nn}-\theta)\xrightarrow[n\to\infty]{P_\theta}0\quad\forall\delta>0$ και $\forall\theta\in\Theta$.

Αρα, η $\hat{\theta}_n$ ειναι συνεπης και μαλιστα με πολυ μεγαλη ταχυτητα συγκλισεως (n). Επισης, η d_n ειναι συνεπης:

$$n(\theta-d_n)=n[\theta-(1+\tfrac{1}{n})X_{nn}]=n(\theta-X_{nn})-X_{nn}$$
$$\xrightarrow[n\to\infty]{d}\mathcal{E}(\lambda=\tfrac{1}{\theta})-\theta\qquad\forall\theta\in\Theta,$$

αφο τις (3.125) και (3.126) και το Θεωρημα του Slutsky. Η συνεπεια λοιπον της d_n — με την ιδια με την $\hat{\theta}_n$ ταχυτητα συγκλισεως — συναγεται αφο την τελευταια συγκλιση κατα κατανομη, οπως η (3.126) αφο την (3.125).

Εδω, επι πλεον της (ταχειας) συνεπειας των d_n, $\hat{\theta}_n$, αποδειξαμε και ισχυροτερο την (ταχεια) συγκλιση τους κατα κατανομη. Δηλαδη οτι: $\forall\theta\in\Theta$

(3.127) $n(\hat{\theta}_n-\theta)\xrightarrow[n\to\infty]{d}-Y$, οπου $Y\sim f(y|\theta)=\frac{1}{\theta}e^{-y/\theta}1(y>0)$

(3.128) $n(d_n-\theta)\xrightarrow[n\to\infty]{d}-Z$, οπου $Z\sim f(z|\theta)=\frac{1}{\theta}e^{-(z+\theta)/\theta}1(z>-\theta)$.

Οι ασυμπτωτικες κατανομες των (3.127) και (3.128) ειναι η εξαιρεση, συνηθως η ασυμπτωτικη κατανομη μιας ακριβους, ασυμπτωτικα -τουλαχιστον- αμεροληπτης εκτιμητριας ειναι, οπως θα δουμε αργοτερα, μια $N(0, c^2(\theta))$ με το $c^2(\theta)$ της (3.120)

Απο την (3.117) και τη σχεση μεταξυ του ΜΤΣ και της διασπορας μιας εκτιμητριας, βλεπουμε οτι η καταλληλοτητα μιας εκτιμητριας d_n της $q(\theta)$, δηλαδη, το ποσο κοντα στην $q(\theta)$ βρισκεται και το με ποια ταχυτητα την πλησιαζει -ταχυτητα συνεπειας της- εξαρταται κατα κριτιμο τροπο απο την διασπορα της - το μετρο $\frac{1}{\mathcal{D}_\theta(d_n)}$ της ακριβειας της. Ειναι λοιπον λογικο να συγκρινουμε την αποδοση δυο εκτιμητριων, οι οποιες ειναι τουλαχιστον ασυμπτωτικα συνεπεις, μεσω των διασπορων τους.

(3.129) Ορισμος. Η σχετικη αποδοση (σ.α.) μιας εκτιμητριας $d_n^{(1)}(\underline{X}_n)$ ως προς την $d_n^{(2)}(\underline{X}_n)$ της $q(\theta)$, $\theta \in \Theta$, οριζεται ως:

$$e(\theta | d_n^{(1)}, d_n^{(2)}) := \frac{\mathcal{D}_\theta(d_n^{(2)})}{\mathcal{D}_\theta(d_n^{(1)})}, \quad \theta \in \Theta,$$

εφοσον αυτες ειναι αμεροληπτες (ή εστω ασυμπτωτικα αμεροληπτες).

(3.130) Ορισμος. Η ασυμπτωτικη σχετικη αποδοση (α.σ.α.) μιας εκτιμητριας $d_n^{(1)}(\underline{X}_n)$ ως προς την $d_n^{(2)}(\underline{X}_n)$ της $q(\theta)$, $\theta \in \Theta$, οριζεται ως:

$$e_\infty(\theta | d_n^{(1)}, d_n^{(2)}) := \lim_{n\to\infty} \frac{\mathcal{D}_\theta(d_n^{(2)})}{\mathcal{D}_\theta(d_n^{(1)})}, \quad \theta \in \Theta,$$

εφοσον αυτες ειναι (τουλαχιστον) ασυμπτωτικα αμεροληπτες.

Παρατηρουμε, οτι αν οι $d_n^{(1)}, d_n^{(2)}$ ειναι επι πλεον ακριβεις, δηλαδη, $n\mathcal{D}_\theta(d_n^{(i)}) \xrightarrow[n\to\infty]{} c_i^2(\theta)$, $i=1,2$, τοτε,

$$e_\infty(\theta | d_n^{(1)}, d_n^{(2)}) = \frac{C_2^2(\theta)}{C_1^2(\theta)}, \quad \theta \in \Theta.$$

(3.131) <u>Παραδειγμα</u>. Εστω $\underline{X}_n = (X_1, \dots, X_n)$ α.ι. $N(\mu, \sigma^2)$, $\underline{\theta} = (\mu, \sigma^2) \in \Theta = \mathbb{R} \times (0, +\infty)$. Βρητε την σ.α και α.σ.α της εμπ $\hat{\sigma}_n^2$ της σ^2 ως προς την ΟΑΕΕΔ S_n^2.

<u>Απαντ.</u> Απο την (3.123β) η $\hat{\sigma}_n^2$ ειναι ασυμπτωτικα αμεροληπτη. Εχουμε δε οτι η $\mathcal{D}_{\underline{\theta}}(\hat{\sigma}_n^2) = \dfrac{2(n-1)\sigma^4}{n^2}$ και η $\mathcal{D}_{\underline{\theta}}(S_n^2) = \dfrac{2\sigma^4}{n-1}$, $\forall \underline{\theta} \in \Theta$. Αρα, (βλ. και (3.56)), (η συγκριση ειναι καταχρηστικη: $E_{\underline{\theta}}\hat{\sigma}_n^2 \neq \sigma^2$);

$$e(\underline{\theta} | \hat{\sigma}_n^2, S_n^2) = \left(\frac{n}{n-1}\right)^2 = \left(1 + \frac{1}{n-1}\right)^2 > 1 \quad \forall \underline{\theta} \in \Theta,$$

και,

$$e_\infty(\underline{\theta} | \hat{\sigma}_n^2, S_n^2) = \lim_{n \to \infty} \left(\frac{n}{n-1}\right)^2 = 1 \quad \forall \underline{\theta} \in \Theta.$$

(3.132) <u>Παραδειγμα</u>. Εστω $\underline{X}_n =$, α.ι. $U(0, \theta)$, $\theta > 0$. Ειδαμε – βλ., π.χ., (3.124) – οτι η ΟΑΕΕΔ της θ ειναι η $d_n(\underline{X}_n) := \left(1 + \frac{1}{n}\right) X_{nn}$, με $\mathcal{D}_\theta(d_n) = \dfrac{\theta^2}{n(n+2)}$.

Ειναι ευκολο να δη κανεις οτι η εκτιμητρια της μεθοδου των της θ ειναι η $\tilde{\theta}_n := 2\bar{X}_n$, η οποια ειναι αμεροληπτη και $\mathcal{D}_\theta(\tilde{\theta}_n) = 4\mathcal{D}_\theta(\bar{X}_n) = \dfrac{\theta^2}{3n}$. Αρα,

$$e(\theta | d_n, \tilde{\theta}_n) = \frac{\theta^2/(3n)}{\theta^2/[n(n+2)]} = \frac{n+2}{3} >> 1 \quad \forall \theta > 0,$$

και μαλιστα,

$$e_\infty(\theta / d_n, \tilde{\theta}_n) = \lim_{n \to \infty} \frac{n+2}{3} = +\infty \quad \forall \theta > 0.$$

(3.133) <u>Παραδειγμα</u>. Εστω $\underline{X}_n = (X_1, \dots, X_n)$ α.ι. $N(\mu, \sigma^2)$, $\Theta = \mathbb{R} \times (0, +\infty)$. Εχουμε δει οτι η εμπ της μ ειναι η $\bar{X}_n$. Δεδομενου ομως οτι η μ ειναι και η μεση της κατανομης αυτης, μια αλλη λογικη εκτιμητρια της μ ειναι η μεση του δειγματος $\hat{X}_n$.

Θα δουμε οτι η α.σ.α, της εμπ $\bar{X}_n$ ως προς την $\hat{X}_n$ ειναι:

$$e_\infty(\underline{\theta} | \bar{X}_n, \hat{X}_n) = \frac{\pi}{2} \approx 1.57 \quad \forall \underline{\theta} = (\mu, \sigma^2) \in \Theta.$$

Με στοχό αυτό, θα σκιαγραφήσουμε πρώτα τη βασική ιδέα για την απόδειξη του ότι : αν $X_1, \dots, X_n$ α.ι. f, τότε

(3.134) $n\mathcal{D}(\hat{X}_n) \xrightarrow[n\to\infty]{} \dfrac{1}{4f(m)^2}$, όπου $m := F^{-1}(1/2)$:

$$P(\sqrt{n}(\hat{X}_n - m) \le t) = P(\hat{X}_n \le m + \tfrac{t}{\sqrt{n}})$$
$$\approx P(\#\{X_i \le m + \tfrac{t}{\sqrt{n}}, i = 1, \dots, n\} \ge \tfrac{n}{2})$$
$$= P(\mathbb{F}_n(m + \tfrac{t}{\sqrt{n}}) \ge \tfrac{1}{2} = F(m)) =$$
$$= P\left(\sqrt{n}\left[\mathbb{F}_n(m + \tfrac{t}{\sqrt{n}}) - F(m + \tfrac{t}{\sqrt{n}})\right] \ge -t\,\frac{F(m+\frac{t}{\sqrt{n}}) - F(m)}{t/\sqrt{n}}\right),$$

και άρα — βλ. (1.89) — έχουμε:

$$P(\sqrt{n}(\hat{X}_n - m) \le t) \xrightarrow[n\to\infty]{} P(N(0, F(m)[1-F(m)]) \ge -t f(m))$$
$$= P(N(0, \tfrac{1}{4}) \ge -t f(m)) = P(N(0, \tfrac{1}{4f(m)^2}) \ge -t)$$
$$= 1 - P(N(0, \tfrac{1}{4f(m)^2}) \le -t) = P(N(0, \tfrac{1}{4f(m)^2}) \le t),$$

δηλαδή,

(3.135) $\sqrt{n}(\hat{X}_n - m) \xrightarrow[n\to\infty]{d} N(0, \dfrac{1}{4f(m)^2})$.

Βλέπουμε λοιπόν ότι, από την (3.134),

$$n\mathcal{D}_{\underline{\theta}}(\hat{X}_n) \xrightarrow[n\to\infty]{} \frac{1}{4\left[\frac{1}{\sigma}\phi(0)\right]^2} = \frac{\pi\sigma^2}{2} \quad \forall\, \underline{\theta} \in \Theta.$$

Επίσης, από την (3.123a), $n\mathcal{D}_{\underline{\theta}}(\bar{X}_n) = \sigma^2 \ \forall \theta \in \Theta$.

Άρα,

$$e_\infty(\theta \mid \bar{X}_n, \hat{X}_n) = \lim_{n\to\infty} \frac{\mathcal{D}_{\underline{\theta}}(\hat{X}_n)}{\mathcal{D}_{\underline{\theta}}(\bar{X}_n)} = \lim_{n\to\infty} \frac{n\mathcal{D}_{\underline{\theta}}(\hat{X}_n)}{n\mathcal{D}_{\underline{\theta}}(\bar{X}_n)} =$$
$$= \frac{\frac{\pi\sigma^2}{2}}{\sigma^2} = \frac{\pi}{2} \approx 1.57.$$

(3.136) <u>Άσκηση</u>. Έστω $X_1, \dots, X_n$ α.ι. Laplace(θ, λ), $\theta \in \mathbb{R}$, $\lambda > 0$. Από το (3.87) έχουμε ότι η εμπ της θ είναι η $\hat{X}_n$. Αλλά,

η εκτιμήτρια της μεθόδου των ροπών της θ είναι η $\bar{X}_n$.
Δείξτε ότι η α.σ.α. της εμπ $\hat{X}_n$ ως προς την $\bar{X}_n$ είναι:
$e_\infty(\theta \mid \hat{X}_n, \bar{X}_n) = 2$.

Από τα (3.133) και (3.136) φαίνεται ότι για την εκτίμηση της παραμέτρου τοποθεσίας θ μιας συμμετρικής κατανομής (με $\theta = F^{-1}(1/2) = E_\theta X$) δεν είναι πάντα η $\hat{X}_n$ προτιμότερη της $\bar{X}_n$ ή αντιστρόφως, εξαρτάται από την κατανομή. Πάντως και στις δύο αυτές περιπτώσεις η εμ.π. ήταν προτιμότερη της ανταγωνίστριας της.

(3.137) <u>Παρατήρηση.</u> Η πρακτική σημασία της ασυμπτωτικής σχετικής απόδοσης e_∞ μιας εκτιμήτριας $d_n^{(1)}$ ως προς την $d_n^{(2)}$ — τουλάχιστον όταν και οι δύο είναι ακριβείς — είναι η εξής:
εφόσον η $\mathcal{D}_\theta(d_n^{(i)}) \xrightarrow[n\to\infty]{} c_i^2(\theta)$, $i=1,2$,

τότε, $\mathcal{D}_\theta(d_n^{(i)}) \approx \dfrac{c_i^2(\theta)}{n}$, $i=1,2$.
Εφόσον όμως, $\dfrac{c_2^2(\theta)}{c_1^2(\theta)} = e_\infty$, έχουμε ότι

$$\mathcal{D}_\theta(d_n^{(1)}) \approx \frac{c_1^2(\theta)}{n} = \frac{c_2^2(\theta)}{e_\infty n} = \frac{c_2^2(\theta)}{m} \approx \mathcal{D}_\theta(d_m^{(2)}),$$

όπου $m = e_\infty \cdot n$. Δηλαδή, η δεύτερη εκτιμήτρια για να πετύχει την ίδια διασπορά και άρα ακρίβεια όπως η πρώτη, χρειάζεται μέγεθος δείγματος m περίπου ίσο προς e_∞ φορές το μέγεθος δείγματος της πρώτης. Έτσι, π.χ., στην περίπτωση της $N(\mu, \sigma^2)$ — βλ. (1.133) — η $\hat{X}_n$ χρειάζεται περίπου 57% περισσότερο δείγμα για να πετύχει την ίδια ακρίβεια με την $\bar{X}_n$, ενώ στην περίπτωση της Laplace(θ, λ) — βλ. (1.136) — της αρκεί το μισό περίπου δείγμα της $\bar{X}_n$.

Θα αναπτυξουμε τωρα ορισμενα απολυτα μετρα αποδοσης μιας εκτιμητριας. Θα συγκρινουμε δηλαδη την διασπορα της προς καποιο standard το οποιο εξαρταται απο την οικογενεια κατανομων που θεωρουμε και οχι τη συγκεκριμμενη εκτιμητρια – βλ.(3.   ).

Θα θεωρουμε στο εξης κατανομες $F(\cdot|\theta)$, $\theta\in\Theta$, των οποιων οι πυκνοτητες $f(\cdot|\theta)$ ή οι σ.μ.π. $p(\cdot|\theta)$ ειναι διαφορισιμες ως προς θ, και οτι επισης μπορουμε να εναλλασουμε την διαφορισιν ως προς θ και την ολοκληρωση ως προς x, π.χ.,

$$\int_{-\infty}^{+\infty}\left[\frac{\partial}{\partial\theta}\log f(x|\theta)\right]f(x|\theta)dx=\int_{-\infty}^{+\infty}\frac{\partial}{\partial\theta}f(x|\theta)dx=\frac{\partial}{\partial\theta}\int_{-\infty}^{+\infty}f(x|\theta)dx=\frac{\partial}{\partial\theta}1=0,$$

δηλαδη,

$$(3.138)\qquad E_\theta\left\{\frac{\partial}{\partial\theta}\log f(X_1|\theta)\right\}=0\qquad\forall\,\theta\in\Theta,$$

και,

$$E_\theta\{l_n'(\theta|\underline{X}_n)\}=0\left(=l_n'(\hat\theta_n|\underline{X}_n),\text{ αν υπαρχη ε.μ.π. }\hat\theta_n\right).$$

Θα δουλευουμε μονο με πυκνοτητες $f(\cdot|\theta)$, αλλα τα αποτελεσματα μας – με αναλογες αποδειξεις – θα ισχυουν και για σ.μ.π. $p(\cdot|\theta)$.

Οριζουμε ως πληροφορια κατα Fisher ή απλως πληροφορια του δειγματος $\underline{X}_n$ την συναρτηση – αν υπαρχη –

$$(3.139)\qquad I_n(\theta)\equiv I_{\underline{X}_n}(\theta):=\mathcal{D}_\theta\{l'(\theta|\underline{X}_n)\}=E_\theta\{l'(\theta|\underline{X}_n)^2\},\ \theta\in\Theta.$$

Παρατηρουμε, οτι αν οι $\underline{X}_n=(X_1,\dots,X_n)$ ειναι α.ι., τοτε,

$$(3.140)\qquad I_n(\theta)\equiv I_{\underline{X}_n}(\theta)=\mathcal{D}_\theta\left\{\Sigma_{i=1}^n\frac{\partial}{\partial\theta}\log f(X_i|\theta)\right\}=\Sigma_{i=1}^n\mathcal{D}_\theta\left\{\frac{\partial}{\partial\theta}\log f(X_i|\theta)\right\}$$
$$=\Sigma_{i=1}^n I_{X_i}(\theta)=nI_{X_1}(\theta)=nI_1(\theta)=nI(\theta),$$

οπου η συναρτηση $I(\theta)$ της πληροφοριας της κατανομης $F(\cdot|\theta)$ οριζεται ως:

$$(3.141)\qquad I(\theta):=\mathcal{D}_\theta\left\{\frac{\partial}{\partial\theta}\log f(X_1|\theta)\right\}=E_\theta\left\{\frac{\partial}{\partial\theta}\log f(X_1|\theta)\right\}^2=$$
$$=\int_{-\infty}^{+\infty}\left[\frac{\partial}{\partial\theta}\log f(x|\theta)\right]^2f(x|\theta)\,dx=$$
$$=\int_{-\infty}^{+\infty}\frac{\left[\frac{\partial}{\partial\theta}f(x|\theta)\right]^2}{f(x|\theta)}dx=4\int_{-\infty}^{+\infty}\left\{\frac{\partial}{\partial\theta}\sqrt{f(x|\theta)}\right\}^2dx,\quad\theta\in\Theta,$$

οπου εχουμε κανει χρηση, οπως και για την (3.139), της (3.138).

Επισης, αν υπαρχη η $l''(\theta)$, παραγωγιζοντας την (3.138), εχουμε:

$$0=\frac{\partial}{\partial\theta}\int_{-\infty}^{+\infty}\left[\frac{\partial}{\partial\theta}\log f(x|\theta)\right]f(x|\theta)dx=\int_{-\infty}^{+\infty}\left[\frac{\partial^2}{\partial\theta^2}\log f(x|\theta)\right]f(x|\theta)dx+$$
$$+\int_{-\infty}^{+\infty}\left[\frac{\partial}{\partial\theta}\log f(x|\theta)\right]\frac{\partial}{\partial\theta}f(x|\theta)\,dx=\int_{-\infty}^{+\infty}\left[\frac{\partial^2}{\partial\theta^2}\log f(x|\theta)\right]f(x|\theta)dx+$$

$$+\int_{-\infty}^{+\infty}\left[\frac{\partial}{\partial\theta}\log f(x|\theta)\right]^2 f(x|\theta)dx = E_\theta\left\{\frac{\partial^2}{\partial\theta^2}\log f(X_1|\theta)\right\} + I(\theta) \Rightarrow$$

(3.142) $$I(\theta) = -E_\theta\left\{\frac{\partial^2}{\partial\theta^2}\log f(X_1|\theta)\right\}, \quad \theta\in\Theta,$$

και,

$$I_n(\theta) \equiv I_{\underline{X}_n}(\theta) = -E_\theta\{\ell''(\theta|\underline{x})\}, \quad \theta\in\Theta .$$

Απο τις διαφορες εκφρασεις της πληροφοριας κατα Fisher, ειναι σαφες οτι αυτη ειναι ενα μετρο της μεταβλητοτητος της συναρτησης πιθανοφανειας $\ell_n(\theta) = \ell_n(\theta|\underline{X}_n)$. Απο δε τον N.M.A.-βλ.(1.86), αν υπαρχουν οι $\ell_n''(\theta)$, $I(\theta)$ και η συναρτηση εντροπιας της $F(\cdot|\theta)$:

(3.143) $$E(\theta) := -E_\theta\{\log f(X_1|\theta)\} = -\int_{-\infty}^{+\infty}[\log f(x|\theta)]f(x|\theta)dx, \quad \theta\in\Theta,$$

τοτε, για $\theta\in\Theta$,

(3.144) $$\frac{1}{n}\ell_n(\theta) = \frac{1}{n}\sum_{i=1}^n \log f(X_i|\theta) \xrightarrow[n\to\infty]{P} E_\theta\{\log f(X_1|\theta)\} = -E(\theta),$$

(3.145) $$\frac{1}{n}\ell_n'(\theta) = \frac{1}{n}\sum_{i=1}^n \frac{\partial}{\partial\theta}\log f(X_i|\theta) \xrightarrow[n\to\infty]{P} E_\theta\left\{\frac{\partial}{\partial\theta}\log f(X_1|\theta)\right\} = 0,$$

(3.146) $$\frac{1}{n}\ell_n''(\theta) = \frac{1}{n}\sum_{i=1}^n \frac{\partial^2}{\partial\theta^2}\log f(X_i|\theta) \xrightarrow[n\to\infty]{P} E_\theta\left\{\frac{\partial^2}{\partial\theta^2}\log f(X_1|\theta)\right\} = -I(\theta).$$

Επισης, απο την στοιχειωδη ανισοτητα $\log(1+x)\le x \ \forall x>-1$ με ισοτητα αν και μονο αν $x=0$, εχουμε οτι:

$$E(\theta',\theta) := -E_\theta\{\log f(X_1|\theta')\} = -E_\theta\left\{\log\frac{f(X_1|\theta')}{f(X_1|\theta)}\right\} + E(\theta,\theta)$$

$$= -E_\theta\left\{\log\left[1+\left(\frac{f(X_1|\theta')}{f(X_1|\theta)}-1\right)\right]\right\} + E(\theta,\theta) \geqq$$

$$\geq E(\theta,\theta) - E_\theta\left\{\frac{f(X_1|\theta')}{f(X_1|\theta)}-1\right\} = E(\theta,\theta) - \int_{-\infty}^{+\infty} f(x|\theta')dx + 1 = E(\theta,\theta),$$

και οριζουμε την πληροφορια κατα Kullback-Leibler, ως:

(3.147) $$K(\theta',\theta) := E(\theta',\theta) - E(\theta,\theta) = E(\theta',\theta) - E(\theta) =$$
$$= E_\theta\{\log f(X_1|\theta)\} - E_\theta\{\log f(X_1|\theta')\} \geqslant 0$$

και $K(\theta',\theta) = 0 \Leftrightarrow \theta'=\theta$.

Τωρα, απ' το Θεωρημα Taylor εχουμε οτι:

$$\log f(x|\theta') \approx \log f(x|\theta) + (\theta'-\theta)\frac{\partial}{\partial\theta}\log f(x|\theta) + \frac{(\theta'-\theta)^2}{2}\frac{\partial^2}{\partial\theta^2}\log f(x|\theta),$$

δηλαδη, απο τις (3.138), (3.142), (3.147) εχουμε οτι:

(3.148) $$K(\theta',\theta) \approx \frac{1}{2} I(\theta)(\theta'-\theta)^2 , \quad \theta, \theta' \in \Theta .$$

Επισης, οπως στην περιπτωση της (3.144), εχουμε οτι:

$\frac{1}{n} l_n(\theta') \xrightarrow[n\to\infty]{P_\theta} -E(\theta',\theta)$, και αρα, για μεγαλα n

$l_n(\theta') \approx -E(\theta',\theta)$, ουτως ωστε απο τις (3.147), (3.148) με $\theta'=\hat{\theta}_n$, εχουμε οτι:

(3.149) $$E_\theta\{l_n(\theta|\underline{X}_n)\} - l_n(\hat{\theta}_n|\underline{X}_n) \approx nK(\hat{\theta}_n,\theta) \approx \frac{1}{2} I_{\underline{X}_n}(\theta)(\hat{\theta}_n-\theta)^2 \geq 0 .$$

Η σχεση αυτη δειχνει κατα σαφη, αν και ευρισικο τροπο, την αντιστοιχεια αναμεσα στην μεγιστοποιηση της πιθανοφανειας $l_n(\hat{\theta}_n|\underline{X}_n) \gtrsim E_\theta\{l_n(\theta|\underline{X}_n)\}$ (οπως δειχνει η (3.147)), και την ελαχιστοποιηση της αποστασεως $|\hat{\theta}_n-\theta|$ της εκτιμητριας απο την παραμετρο. Επισης η (3.149) προσφερει μια ερμηνεια της $I_{\underline{X}_n}(\theta)$ ως "πληροφοριας", δεδομενου οτι δειχνει οτι εστω και αν η μεγιστοποιηση της πιθανοφανειας εχει επιτυχει να κανη "μικρο" το αριστερο σκελος της (3.149), αν η $I_{\underline{X}_n}(\theta)$ ειναι "μικρη" τοτε η $|\hat{\theta}_n-\theta|$ μπορει να ειναι "μεγαλη" και αρα η "πληροφορια" του δειγματος δεν ηταν αρκετη για να μαθουμε με καποια ακριβεια τα περι την αληθη τιμη της θ. Αντιθετα, αν η $I_{\underline{X}_n}(\theta)$ ειναι "μεγαλη", η $|\hat{\theta}_n-\theta|$ πρεπει να ειναι "μικρη" και αρα η "πληροφορια" του δειγματος ηταν ικανη για μια καλη προσεγγιση της αληθους τιμης της αγνωστης παραμετρου θ.

Θα δουμε μαλιστα οτι — οπως θα επρεπε να αναμενεται απο την παραπανω συζητηση — η πληροφορια κατα Fisher του δειγματος συνδεεται με την διασπορα της εκτιμητριας της θ, συγκεκριμμενα εν γενει $n\mathcal{D}_\theta(\hat{\theta}_n) \xrightarrow[n\to\infty]{} \frac{1}{I(\theta)}$, δηλαδη, $\mathcal{D}_\theta(\hat{\theta}_n) \approx \frac{1}{I_{\underline{X}_n}(\theta)}$, για "μεγαλα" n — βλ. (3.).

Προς το παρον θα δειξουμε μια σχεσιμη με τα ανωτερω και βασικη για την εκτιμητικη ανισοτητα η οποια θα αποτελεσει τη βαση για τον ορισμο της (απολυτης) αποδοσης μιας εκτιμητριας.

(3.150) <u>Προταση</u>. Εστω $\underline{X}_n \sim f(\cdot|\theta) \overset{n}{=} p(\cdot|\theta)$, $\theta \in \Theta$, με $I_{\underline{X}_n}(\theta) < +\infty$. Εστω δε στατιστικη συναρτηση $d_n(\underline{X}_n)$ τετοια ωστε $\exists\ \psi(\theta) := E_\theta\{d_n(\underline{X}_n)\}$, $\theta \in \Theta$ και ειναι παραγωγισιμη. Τοτε, ισχυει η ανισοτητα των <u>Cramér</u>-<u>Frechet</u>-<u>Rao</u> :

$$\mathcal{D}_\theta\{d_n(\underline{X}_n)\} \geq \frac{[\psi'(\theta)]^2}{I_{\underline{X}_n}(\theta)} \qquad \forall\, \theta \in \Theta ,$$

με "=" αν και μονο αν $\exists\ \alpha(\theta), \beta(\theta)$ τετοιες ωστε

$$P_\theta\{l_n'(\theta|\underline{X}_n) = \alpha(\theta)\, d_n(\underline{X}_n) + \beta(\theta)\} = 1 , \quad \theta \in \Theta .$$

<u>Αποδ.</u> Εχουμε : $|\psi'(\theta)| = \left|\frac{\partial}{\partial\theta}\int d_n(\underline{x})\, f(\underline{x}|\theta)\, d\underline{x}\right|$

$$= \left|\int d_n(\underline{x}) \frac{\partial}{\partial\theta} f(\underline{x}|\theta)\, dx\right| = \left|\int\left\{d_n(\underline{x}) \frac{\partial}{\partial\theta}\log f(\underline{x}|\theta)\right\} f(\underline{x}|\theta)\, dx\right|$$

$$= \left|E_\theta\{d_n(\underline{X}_n)\, l'(\theta|\underline{X}_n)\}\right| = \left|cov(d_n(\underline{X}_n), l'(\theta|\underline{X}_n))\right|$$

$$\leq \sqrt{\mathcal{D}_\theta\{d_n(\underline{X}_n)\}}\sqrt{\mathcal{D}_\theta\{l'(\theta|\underline{X}_n)\}} = \sqrt{\mathcal{D}_\theta\{d_n(\underline{X}_n)\}}\sqrt{I_{\underline{X}_n}(\theta)} ,$$

οπου καναμε χρηση των (3.138) και (1.48γ). Τωρα, απο την (1.49ε) εχουμε οτι ισοτητα εχουμε αν και μονο αν με πιθανοτητα (P_θ) 1, $l'(\theta|\underline{X}_n) = \alpha(\theta)\, d_n(\underline{X}_n) + \beta(\theta)$ για καποιες συναρτησεις $\alpha(\theta), \beta(\theta)$ ανεξαρτητες του δειγματος.

Τελος, με τετραγωνισμο των μελων της ανισοτητος αυτης παιρνουμε την ανισοτητα C-F-R.

Παρατηρουμε οτι αν η εκτιμητρια $d_n(\underline{X}_n)$ ειναι αμεροληπτη εκτιμητρια της θ, δηλαδη, $\psi(\theta) = \theta\ \forall \theta \in \Theta$, τοτε

(3.151) $$\mathcal{D}_\theta\{d_n(\underline{X}_n)\} \geq \frac{1}{I_\nu(\theta)} \qquad \forall\, \theta \in \Theta .$$

Επισης αν οι $\underline{X}_n = (X_1, \dots, X_n)$ ειναι α.ι. τοτε

(3.152) $$\mathcal{D}\{d_n(\underline{X}_n)\} \geq \frac{[\psi'(\theta)]^2}{nI(\theta)} \quad \forall \theta \in \Theta$$

$$= \frac{1}{nI(\theta)} \quad \forall \theta \in \Theta,$$ αν επιπλεον η $d_n(\underline{X}_n)$ ειναι αμεροληπτη εκτιμητρια της θ.

(3.152) Ορισμος. Η αποδοση μιας αμεροληπτης (ή εστω ασυμπτωτικα αμεροληπτης) εκτιμητριας $d_n(\underline{X}_n)$ της $q(\theta)$, με $q(\cdot)$ διαφορισιμη, οριζεται ως:

$$e(\theta|d_n) := \frac{[q'(\theta)]^2}{I_{\underline{X}_n}(\theta)\,\mathcal{D}_\theta\{d_n(\underline{X}_n)\}} = \frac{[q'(\theta)]^2}{nI(\theta)\mathcal{D}_\theta(d_n)} \leq 1, \quad \theta \in \Theta.$$

Αν η $e(\theta|d_n) = 1$ η d_n καλειται πληρως αποδοτικη.

Απο τις (3.129) και (3.152) εχουμε οτι:

(3.153) $$e(\theta|d_n^{(1)}, d_n^{(2)}) = \frac{e(\theta|d_n^{(1)})}{e(\theta|d_n^{(2)})}, \quad \theta \in \Theta.$$

(3.154) Παραδειγμα. Εστω $X_1, \dots, X_n$ α.ι. $\mathcal{E}(\lambda = \frac{1}{\theta})$, $\theta \in \Theta = (0, +\infty)$.

(α) Βρητε την πληροφορια κατα Fisher της $\mathcal{E}(\frac{1}{\theta})$ και της $\mathcal{E}(\lambda)$.

(β) Βρητε την πληροφορια του δειγματος $I_{\underline{X}_n}(\theta)$, καθως και $I_{\underline{X}_n}(\lambda)$.

(γ) Η $\bar{X}_n$ ειναι η ε.μ.π. και η ΟΑΕΕΔ της θ. Δειξτε οτι,

$e(\theta|\bar{X}_n) = 1$, δηλαδη, η $\bar{X}_n$ ειναι πληρως αποδοτικη για την θ.

(δ) Δειξτε οτι η ΟΑΕΕΔ $d_n(\underline{X}_n) = (1 - \frac{1}{n})\frac{1}{\bar{X}_n}$ της λ δεν ειναι πληρως αποδοτικη (υποθεσατε οτι $n \geq 3$).

Απαντ. (α) $l_1(\theta|X_1) = \log f(X_1|\theta) = -\frac{1}{\theta}X_1 - \log\theta \Rightarrow l_1'(\theta|X_1) = \frac{1}{\theta^2}X_1 - \frac{1}{\theta} \Rightarrow$

$\Rightarrow I(\theta) = \mathcal{D}_\theta\{l_1'(\theta|X_1)\} = \mathcal{D}_\theta\{\frac{1}{\theta^2}X_1 - \frac{1}{\theta}\} = \frac{1}{\theta^4}\mathcal{D}_\theta(X_1) = \frac{1}{\theta^2}$.

Παρομοια, $I(\lambda) = \mathcal{D}_\lambda\{l_1'(\lambda|X_1)\} = \mathcal{D}_\lambda\{-X_1 + \frac{1}{\lambda}\} = \mathcal{D}_\lambda(X_1) = \frac{1}{\lambda^2}$.

Παρατηρουμε οτι $I(\lambda)\Big|_{\lambda=\frac{1}{\theta}} = \theta^2 \neq \frac{1}{\theta^2} = I(\theta)$ (για $\lambda \neq 1$).

Γενικοτερα, αν η $q(\theta) = \eta \in H = q(\Theta)$, $\theta \in \Theta$ ειναι διαφορισιμη και 1:1, τοτε,

(3.155) $$I(q(\theta)) = I^*(\theta)/[q'(\theta)]^2, \quad \theta \in \Theta,$$

οπου $I(\eta)$, $I^*(\theta)$ ειναι η πληροφορια κατα Fisher για τις η, θ αντιστοιχα:

έχουμε ότι, $l_1'(\eta) = \frac{d}{dq(\theta)} l_1(q(\theta)) \Big|_{\theta=q^{-1}(\eta)}$, δηλαδή,
$l_1'(q(\theta)) = \frac{d}{d\theta} l_1(q(\theta)) / q'(\theta) = [l_1^*(\theta)]' / q'(\theta)$, όπου $l_1^* = l_1 \circ q$. Άρα,
$I(\eta)\Big|_{\eta=q(\theta)} = \mathcal{D}_\eta\{l_1'(\eta)\}\Big|_{\eta=q(\theta)} = E_\theta\{[l_1^*(\theta)]'\} / [q'(\theta)]^2 = I^*(\theta)/[q'(\theta)]^2$.

(β) $I_{\underline{X}_n}(\theta) = n\, I(\theta) = n\theta^{-2}$ και $I_{\underline{X}_n}(\lambda) = n\, I(\lambda) = n\lambda^{-2}$.

(γ) $\mathcal{D}_\theta(\bar{X}_n) = \mathcal{D}_\theta(X_1)/n = \theta^2/n \Rightarrow$
$\Rightarrow e(\theta\,|\,\bar{x}_n) = [I_{\underline{X}_n}(\theta)\, \mathcal{D}_\theta(\bar{X}_n)]^{-1} = \left(\frac{n}{\theta^2}\, \frac{\theta^2}{n}\right)^{-1} = 1$.

(δ) Από την (3.52α) έχουμε ότι η μοναδική ΟΑΕΕΔ της λ είναι
η $d_n(\underline{X}_n) = \frac{n-1}{\sum_{i=1}^n X_i}$. Έχουμε δε ότι, η διασπορά
$\mathcal{D}_\lambda(d_n) = (n-1)^2 \mathcal{D}_\lambda(Y^{-1}) = (n-1)^2[E_\lambda Y^{-2} - (EY^{-1})^2]$, όπου $Y \sim \mathcal{G}(n,1)$.
Άρα, $EY^{-1} = \frac{\Gamma(n-1)}{\Gamma(n)}\lambda = \frac{\lambda}{n-1}$ και $EY^{-2} = \frac{\Gamma(n-2)}{\Gamma(n)}\lambda^2 = \frac{\lambda^2}{(n-1)(n-2)}$,
και άρα,
$\mathcal{D}_\lambda(d_n) = \frac{\lambda^2}{n-2} \Rightarrow e(\lambda\,|\,d_n) = \frac{\lambda^2}{n}\cdot\frac{(n-2)}{\lambda^2} = 1 - \frac{2}{n} < 1$

Για την εμπ $\hat{\lambda}_n = \frac{1}{\bar{X}_n}$, τώρα, έχουμε ότι
$\psi(\lambda) = n\, E_\lambda Y^{-1} = \frac{n}{n-1}\lambda$, $\mathcal{D}_\lambda(\hat{\lambda}_n) = n^2 \mathcal{D}_\lambda(Y^{-1}) = \frac{n^2\lambda^2}{(n-1)^2(n-2)}$,
και άρα,
$\bar{b}(\lambda;\hat{\lambda}_n) = \frac{\lambda/(n-1)}{n\lambda/[(n-1)\sqrt{n-2}]} = \frac{\sqrt{n-2}}{n} \xrightarrow[n\to\infty]{} 0 \;\; \forall \lambda > 0$, είναι δηλαδή
η $\hat{\lambda}_n$ ασυμπτωτικά αμερόληπτη, η δε απόδοση της (καταχρηστικά) είναι:
$e(\lambda\,|\,\hat{\lambda}_n) = \frac{[\psi'(\lambda)]^2}{I_{\underline{X}_n}(\lambda)\,\mathcal{D}_\lambda(\hat{\theta}_n)} = \frac{n-2}{n} = 1 - \frac{2}{n} = e(\lambda\,|\,d_n)\ (<1)$,

όπως θα έπρεπε να αναμένουμε.

(3.156) <u>Παράδειγμα</u>: Έστω $X_1, \dots, X_n$ α.ι. Bernoulli (p), $p \in \Theta = [0,1]$.
(α) Βρητε την πληροφορία κατά Fisher $I(p)$ και την $I_{\underline{X}_n}(p)$.
(β) Βρητε της απόδοση της εμπ και ΟΑΕΕΔ $\bar{X}_n$ της p.
<u>Απαντ.</u> (α) $l_1(p\,|\,X_1) = X_1 \log p + (1-X_1)\log(1-p) \Rightarrow l_1'(p\,|\,X_1) = \frac{X_1 - p}{p(1-p)} \Rightarrow$
$\Rightarrow I(p) = \mathcal{D}_p\left(\frac{X_1 - p}{p(1-p)}\right) = \frac{\mathcal{D}_p(X_1)}{[p(1-p)]^2} = \frac{1}{p(1-p)} \Rightarrow I_{\underline{X}_n}(p) = \frac{n}{p(1-p)}$.

(β) $\mathcal{D}_p(\bar{X}_n) = \frac{\mathcal{D}_p(X_1)}{n} = \frac{p(1-p)}{n} \Rightarrow e(p\,|\,\bar{X}_n) = 1$.

(3.157) Ασκηση. Εστω $X_1, \ldots, X_n$ α.ι. Bernoulli(p), $p \in [0,1]$, $n \geq 5$.

(α) Βρητε την αποδοση της ΟΑΕΕΔ $d_n(\underline{X}_n) = \frac{n}{n-1} \bar{X}_n (1-\bar{X}_n)$ της $q(p) = p(1-p)$. (βλ. 3.49).

(β) Βρητε (καταχρηστικα) την αποδοση της εμπ $q(\hat{p}) = \bar{X}_n(1-\bar{X}_n)$ της $q(p)$.

(Υποδ. $Y = \Sigma_1^n X_i \sim \mathcal{B}(n,p)$ και αρα για $m \leq n$, $E_p\{\prod_{l=0}^{m-1}(X-l)\} = [\prod_{l=0}^{m-1}(n-l)]\, p^m$.).

(3.158) Ασκηση. Για καθε μια απο τις ακολουθες κατανομες:

(i) $N(\theta, \sigma^2)$, $\theta \in \Theta = \mathbb{R}$, με σ^2 γνωστο,

(ii) $N(\mu, \theta)$, $\theta(=\sigma^2) \in \Theta = (0,+\infty)$, με μ γνωστο,

(iii) $\mathcal{G}(\alpha, \lambda = \frac{1}{\theta})$, $\theta \in \Theta = (0,+\infty)$, με α γνωστο,

(iv) $W(\alpha, \lambda = \frac{1}{\theta})$, $\theta \in \Theta = (0,+\infty)$, με α γνωστο,

(v) Γεωμετρικη$(p = \frac{1}{\theta})$, $\theta \in \Theta = (1,+\infty)$,

(vi) Poisson$(\lambda = \theta)$, $\theta \in \Theta = (0,+\infty)$,

(α) Βρητε τις πληροφοριες κατα Fisher $I(\theta)$, $I_{\underline{X}_n}(\theta)$,

(β) δειξτε οτι οι ΟΑΕΕΔ και οι εμπ των θ συμπιπτουν και ειναι πληρως αποδοτικες.

[Υποδ. Εδω, τις ΟΑΕΕΔ της θ μπορει κανεις να τις συναγη και ωςεξης: να βρη την εμπ της θ να δειξη οτι ειναι αμεροληπτη και πληρως αποδοτικη $\forall \theta \in \Theta$. Τοτε, απο την ανισοτητα C-F-R, η εμπ πρεπει να ειναι και η ΟΑΕΕΔ της θ.].

(3.159) Ασκηση. Εστω $X_1, \ldots, X_n$ α.ι. $N(\mu, \sigma^2)$, $\underline{\theta} = (\mu, \sigma^2) \in \Theta = \mathbb{R} \times (0,+\infty)$. Συγκρινατε τις διασπορες των εμπ και ΟΑΕΕΔ της σ^2 με το κατω φραγμα της ανισοτητος C-F-R.

(3.160) Ασκηση. Υπολογιστε τις $I(\lambda)$, $I_{\underline{X}_n}(\lambda)$ της $\mathcal{G}(\alpha, \lambda)$ με α γνωστο, και συγκρινατε τις διασπορες των ε.μ.π. και ΟΑΕΕΔ της λ με το κατω φραγμα της ανισοτητος C-F-R.

(3.161) <u>Ασκηση</u>. Εστω $X_1, \ldots, X_n \sim f(x|\theta) =$
$= \exp\{c(\theta)T(x) + d(\theta) + S(x)\}\, 1(x \in A)$, $\theta \in \Theta \subseteq \mathbb{R}$, η μονοπαραμετρικη εκθετικη οικογενεια των (3.18) και (3.19) και $\exists\, c'(\theta), \theta \in \Theta$.
Θεσατε $\bar{T}_n(\underline{X}_n) := \frac{1}{n}\sum_{i=1}^{n} T(X_i)$, $\eta \equiv c(\theta)$ και δειξτε οτι:

(α) $\psi(\theta) := E_\theta(\bar{T}_n) = -d'(\theta)/c'(\theta)$, $\theta \in \Theta$.

$\psi_0(\eta) := E_\eta(\bar{T}_n) = -d_0'(\eta)$, $\eta \in H = c(\Theta)$.

(β) $I_0(\eta) = -d_0''(\eta)$, $\eta \in \mathring{H}$.

$I(\theta) = -[c'(\theta)]^2\, d_0''(q(\theta))$

(γ) $\mathcal{D}_\eta(\bar{T}_n) = \dfrac{[\psi_0'(\eta)]^2}{n I_0(\eta)}$, $\eta \in \mathring{H}$.

$$\mathcal{D}_\theta(\bar{T}_n) = \frac{[\psi'(\theta)]^2}{n I(\theta)}, \quad \theta \in \Theta.$$

(<u>Σημ.</u> Μπορει να αποδειχθη <u>και</u> οτι – κατω απο ασθενεις συνθηκες ομαλοτητος τη $f(\cdot|\theta)$ — αν ισχυει η (γ) τοτε η $f(\cdot|\theta)$, $\theta \in \Theta$ ειναι μια μονοπαραμετρικη εκθετικη οικογενεια.)

(3.162) <u>Σημειωση</u>. Η πληροφορια κατα Fisher και η ανισοτητα C-F-R γενικευονται και στην περιπτωση που η διασταση του παραμετρικου χωρου Θ ειναι μεγαλυτερη του ένα. Θα υποδειξουμε τωρα αυτες τις γενικευσεις.

Κατα αναλογια προς την (3.142) ο πινακας πληροφοριας $I(\underline{\theta})$, $\underline{\theta} = (\theta_1, \ldots, \theta_k)^T \in \Theta \subseteq \mathbb{R}^k$, οριζεται ως:

(3.163) $I(\underline{\theta}) := \left[-E_{\underline{\theta}}\left\{ \frac{\partial^2}{\partial\theta_i \partial\theta_j} \log f(X_1|\underline{\theta}) \right\} \right] \in \mathbb{R}^k \times \mathbb{R}^k$, $i, j = 1, 2, \ldots, k$,

ειναι δε συμμετρικος και θετικα ορισμενος. Ειναι, δηλαδη, ο $I(\underline{\theta})$ η μεση τιμη της Hessian της $-l_1(\underline{\theta}|X_1)$.

Αναλογα, $I_{\underline{X}_n}(\underline{\theta}) := \left[-E_{\underline{\theta}}\left\{ \frac{\partial^2}{\partial\theta_i \partial\theta_j} l_n(\underline{\theta}|\underline{X}_n) \right\} \right] \in \mathbb{R}^k \times \mathbb{R}^k$.

Η ανισοτητα των Cramér-Frechét-Rao γενικευεται ως εξης: Εστω $d_n^{(i)}(\underline{X}_n)$ εκτιμητρια της $q_i(\underline{\theta})$, $i = 1, \ldots, r$; $\underline{d}_n := (d_n^{(1)}, \ldots, d_n^{(r)})^T$, εστω δε $\psi_i(\underline{\theta}) := E_{\underline{\theta}}\{d_n^{(i)}(\underline{X}_n)\}$, $i = 1, \ldots, r$, $\dot{\Psi}(\underline{\theta}) := \left[\frac{\partial \psi_i(\underline{\theta})}{\partial \theta_j} \right] \in \mathbb{R}^r \times \mathbb{R}^k$

ο Ιακωβιανος πινακας των $\psi_i(\theta)$, $i=1,\dots,r$. Αν με $\mathcal{D}_{\underline{\theta}}(\underline{d}_n)$ συμβολισουμε τον πινακα διασπορας της $\underline{d}_n=(d_n^{(1)},\dots,d_n^{(r)})^T$, δηλαδη,
$\mathcal{D}_{\underline{\theta}}(\underline{d}_n) = [\operatorname{cov}(d_n^{(i)},d_n^{(j)})] \in \mathbb{R}^r\times\mathbb{R}^r$, $i,j=1,\dots,r$, εχουμε οτι:
ο πινακας:

$$(3.164)\quad \mathcal{D}_{\underline{\theta}}(\underline{d}_n) - \dot{\Psi}(\theta)[I_{\underline{X}_n}(\underline{\theta})]^{-1}\dot{\Psi}(\theta)^T \geq 0 ,$$

ειναι δηλαδη θετικα ορισμενος.

(3.165) Παραδειγμα: Εστω $X_1,\dots,X_n$ α.ι. $N(\mu,\sigma^2)$, $\underline{\theta}=(\mu,\sigma^2)\in\Theta=\mathbb{R}\times(0,+\infty)$.

(α) Βρητε τον πινακα πληροφοριας της κατανομης $I(\theta)$ και τον πινακα πληροφοριας του δειγματος $I_{\underline{X}_n}(\theta)$.

(β) Επαληθευσατε την (3.164) για $d_n^{(1)}(\underline{X}_n)=\bar{X}_n$, $d_n^{(2)}(\underline{X}_n)=\hat{\sigma}_n^2$.

(γ) Επαληθευσατε την (3.164) για $d_n(\underline{X}_n)=S_n^2$ (βλ. και (3.159).).

Απαντ. (α) $l_n(\underline{\theta}|\underline{X}_n) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^n(X_i-\mu)^2 \Rightarrow$

$$\Rightarrow \frac{\partial^2}{\partial\mu^2}l_n(\theta|\underline{X}_n) = -\frac{n}{\sigma^2}, \quad \frac{\partial^2}{\partial\mu\,\partial\sigma^2}l_n(\underline{\theta}|\underline{X}_n) = -\frac{n(\bar{X}_n-\mu)}{\sigma^4},$$

$$\frac{\partial^2}{\partial(\sigma^2)^2}l_n(\underline{\theta}|\underline{X}_n) = \frac{n}{2\sigma^4} - \frac{1}{\sigma^6}\sum_{i=1}^n(X_i-\mu)^2 .$$

Αρα,

$$I_{\underline{X}_n}(\underline{\theta}) = \begin{pmatrix} \frac{n}{\sigma^2} & 0 \\ 0 & \frac{n}{2\sigma^4}\end{pmatrix} = n\,I(\underline{\theta}), \quad [I_{\underline{X}_n}(\underline{\theta})]^{-1} = \begin{pmatrix} \frac{\sigma^2}{n} & 0 \\ 0 & \frac{2\sigma^4}{n}\end{pmatrix} = \frac{1}{n}[I(\underline{\theta})]^{-1}.$$

(β) $\psi_1(\underline{\theta}) = E_{\underline{\theta}}\bar{X}_n = \mu = q_1(\underline{\theta})$, $\psi_2(\underline{\theta}) = E_{\underline{\theta}}(\hat{\sigma}_n^2) = \frac{n-1}{n}\sigma^2 = \frac{n-1}{n}q_2(\underline{\theta})$.

Αρα,

$$\dot{\Psi}(\underline{\theta}) = \begin{pmatrix} 1 & 0 \\ 0 & \frac{n-1}{n}\end{pmatrix}.$$

Επισης, $\mathcal{D}_{\underline{\theta}}(\bar{X}_n) = \frac{\sigma^2}{n}$, $\mathcal{D}_{\underline{\theta}}(\hat{\sigma}_n^2) = \frac{2(n-1)}{n^2}\sigma^4$ (βλ. (3.123)),
$\operatorname{cov}_{\underline{\theta}}(\bar{X}_n,\hat{\sigma}_n^2)=0$ (βλ. (1.62)). Αρα:

$$\mathcal{D}_{\underline{\theta}}(\underline{d}_n) = \begin{pmatrix} \frac{\sigma^2}{n} & 0 \\ 0 & \frac{2(n-1)\sigma^4}{n^2}\end{pmatrix}.$$ Εχουμε, λοιπον, οτι:

$$\mathcal{D} - \dot{\Psi}I_{\underline{X}_n}^{-1}\dot{\Psi}^T = \begin{pmatrix} 0 & 0 \\ 0 & \frac{2(n-1)\sigma^4}{n^3}\end{pmatrix} \geq 0 ,$$

δηλαδη, μη αρνητικα ορισμενος.

(γ) Εδω, $r=1$ με $\psi_1(\underline{\theta}) = \psi(\underline{\theta}) = E_{\underline{\theta}}(S_n^2) = \sigma^2 = q(\underline{\theta})$, και αρα $\dot{\psi}(\underline{\theta}) = (0, 1)$. Εχουμε, λοιπον, $\dot{\psi} I_{\underline{X}_n}^{-1} \dot{\psi}^T = \frac{2\sigma^4}{n}$ και απο την (3.123), $\mathcal{D}_{\underline{\theta}}(S_n^2) = \frac{2\sigma^4}{n-1} > \frac{2\sigma^4}{n} = \dot{\psi} I_{\underline{X}_n}^{-1} \dot{\psi}^T$.

Μεχρι τωρα, κριναμε την αποδοση μιας εκτιμητριας $d_n(\underline{X}_n)$ της $q(\theta)$, $\theta \in \Theta$, βασιζομενοι μονο στη μεση της τιμη $E_\theta\{d_n(X_n)\}$ και τη διασπορα της $\mathcal{D}_\theta\{d_n(\underline{X}_n)\}$. Μεσω δε, κυριως, των (3.117), (3.137) και (3.152), δειξαμε οτι ειναι δυνατον με την γνωση μονο της μεσης τιμης και διασπορας της d_n να ελεγξουμε την ακριβεια, τη συνεπεια και την ταχυτητα με την οποια η d_n πλησιαζει την $q(\theta)$, καθως επισης και να συγκρινουμε την καταλληλοτητα της ως προς αλλες εκτιμητριες.
Για μια πληρεστερη ομως γνωση της αποδοσης της d_n θα μας χρειαζοταν η γνωση της κατανομης της. Αυτο ομως ειναι εν γενει πολυ δυσκολο ή και αδυνατο. Ειναι ομως συνηθως εφικτο να βρουμε την ασυμπτωτικη κατανομη της εκτιμητριας μας, αν δε η ταχυτητα συγκλισης προς αυτην την ασυμπτωτικη κατανομη ειναι μεγαλη, η γνωση της ειναι χρησιμη και για την συμπεριφορα της εκτιμητριας d_n για "μετρια" μεγεθη δειγματος.

Θα δειξουμε τωρα, ευρισικα, οτι εν γενει οι εμπ – κατω απο σχετικα ασθενεις συνθηκες ομαλοτητος των μοντελων μας – ειναι ασυμπτωτικα κανονικες.
Συγκεκριμμενα:

(3.166) $\quad n^{1/2-\delta}(\hat{\theta}_n - \theta) \xrightarrow[n\to\infty]{P_\theta} 0 \quad, \theta \in \Theta, \ \forall \delta > 0,$

(3.167) $\quad \sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow[n\to\infty]{d} N\left(0, \frac{1}{I(\theta)}\right), \quad \theta \in \Theta.$

Εχουμε: $0 = l_n'(\hat{\theta}_n) \approx l_n'(\theta) + (\hat{\theta}_n - \theta)\, l_n''(\theta) \Rightarrow$

$\Rightarrow \sqrt{n}(\hat{\theta}_n - \theta) \approx \frac{-1}{\frac{1}{n} l_n''(\theta)} \frac{1}{\sqrt{n}} l_n'(\theta) = -\frac{1}{\frac{1}{n} l_n''(\theta)} \cdot \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{\partial}{\partial\theta} \log f(X_i|\theta).$

Τωρα, απο την (3.146), εχουμε οτι $-\frac{1}{n} l_n''(\theta) \xrightarrow[n\to\infty]{P} I(\theta)$, $\theta \in \Theta$,
και αφο τις (1.138), (1.341) και το Κ.Ο.Θ. εχουμε οτι:
$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{\partial}{\partial\theta} \log f(X_i|\theta) \xrightarrow[n\to\infty]{d} N(0, I(\theta))$.
Αρα αφο το Θεωρημα του Slutsky εχουμε οτι:
$\sqrt{n}\,(\hat{\theta}_n - \theta) \xrightarrow[n\to\infty]{d} \frac{1}{I(\theta)} \cdot N(0, I(\theta)) \overset{d}{=} N\left(0, \frac{1}{I(\theta)}\right)$, $\theta \in \Theta$.
Παλι αφ' το Θεωρημα του Slutsky εχουμε οτι: $\forall \delta > 0$,
$n^{1/2-\delta}(\hat{\theta}_n - \theta) = \frac{1}{n^{\delta}} \sqrt{n}\,(\hat{\theta}_n - \theta) \xrightarrow[n\to\infty]{d} 0 . N(0, I(\theta)^{-1}) = 0$;
και αρα αφ' την (1.848), $n^{1/2-\delta}(\hat{\theta}_n - \theta) \xrightarrow[n\to\infty]{P} 0$.

Σημειωνουμε εδω, οτι ασυμπτωτικες κατανομες και ταχυτητες συγκλισεως σαν των (3.127) και (3.128) ειναι η εξαιρεση —παρατηρουμε οτι η $f(\cdot|\theta)$ εκει, δεν διαφοριζεται ως προς θ και δεν υπαρχει η πληροφορια κατα Fisher $I(\theta)$ της $U(0,\theta)$— κατα κανονα περιμενουμε αφο μια καλη εκτιμητρια $d_n(\underline{X}_n)$ της θ να εχη μια ασυμπτωτικα κανονικη κατανομη, δηλαδη,

$$(3.168) \quad \sqrt{n}\left(d_n(\underline{X}_n) - \theta\right) \xrightarrow[n\to\infty]{d} N(0, C^2(\theta)), \quad \theta \in \Theta.$$

Σ' αυτην την περιπτωση η ασυμπτωτικη αποδοση της d_n οριζεται ως:

$$(3.169) \quad e_{\infty}(\theta|d_n) := \frac{1}{I(\theta)\, C^2(\theta)}, \quad \theta \in \Theta.$$

Σημειωνουμε εδω οτι ενω αφο την (3.152), λογω της (3.152), θα περιμεναμε να εχουμε για μια ακριβη εκτιμητρια d_n οτι
$$1 \geq e(\theta|d_n) = \frac{1}{I(\theta)\, n\, \mathcal{D}_\theta(d_n)} \xrightarrow[n\to\infty]{} \frac{1}{I(\theta)\, C^2(\theta)} = e_{\infty}(\theta|d_n)$$
και αρα και
$e_{\infty}(\theta|d_n) \leq 1$, εν τουτοις αυτο δεν συμβαινει παντοτε.
Παντως η υπεραποδοση ($e_{\infty} > 1$) απαντιεται μονο σε παθολογικες περιπτωσεις και μια εκτιμητρια d_n για την οποια ισχυει η (3.168) καλειται ασυμπτωτικα πληρως αποδοτικη αν επι πλεον

$e_\infty(\theta|d_n) = 1 \quad \forall\, \theta \in \Theta$. Μια τετοια εκτιμητρια λεγεται επισης <u>βελτιστη</u> <u>ασυμπτωτικα</u> <u>κανονικη</u> (Β.Α.Κ.).

Παρατηρουμε οτι η (3.167) οτι οι εμπ ειναι εν γενει Β.Α.Κ..

Επισης επειδη $\widehat{q(\theta)} = q(\hat{\theta}_n)$, αν υπαρχει η $q'(\theta)$ τοτε, απο το (1.95) και την (1.168), εχουμε – εν γενει – οτι:

(1.170) $\sqrt{n}\,(\widehat{q(\theta)} - q(\theta)) \xrightarrow[n\to\infty]{d} N(0, C_q^2(\theta))$, $\theta \in \Theta$,

οπου, $C_q^2(\theta) = [q'(\theta)]^2 C^2(\theta)$, και αρα αν η $\hat{\theta}_n$ ειναι ΒΑΚ τοτε και η $q(\hat{\theta}_n)$ ειναι ΒΑΚ. Η συνεπεια των εκτιμητριων ειναι φυσικα αμεση απο την (1.170) και το Θεωρημα του Slutsky.

(1.171) <u>Παραδειγμα</u>. Εστω $X_1, \dots, X_n$ α.δ. Bernoulli(p), $p \in \Theta = [0,1]$.

(α) Δειξτε οτι η εμπ και ΟΑΕΕΔ $\bar{X}_n$ της p ειναι ΒΑΚ.

) Δειξτε οτι η $\widehat{q(p)} = \bar{X}_n(1-\bar{X}_n)$ της $q(p) = p(1-p)$ και ΟΕΕΕΔ $d_n(\underline{X}_n) = \frac{n}{n-1}\bar{X}_n(1-\bar{X}_n)$ της $q(p)$, ειναι Β.Α.Κ, για $p \neq 1/2$.

<u>Απαντ.</u> (α) Απο το ΚΟΘ εχουμε οτι:

$$\sqrt{n}\,(\bar{X}_n - p) \xrightarrow[n\to\infty]{d} N(0, p(1-p)) = N\left(0, \frac{1}{I(p)}\right) \quad - \text{βλ. } (3.156\alpha).$$

(β) Απο το (1.95) και το μερος (α), εφοσον $q'(p) = 1 - 2p$, εχουμε οτι:

$$\sqrt{n}\left[\bar{X}_n(1-\bar{X}_n) - p(1-p)\right] \xrightarrow[n\to\infty]{d} N(0, (1-2p)^2 p(1-p))$$

εφοσον $p \neq 1/2$ (ή των 0, 1 φυσικα). Αν $p = 1/2$ εχουμε,

$$q(\hat{p}_n) = q(p) + (\hat{p}_n - p)\,q'(p) + \frac{1}{2}(\hat{p}_n - p)^2 q''(p)$$

$$= q(\tfrac{1}{2}) - (\hat{p}_n - \tfrac{1}{2})^2 \Rightarrow$$

$$\Rightarrow n\left[\frac{1}{4} - \bar{X}_n(1-\bar{X}_n)\right] = \left[\sqrt{n}(\bar{X}_n - \tfrac{1}{2})\right]^2 \xrightarrow[n\to\infty]{d} \left[N(0, \tfrac{1}{4})\right]^2 \Rightarrow$$

$$\Rightarrow 2n\left[\frac{1}{4} - \bar{X}_n(1-\bar{X}_n)\right] \xrightarrow[n\to\infty]{d} [N(0,1)]^2 = \chi_1^2 \quad \text{αν } p = 1/2,$$

οπου καναμε χρηση του (1.94γ).

Παρομοια, εφοσον, $d_n(\underline{X}_n) = \frac{n}{n-1} q(\bar{X}_n)$ εχουμε οτι

$$\sqrt{n}\,(d_n(\underline{X}_n) - q(p)) = \sqrt{n}\left[\frac{n}{n-1} q(\bar{X}_n) - q(p)\right] =$$

$$= \sqrt{n}\,(q(\bar{X}_n) - q(p)) + \frac{\sqrt{n}}{n-1} q(\bar{X}_n) \xrightarrow[n\to\infty]{d} N\left(0, \frac{(1-2p)^2}{I(p)}\right) + 0 . q(p)$$

$\stackrel{d}{=} N\left(0, \frac{(1-2p)^2}{I(p)}\right)$ εφοσον $p \neq 1/2$, οπου καναμε χρηση του Θεωρηματος του Slutsky. Με τον ιδιο τροπο μπορουμε να δουμε οτι η d_n εχει παλι την ιδια ασυμπτωτικη κατανομη με την $q(\hat{p}_n)$, αν $p=1/2$.

(1.172) <u>Παράδειγμα.</u> Έστω $X_1,\dots,X_n$ α.ι. $N(\mu,\sigma^2)$, $\underline{\theta}=(\mu,\sigma^2)\in$
$\in \Theta = \mathbb{R}\times(0,+\infty)$. Δείξτε ότι,

(α) η $\bar{X}_n$ είναι ΒΑΚ – και άρα και συνεπής – για την μ.

(β) η $\hat{X}_n$ είναι συνεπής και ασυμπτωτικά κανονική αλλά όχι ΒΑΚ για την μ.

(γ) η $\hat{\sigma}_n^2$ και η S_n^2 είναι ΒΑΚ για την σ^2.

<u>Απαντ.</u> (α) Άμεσος συνέπεια του ΚΟΘ και του ότι $I(\mu)=\frac{1}{\sigma^2}$ — βλ. (3.159) και (3.165).

(β) Από την (3.135),

$$\sqrt{n}\,(\hat{X}_n-\mu)\xrightarrow[n\to\infty]{d} N\left(0,\frac{\pi\sigma^2}{2}\right)\quad \forall\,\theta\in\Theta.$$

Άρα, $$e_\infty(\underline{\theta}\,|\,\hat{X}_n)=\frac{1}{\frac{1}{\sigma^2}\cdot\frac{\pi\sigma^2}{2}}=\frac{2}{\pi}\approx 0{,}636<1$$

(γ) Από την (3.159) και (3.165), $I(\sigma^2)=\frac{1}{2\sigma^4}$. Επίσης από το (1.97α) έχουμε ότι:

$$\sqrt{n}\,(S_n^2-\sigma^2)\xrightarrow[n\to\infty]{d} N(0,\mu_4-\sigma^4),$$

όπου $\mu_4:=E_{\underline{\theta}}(X_1-\mu)^4=\sigma^4 E_{\underline{\theta}}\left(\frac{X_1-\mu}{\sigma}\right)^4=\sigma^4 E[N(0,1)]^4=3\sigma^4$,

άρα, $\sqrt{n}\,(S_n^2-\sigma^2)\xrightarrow[n\to\infty]{d} N(0,2\sigma^4)=N\left(0,\frac{1}{I(\sigma^2)}\right)$.

Τότε, εφόσον, $\hat{\sigma}_n^2=\frac{n-1}{n}S_n^2$, έχουμε,

$$\sqrt{n}\,(\hat{\sigma}_n^2-\sigma^2)=\sqrt{n}\,(S_n^2-\sigma^2)-\frac{1}{n}S_n^2,$$

και άρα, εφόσον και $S_n^2-\sigma^2=\frac{1}{\sqrt{n}}\sqrt{n}(S_n^2-\sigma^2)\xrightarrow[n\to\infty]{d} 0.N(0,2\sigma^4)=0$
$\Rightarrow S_n^2\xrightarrow[n\to\infty]{P}\sigma^2$, έχουμε από το Θεώρημα του Slutsky ότι:

$$\sqrt{n}\,(\hat{\sigma}_n^2-\sigma^2)\xrightarrow[n\to\infty]{d} N(0,2\sigma^4)-0.\sigma^2=N(0,2\sigma^4).$$

(3.173) <u>Παράδειγμα.</u> Έστω $X_1,\dots,X_n$ α.ι. $\mathcal{E}(\lambda)$, $\lambda\in\Theta=(0,+\infty)$.
Δείξτε ότι η εμ.π. $\hat{\lambda}_n=(\bar{X}_n)^{-1}$ και η ΟΑΕΕΔ $d_n(\underline{X}_n)=$
$=(1-\frac{1}{n})(\bar{X}_n)^{-1}$ της λ είναι ΒΑΚ. (βλ. και (3.154).).

<u>Απαντ.</u> Από το ΚΟΘ έχουμε ότι:

$$\sqrt{n}\,\left(\bar{X}_n-\frac{1}{\lambda}\right)\xrightarrow[n\to\infty]{d} N\left(0,\frac{1}{\lambda^2}\right)\quad\forall\,\lambda>0.$$

Αρα, $\sqrt{n}(\hat{\lambda}_n - \lambda) = \sqrt{n}\left(\frac{1}{\bar{X}_n} - \lambda\right) = -\frac{\lambda}{\bar{X}_n}\sqrt{n}\left(\bar{X}_n - \frac{1}{\lambda}\right)$.
Τωρα, απο τον ΝΜΑ $\bar{X}_n \xrightarrow[n\to\infty]{P} \frac{1}{\lambda}$ και αρα απ' το Θεωρημα του Slutsky: $\frac{\lambda}{\bar{X}_n} \xrightarrow[n\to\infty]{P} \lambda^2$. Αρα, απο Θεωρημα του Slutsky παλι,

$$\sqrt{n}(\hat{\lambda}_n - \lambda) \xrightarrow[n\to\infty]{d} -\lambda^2 N\left(0, \frac{1}{\lambda^2}\right) \stackrel{d}{=} -N(0, \lambda^2) \stackrel{d}{=} N(0, \lambda^2)$$

και αρα $e_\infty(\lambda | \hat{\lambda}_n) = 1$, εφοσον $I(\lambda) = \lambda^{-2}$ –βλ. (3.154).

Το οτι και η d_n εχει την ιδια ασυμπτωτικη κατανομη με την $\hat{\lambda}_n$ επεται παλι απ' το Θεωρημα του Slutsky ως εξης:

$$\sqrt{n}(d_n - \lambda) = \sqrt{n}(\hat{\lambda}_n - \lambda) - \frac{1}{n}\hat{\lambda}_n \xrightarrow[n\to\infty]{d} N(0, \lambda^2) - 0.\lambda \stackrel{d}{=}$$
$$\stackrel{d}{=} N(0, \lambda^2) = N\left(0, \frac{1}{I(\lambda)}\right).$$

(3.174) <u>Ασκηση</u>. Δειξτε οτι καθε μια απο τις εκτιμητριες της ασκησης (3.158) ειναι, για τις αντιστοιχες παραμετρους:

(α) ΒΑΚ

(β) συνεπεις με ταχυτητα $n^{1/2-\delta}$ $\forall \delta > 0$.

(3.175) <u>Ασκηση</u>. Δειξτε οτι η Εμ.Π. και η ΟΑΕΕΔ της (3.160) ειναι ΒΑΚ και συνεπεις με ταχυτητα $n^{1/2-\delta}$ $\forall \delta > 0$.

3.5. ΠΕΡΙΟΧΕΣ ΕΜΠΙΣΤΟΣΥΝΗΣ ΒΑΣΙΣΜΕΝΕΣ ΣΕ ΕΚΤΙΜΗΤΡΙΕΣ.

Εστω στοχαστικο δειγμα $\underline{X}_n = (X_1, \ldots, X_n)$ α.ι. $F(\cdot|\theta)$, $\theta \in \Theta$. Μια εκτιμητρια $d_n(\underline{X}_n)$ απ' αυτες που αναπτυξαμε προηγουμενως, εφοσον ειναι καλη — με τις εννοιες που ορισαμε πριν — μπορουμε να περιμενουμε οτι θα ειναι κοντα στην αγνωστη αληθη τιμη της $q(\theta)$, $\theta \in \Theta$, τουλαχιστον για "μεγαλα" δειγματικα μεγεθη n, δηλαδη, $\forall \theta \in \Theta$, $\forall \varepsilon > 0$:

$$1 \geq P_\theta\{ d_n(\underline{X}_n) - \varepsilon \leq q(\theta) \leq d_n(\underline{X}_n) + \varepsilon \} \geq 1 - \frac{n\mathcal{D}_\theta(d_n)[1 + \bar{b}(\theta, d_n)]}{n\,\varepsilon^2} \xrightarrow[n\to\infty]{} 1,$$

εφοσον τουλαχιστον η d_n ειναι ακριβης και ασυμπτωτικα αμεροληπτη — και αρα και συνεπης (με ταχυτητα συγκλισης περιπου $n^{1/2}$).

Εχουμε, δηλαδη, οτι καθως το $n \to \infty$ $\forall \varepsilon > 0$ το στοχαστικο διαστημα, μηκους 2ε, $[d_n(\underline{X}_n) - \varepsilon, d_n(\underline{X}_n) + \varepsilon]$ περιεχει την αγνωστη αληθη τιμη της παραμετρικης συναρτησης $q(\theta)$ με "μεγαλη" P_θ - πιθανοτητα, η οποια μαλιστα τεινει στο 1, $\forall \theta \in \Theta$.

Δεν γνωριζουμε ομως, για πεπερασμενο n, αυτην την πιθανοτητα, ουτε το ε, ουτε την σχεση που εχουν αυτα τα δυο. Για κατι τετοιο θα χρειαζοταν να γνωριζουμε την κατανομη της $d_n(\underline{X}_n)$ — δεν αρκει ο προσδιορισμος της συμπεριφορας της $d_n(\underline{X}_n)$ μεσω της μεσης τιμης και διασπορας της μονο.

Η θεωρησης της ασυμπτωτικης κατανομης — συνηθως κανονικης — των εκτιμητριων μας, με την οποια ασχοληθηκαμε στο τελος της προηγουμενης ενοτητας, ηταν ακριβως μια τετοια προσπαθεια καθορισμου της ακριβειας της εκτιμητριας $d_n(\underline{X}_n)$, μετρουμενης μεσω του ε και της σχεσης του με την πιθανοτητα $P_\theta(d_n - \varepsilon \leq q(\theta) \leq d_n + \varepsilon)$. Συγκεκριμμενα, εστω οτι

$$\sqrt{n}\,[d_n(\underline{X}_n) - q(\theta)] \xrightarrow[n\to\infty]{d} N(0, c_q^2(\theta)), \quad \theta \in \Theta,$$

τοτε, θετοντας

$$(3.176) \quad z(1-\alpha) := \Phi^{-1}(1-\alpha), \quad \alpha \in [0,1],$$

εχουμε, για προκαθωρισμενο $\alpha \in [0,1]$, οτι:

$$P_\theta\left\{ d_n(\underline{X}_n) - \frac{C_q(\theta) Z(1-\alpha/2)}{\sqrt{n}} \leq q(\theta) \leq d_n(\underline{X}_n) + \frac{C_q(\theta) Z(1-\alpha/2)}{\sqrt{n}} \right\} =$$

$$= P_\theta\left\{ -Z(1-\alpha/2) \leq \frac{\sqrt{n}\,[d_n(\underline{X}_n) - q(\theta)]}{C_q(\theta)} \leq Z(1-\alpha/2) \right\} \approx$$

$$\approx \Phi(Z(1-\alpha/2)) - \Phi(-Z(1-\alpha/2)) = 2\Phi(Z(1-\alpha/2)) - 1 = 2(1-\alpha/2) - 1 = 1-\alpha, \ \forall \theta \in \Theta.$$

Δηλαδη, με προκαθωρισμενη – αν θελουμε – πιθανοτητα (P_θ) $1-\alpha$ περιπου (διοτι η κατανομη $N(0, C_q^2(\theta))$ ειναι ασυμπτωτικη μονο), $\forall \theta \in \Theta$ – δηλαδη οποιαδηποτε και αν ειναι η αληθης αγνωστη τιμη της παραμετρου θ – το στοχαστικο διαστημα, μηκους $\frac{2C_q(\theta)Z(1-\alpha/2)}{\sqrt{n}}$,

$$\left[d_n(\underline{X}_n) - \frac{C_q(\theta)Z(1-\alpha/2)}{\sqrt{n}},\ d_n(\underline{X}_n) + \frac{C_q(\theta)Z(1-\alpha/2)}{\sqrt{n}} \right]$$ θα περιεχη την $q(\theta)$.

Επειδη, εν γενει, η $C_q(\theta)$ – και αρα το στοχαστικο διαστημα – εξαρταται απο την αγνωστη παραμετρο θ, αν χρειασθη, μπορουμε να εκτιμησουμε την θ με μια συνεπη εκτιμητρια της $\hat{\theta}$, και το $C_q(\theta)$ με την $C_q(\hat{\theta})$, και τοτε – εν γενει – η εμπιστοσυνη μας στο στοχαστικο διαστημα παραμενει περιπου στο ιδιο επιπεδο $1-\alpha$.

Επισης, παρατηρουμε οτι οσο μεγαλυτερη ειναι η ασυμπτωτικη αποδοση της εκτιμητριας d_n τοσο η $C_q(\theta)$ πλησιαζει την – στην ουσια – ελαχιστη τιμη της $\frac{[q'(\theta)]^2}{I(\theta)}$, και αρα το στοχαστικο διαστημα το ελαχιστο δυνατο μηκος του – δηλαδη, την μεγιστη δυνατη ακριβεια του – για το προκαθωρισμενο επιπεδο εμπιστοσυνης $1-\alpha$.

(3.177) <u>Ορισμος</u>. Ενα στοχαστικο διαστημα $[L_n(\underline{X}_n), U_n(\underline{X}_n)]$, βασισμενο σε δυο στατιστικες συναρτησεις L_n, U_n, τετοιες ωστε $P_\theta(L_n \leq U_n) = 1$, καλειται <u>διαστημα εμπιστοσυνης</u> <u>επιπεδου $1-\alpha$</u> <u>για την $q(\theta)$</u> αν

$$P_\theta\left(L_n(\underline{X}_n) \leq q(\theta) \leq U_n(\underline{X}_n) \right) \geq 1-\alpha \quad \forall \theta \in \Theta,\ \alpha \in [0,1].$$

Η πιθανοτητα στα αριστερα της ανισοτητος λεγεται <u>πιθανοτητα καλυψης</u> της $q(\theta)$ απο το (στοχαστικο) διαστημα εμπιστοσυνης $[L_n, U_n]$, και η μικροτερη της τιμη πανω στο Θ, καλειται συντελεστης εμπιστοσυνης του $[L_n, U_n]$.

Θα ασχοληθουμε εδω, κυριως, με περιπτωσεις που οι κατανομες των $L_n(\underline{X}_n) - q(\theta)$, $U_n(\underline{X}_n) - q(\theta)$ ή των $L_n(\underline{X}_n)/q(\theta)$, $U_n(\underline{X}_n)/q(\theta)$ ειναι ανεξαρτητες της θ. Σ' αυτες τις περιπτωσεις:

$$P_\theta(L_n \le q(\theta) \le U_n) = P(L_n - q(\theta) \le 0 \le U_n - q(\theta))$$
$$\text{ή} = P(L_n/q(\theta) \le 1 \le U_n/q(\theta)),$$

ειναι, δηλαδη, η πιθανοτητα καλυψης ανεξαρτητη της θ — σταθερη πανω στο Θ — και αρα ταυτιζεται με τον συντελεστη εμπιστοσυνης, ο οποιος τιθεται ισος προς το επιπεδο εμπιστοσυνης $1-\alpha$. Μπορουμε λοιπον να εχουμε, σ' αυτη την περιπτωση,

$$(3.178)\quad P_\theta(L_n(\underline{X}_n) \le q(\theta) \le U_n(\underline{X}_n)) = 1-\alpha \quad \forall \theta \in \Theta,\ \alpha \in [0,1].$$

Αν γνωριζουμε μονο τις ασυμπτωτικες κατανομες των L_n, U_n τα αντιστοιχα διαστηματα εμπιστοσυνης καλουνται ασυμπτωτικα ή προσεγγιστικα διαστηματα εμπιστοσυνης.

Το κριτηριο ακριβειας ενος διαστηματος εμπιστοσυνης που θεωρουμε εδω ειναι το (μικρο) μηκος του $U_n - L_n$ ή μαλλον η μεση τιμη του μηκους αυτου $E_\theta\{U_n - L_n\}$, $\theta \in \Theta$. Ενα διαστημα εμπιστοσυνης $[L_n^0, U_n^0]$ καλειται ομοιομορφως μεγιστης ακριβειας (ΟΜΑ) διαστημα εμπιστοσυνης, αν

$$(3.179)\ P_\theta\{L_n^0(\underline{X}_n) \le q(\theta') \le U_n^0(\underline{X}_n)\} \le P_\theta\{L_n(\underline{X}_n) \le q(\theta') \le U_n(\underline{X}_n)\} \quad \forall \theta' \ne \theta, \forall \theta \in \Theta$$

και για καθε αλλο διαστημα εμπιστοσυνης $[L_n, U_n]$.

Εν γενει, δεν υπαρχουν Ο.Μ.Α. διαστηματα εμπιστοσυνης, οπως δεν υπαρχουν και εκτιμητριες ελαχιστης διασπορας ή κινδυνου, υπαρχουν ομως συνηθως ΟΜΑ αμεροληπτα (ΟΜΑΑ) διαστηματα εμπιστοσυνης, ή ΟΜΑ συμμετρικα (Ο.Μ.Α.Σ.) διαστηματα εμπιστοσυνης (ΔΕ). Καλουμε δε ενα Δ.Ε. $[L_n, U_n]$ αμεροληπτο αν

$$(3.180)\quad P_\theta(L_n \le q(\theta) \le U_n) \geq P_\theta(L_n \le q(\theta') \le U_n) \quad \forall \theta, \theta' \in \Theta,$$

και συμμετρικο αν

(3.181) $\quad P_\theta\left(L_n > q(\theta) \right) = P_\theta\left(U_n < q(\theta) \right) \quad \forall\, \theta \in \Theta$.

Δεν θα αποδειξουμε εδω την υπαρξη ΟΜΑΑ και ΟΜΑΣ διαστηματων εμπιστοσυνης. Εν γενει, παντως, μπορουμε να αναμενουμε οτι ενα διαστημα εμπιστοσυνης το οποιο βασιζεται σε καλη ή καλες — με τις εννοιες που εχουμε ηδη αναπτυξει — εκτιμητριες της $q(\theta)$, θα εχει μεγαλη ακριβεια, δηλαδη μικρο μεσο μηκος, για το προκαθορισμενο επιπεδο εμπιστοσυνης του $1-\alpha$, $\alpha \in [0,1]$.

(3.182) <u>Παραδειγμα</u>. Εστω $X_1, \dots, X_{30}$, α.ι. $\mathcal{E}(\lambda = \frac{1}{\theta})$, λ ή $\theta \in (0, +\infty)$.
(α) Βρητε ενα επιπεδου $1-\alpha = 0{,}90$ διαστημα εμπιστοσυνης για την θ.
(β) Βρητε ενα επιπεδου $1-\alpha = 0{,}90$ διαστημα εμπιστοσυνης για την λ.
<u>Απαντ.</u> (α) Θα βασισουμε κατασκευη του ΔΕ για την θ στην ΕΜΠ και ΟΑΕΕΔ $\bar{X}_n$ της θ. Γνωριζουμε οτι η κατανομη της επαρκους και πληρους στατιστικης $\sum_{i=1}^{n} X_i$ του προβληματος ειναι η $\mathcal{G}(n, \frac{1}{\theta})$ και αρα — βλ. (1.35) — $\frac{2n}{\theta}\bar{X}_n \sim \mathcal{G}(n, \frac{1}{2}) \overset{d}{=} \chi^2_{2n}$. Θετοντας λοιπον, κατ' αναλογιαν της (3.176),

(3.183) $\quad \chi^2_n(1-\alpha) := F^{-1}(1-\alpha)$, $\alpha \in [0,1]$, F η σ.κ. της χ^2_n,

εχουμε οτι για $\alpha_1 + \alpha_2 = \alpha = 0{,}1$,

$$1-\alpha = P_\theta\left(\chi^2_{2n}(\alpha_1) \le \frac{2n}{\theta}\bar{X}_n \le \chi^2_{2n}(1-\alpha_2) \right) =$$
$$= P_\theta\left(\frac{2n\bar{X}_n}{\chi^2_{2n}(1-\alpha_2)} \le \theta \le \frac{2n\bar{X}_n}{\chi^2_{2n}(\alpha_1)} \right) \quad \forall\, \theta \in \Theta = (0, +\infty).$$

Αν λοιπον θελουμε το διαστημα εμπιστοσυνης της θ να ειναι συμμετρικο, δηλαδη $\alpha_1 = \alpha_2 = 0{,}05$, τοτε, το στοχαστικο διαστημα

$$\left[\frac{2n\bar{X}_n}{\chi^2_{2n}(1-\alpha/2)}, \frac{2n\bar{X}_n}{\chi^2_{2n}(\alpha/2)} \right] = \left[\frac{60\bar{X}_{30}}{\chi^2_{60}(0{,}95)}, \frac{60\bar{X}_{30}}{\chi^2_{60}(0{,}05)} \right]$$

ειναι ενα συμμετρικο Δ.Ε. (το οποιο σ' αυτη την περιπτωση συμβαινει να ειναι ΟΜΑΣ και αμεροληπτο). Την τιμη $\chi^2_{60}(0{,}95)$

μπορουμε να τη διαβασουμε απ' ευθειας απο τον Πινακα (1.40) και ειναι $\chi^2_{60}(0,95) = 79,082$. Η τιμη $\chi^2_{60}(0,05)$ δεν ειναι καταχωρημενη στον πινακα της χ^2_n, ισχυει ομως οτι

(3.184) $\quad \chi^2_n(\alpha) = n / f_{\infty,n}(1-\alpha) \quad , \alpha \in [0,1]$,

οπου,

$f_{k,m}(1-\alpha) = F^{-1}(1-\alpha)$, με F την σ.κ. της $\mathcal{F}_{k,m}$.

Εχουμε, λοιπον $\chi^2_{60}(0,05) = \dfrac{60}{f_{\infty,60}(0,95)} = \dfrac{60}{1.39} = 43.165$

Η ισχυς της (3.184) επεται απο το εξης λημμα:

(3.185) $\quad \mathcal{F}_{k,m} \overset{d}{=} \dfrac{\chi^2_k/k}{\chi^2_m/m} \xrightarrow[k\to\infty]{d} \dfrac{1}{\chi^2_m/m}$, δηλαδη, $\mathcal{F}_{\infty,m} \overset{d}{=} \dfrac{m}{\chi^2_m}$,

ειναι δε αυτο το αποτελεσμα μια αμεση συνεπεια του Θεωρηματος του Slutsky, διοτι, αν $X_1, \dots, X_k$ ειναι α.ι. χ^2_1 τοτε $\Sigma_{i=1}^k X_i \sim \chi^2_k$, αρα, μπορουμε να θεωρησουμε τον αριθμητη της (3.185), χ^2_k/k, σαν το μεσο ορο των $X_1, \dots, X_k$, δηλαδη $\chi^2_k/k \overset{d}{=} \frac{1}{k}\Sigma_{i=1}^k X_i \xrightarrow[k\to\infty]{P} EX_1 = 1$, απο τον ΝΜΑ. Αρα, απο αυτο και το Θεωρημα του Slutsky εχουμε την (3.185).

Τοτε, $\alpha = P(\chi^2_n < \chi^2_n(\alpha)) = P\left(\dfrac{n}{\chi^2_n} > \dfrac{n}{\chi^2_n(\alpha)}\right)$
$= 1 - P\left(\mathcal{F}_{\infty,n} \le \dfrac{n}{\chi^2_n(\alpha)}\right)$,

απο την οποια συναγεται η (3.184).

(β) Απο το μερος (α) και το οτι $\theta = \frac{1}{\lambda}$, εχουμε οτι το στοχαστικο διαστημα

$$\left[\frac{\chi^2_{2n}(\alpha/2)}{2n\bar{X}_n}, \frac{\chi^2_{2n}(1-\alpha/2)}{2n\bar{X}_n}\right]$$ ειναι ενα

συμμετρικο, επιπεδου $1-\alpha$, Δ.Ε. για την λ.

(3.186) <u>Ασκηση</u>. Κατασκευασατε ενα επιπεδου $1-\alpha = 0,90$, ΔΕ, για την παραμετρο λ της $\mathcal{E}(\lambda)$, βασιζομενοι στην ΟΑΕΕΔ $\dfrac{n-1}{\Sigma_1^n X_i}$ εκτιμητρια της λ.

(3.187) Ασκηση. Εστω $X_1, \ldots, X_n$ α.ι. $W(\beta, \lambda = \frac{1}{\theta})$, $\theta \in \Theta = (0, +\infty)$, β γνωστο. Βρητε ενα επιπεδου $1-\alpha = 0{,}90$ Δ.Ε. για την θ. Συγκεκριμμενοποιειστε το Δ.Ε. στην περιπτωση που $n=10$, $\beta=2$, $x_1 = 5{,}72$, $x_2 = 4{,}25$, $x_3 = 6.21$, $x_4 = 10{,}5$, $x_5 = 7{,}2$, $x_6 = 8.15$, $x_7 = 4.78$, $x_8 = 9{,}5$, $x_9 = 10{,}11$, $x_{10} = 7.5$.

(3.188) Παραδειγμα: Εστω $X_1, \ldots, X_n$ α.ι. $N(\mu, \sigma^2)$, $\underline{\theta} = (\mu, \sigma^2) \in \Theta = \mathbb{R} \times (0, +\infty)$. Κατασκευαστε ενα διαστημα εμπιστοσυνης, επιπεδου $1-\alpha$, για την $q(\theta) = \mu$, το οποιο να ειναι συμμετρικο.

Απαντ. Θα βασισουμε το ζητουμενο διαστημα εμπιστοσυνης στις ΟΑΕΕΔ $\bar{X}_n$, S_n^2 των μ, σ^2, που ειναι και επαρκεις και πληρεις για την $\underline{\theta}$. Συγκεκριμμενα, εχουμε οτι (βλ. (1.62)),

$$\frac{\sqrt{n}\,(\bar{X}_n - \mu)}{\sigma} \sim N(0,1) \quad \text{και} \quad \frac{1}{\sigma^2} \sum_{i=1}^{n} (X_i - \bar{X}_n)^2 \sim \chi^2_{n-1}$$

και οι δυο ειναι ανεξαρτητες. Αρα, η στατιστικη-t,

$$T_n := \frac{\sqrt{n}\,(\bar{X}_n - \mu)}{S_n} \sim t_{n-1} \quad \text{- ανεξαρτητη των } \mu, \sigma^2.$$

Αρα, θετοντας,

(3.189) $\quad t_n(1-\alpha) = F^{-1}(1-\alpha)$, οπου F η σ.κ. της t_n,

εχουμε,

$$1-\alpha = P_{\underline{\theta}}\left(\frac{\sqrt{n}\,|\bar{X}_n - \mu|}{S_n} \le t_{n-1}(1-\alpha/2) \right) =$$

$$= P_{\underline{\theta}}\left(\bar{X}_n - \frac{S_n t_{n-1}(1-\alpha/2)}{\sqrt{n}} \le \mu \le \bar{X}_n + \frac{S_n t_{n-1}(1-\alpha/2)}{\sqrt{n}} \right),$$

Αρα, το στοχαστικο διαστημα

$$\left[\bar{X}_n - \frac{S_n t_{n-1}(1-\alpha/2)}{\sqrt{n}},\ \bar{X}_n + \frac{S_n t_{n-1}(1-\alpha/2)}{\sqrt{n}} \right]$$ ειναι ενα

συμμετρικο, επιπεδου $1-\alpha$, Δ.Ε. για την μ.

Για $\alpha = 0{,}05$, $n = 30$, για παραδειγμα, απο τον πινακα (1.40) εχουμε οτι $t_{n-1}(1-\alpha/2) = t_{29}(0.975) = 2.045$.

(3.190) Παραδειγμα. Εστω $X_1, \ldots, X_n$ α.ι. Poisson(θ), $\theta \in \Theta = (0, +\infty)$. Βρητε ενα επιπεδου $1-\alpha$ ασυμπτωτικο συμμετρικο Δ.Ε. για την θ.

Απαντ. Απο το ΚΟΘ, $\frac{\sqrt{n}\,(\bar{X}_n-\theta)}{\sqrt{\theta}} \xrightarrow[n\to\infty]{d} N(0,1)$ – ανεξαρτητη της θ.
Αρα,

$$1-\alpha \approx P_\theta\left\{ \frac{\sqrt{n}\,|\bar{X}_n-\theta|}{\sqrt{\theta}} \le z(1-\alpha/2)\right\} =$$

$$= P_\theta\left\{ \bar{X}_n - \frac{\sqrt{\theta}\,z(1-\alpha/2)}{\sqrt{n}} \le \theta \le \bar{X}_n + \frac{\sqrt{\theta}\,z(1-\alpha/2)}{\sqrt{n}}\right\} \approx$$

$$\approx P_\theta\left\{ \bar{X}_n - \frac{\sqrt{\bar{X}_n}\,z(1-\alpha/2)}{\sqrt{n}} \le \theta \le \bar{X}_n + \frac{\sqrt{\bar{X}_n}\,z(1-\alpha/2)}{\sqrt{n}}\right\}.$$

Αντι για την τελευταια προσεγγιση που οδηγησε στο Δ.Ε.
$\bar{X}_n \pm \frac{\sqrt{\bar{X}_n}\,z(1-\alpha/2)}{\sqrt{n}}$, είναι δυνατη και η εξης τακτικη για την ανεξαρτητοποιηση της διασπορας απο την αγνωστη θ :
άπο το (1.95) εχουμε οτι εφοσον απο το ΚΟΘ :

$$\sqrt{n}\,(\bar{X}_n-\theta) \xrightarrow[n\to\infty]{d} N(0,\theta),$$

τοτε,

$$\sqrt{n}\,(\sqrt{\bar{X}_n}-\sqrt{\theta}) \xrightarrow[n\to\infty]{d} N(0,\tfrac{1}{4}),$$ και αρα το στοχαστικο διαστημα

$$\left[\sqrt{\bar{X}_n} - \frac{z(1-\alpha/2)}{2\sqrt{n}},\ \sqrt{\bar{X}_n} + \frac{z(1-\alpha/2)}{2\sqrt{n}}\right]$$ ειναι ενα, επιπεδου $1-\alpha$,

ασυμπτωτικο διαστημα εμπιστοσυνης για την $\sqrt{\theta}$.
Την τιμη του $z(1-\alpha/2)$ διαβαζουμε απο τον πινακα (1.40), π.χ., για $\alpha=0.05$, $z(0.975)=1.96$.

(3.191) Ασκηση. Εστω $X_1,\dots,X_n$ α.ι. $N(\mu,\sigma^2)$, $\mu\in\mathbb{R}$, $\sigma^2>0$.
Κατασκευαστε ενα συμμετρικο, επιπεδου $1-\alpha$, Δ.Ε.,
(α) για την μ, αν $\sigma^2=\sigma_0^2$ γνωστο (Παρτε $\alpha=0.05$)
(β) για την σ^2, αν $\mu=\mu_0$ γνωστο (Παρτε $\alpha=0.10$, $n=30$)
(γ) για την σ^2 (μ αγνωστο επισης) (Παρτε $\alpha=0.10$, $n=30$).

(3.192) Ασκηση. Εστω $X_1,\dots,X_n$ α.ι. Bernoulli(p), $p\in\Theta=[0,1]$.
(α) Κατασκευασε ενα συμμετρικο, επιπεδου $1-\alpha$, ασυμπτωτικο Δ.Ε.

για την p.

(β) Βρητε ενα συμμετρικο ασυμπτωτικο Δ.Ε., μεγεθους $1-\alpha$, για την $q(p) = 2\,\text{τοξημ}\sqrt{p}$, $p \in \Theta$.

(3.193) Ασκηση. Εστω $X_1, \ldots, X_n$ α.ι.ι. Laplace$(\theta, 1)$, $\theta \in \Theta = \mathbb{R}$. Βρητε ενα συμμετρικο ασυμπτωτικο Δ.Ε., μεγεθους $1-\alpha$, για την θ. (Βλ. (3.87), (3.135)).

3.6. ΕΚΤΙΜΗΤΡΙΕΣ ΕΛΑΧΙΣΤΩΝ ΤΕΤΡΑΓΩΝΩΝ – ΠΑΛΙΝΔΡΟΜΗΣΗ.

Αρχιζουμε αυτη την ενοτητα με τρια παραδειγματα — σε καπως γενικη μορφη — τα οποια υποδηλωνουν τις βασικες εφαρμογες του μοντελου παλινδρομησης

(3.195) Παραδειγμα. Εστω $(\underline{X}_i, Y_i) = (X_{i1}, X_{i2}, \ldots, X_{ik}, Y_i)$ $i=1,\ldots,n$ ανεξαρτητες και ισονομες $f(\underline{x}, y)$, $(\underline{x}, y) \in \mathbb{R}^{k+1}$ παρατηρησεις επι n ασθενων, π.χ., οι μετρησεις $\underline{X}$ αφορουν το βαρος, την ηλικια, την κατανάλωση καπνου καθως και αλλους κοινωνικο-οικονομικους παραγοντες του ασθενους, η δε μετρηση Y – που μας ενδιαφερει κατα κυριο λογο – αφορα την πιεση του ασθενους.

Σκοπος του πειραματος ειναι η ευρεση της σχεσεως που δυνατον να υπαρχη αναμεσα στις $\underline{X}$ και την Y ή η προβλεψη της Y απο τις $\underline{X}$. Τοτε αν την επιτυχια ή μη της προβλεψης $d(\underline{X})$, την μετρουμε με το μεσο τετραγωνικο σφαλμα (ΜΤΣ): $E[Y - d(\underline{X})]^2$, τοτε, απο την (3.8), η καλυτερη (κατα ΜΤΣ) προβλεψη $d(\underline{X})$ της Y ειναι η:

$$d(\underline{X}) = E(Y|\underline{X}) \equiv \mu(\underline{X}),$$

οπου – βλ. (1.66):

$$\mu(\underline{x}) = E(Y|\underline{X}=\underline{x}) = \int y\, f(y|\underline{x})\, dy \quad , \quad \underline{x} \in \mathbb{R}^k .$$

Η καμπυλη ή εν γενει επιφανεια,

(3.196) $\quad y = \mu(\underline{x}) \quad , \underline{x} \in \mathbb{R}^k,$

καλειται συναρτηση παλινδρομησης, αν δε υποθεσουμε οτι εχει γραμμικη – ως προς τις παραμετρους $\underline{\beta}$ – δομη, δηλαδη,

(3.197) $\quad y = \mu(\underline{x}) \equiv \mu(\underline{x}|\underline{\beta}) = \beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k \quad , \underline{x} \in \mathbb{R}^k,$

τοτε η παλινδρομηση καλειται γραμμικη παλινδρομηση ή γραμμικο μοντελο παλινδρομησης. Εχουμε λοιπον οτι, δεδομενων των $\underline{X}_i = \underline{x}_i$, $i=1,\ldots,n$ οι στοχαστικες μεταβλητες $\varepsilon_i := Y_i - \mu(\underline{x}) \equiv Y_i - (1, \underline{x}^T)\underline{\beta}$, $i=1,\ldots,n$ ειναι α.ι. $F(\cdot|\underline{\eta})$, $\underline{\eta} \in H$. Μπορουμε λοιπον να παραστησουμε το γραμμικο μοντελο ως εξης:

(3.198) $Y_i = \beta_0 + \Sigma_{j=1}^{k} \beta_j x_{ij} + \varepsilon_i \quad , E(\varepsilon_i \mid \underline{X}_i = \underline{x}_i) = 0 , \; i = 1, \ldots, n ,$

με $\varepsilon_1, \ldots, \varepsilon_n$ α.ι. $F(\cdot \mid \underline{\eta})$, $\underline{\eta} \in \mathcal{H}$ — δεδομενου οτι $\underline{X} = \underline{x}$, ή

(3.199) $\underline{Y} = \underline{\underline{X}}\, \underline{\beta} + \underline{\varepsilon} \quad , \; E[\underline{\varepsilon} \mid \underline{\underline{X}} = \underline{\underline{x}}] = \underline{0} \quad ,$

οπου ο $n \times (k+1)$ πινακας των αριθμων x_{ij} — με $x_{i0} \equiv 1$, $i = 1, \ldots, n$ — καλειται πινακας σχεδιασμου. Η παραμετρος του προβληματος ειναι η $\underline{\theta} = (\underline{\beta}, \underline{\eta})$, οπου το τμημα της $\underline{\beta}$ ειναι η προς εκτιμηση παραμετρος που μας ενδιαφερει και το τμημα της $\underline{\eta}$ καλειται ενοχλητικη παραμετρος και ασχολουμεθα με αυτην μονο οσο αυτο μας χρειαζεται για τα συμπερασματα μας για την $\underline{\beta}$.

Η γραμμικοτητα, ως προς $\underline{\beta}$, του μοντελου μας δεν ειναι τοσο περιοριστικη οσο φαινεται εκ πρωτης οψεως, διοτι, εν γενει, η συναρτηση παλινδρομησης $\mu(\underline{x})$ μπορει να προσεγγισθη με βαση καποιο συστημα συναρτησεων $g_1(\underline{x}), g_2(\underline{x}), \ldots$, π.χ., πολυωνυμα, τριγωνομετρικες συναρτησεις κ.τ.λ., δηλαδη,

(3.200) $\mu(\underline{x}) \approx \beta_0 + \Sigma_{j=1}^{k} \beta_j g_j(\underline{x}) \quad , \; \underline{x} \in \mathbb{R}^p ,$

και θετουμε $x_j^* := g_j(\underline{x})$, $j = 1, \ldots, k$, και αρα η (3.198) ή η (3.199) μπορει να θεωρηθη μια προσεγγιση στην πραγματικοτητα.

(3.201) Παραδειγμα. Εστω οτι καποια θεωρια (οικονομικη, της φυσικης κ.τ.λ.) συνδεει τα μεγεθη $\underline{x}$, y (π.χ. επενδυσεις σε τεχνολογικο εξοπλισμο, αποδοση εργαζομενων) μεσω μιας εξισωσεως παλινδρομησης:

(3.202) $y = \mu(\underline{x}) = \beta_0 + \Sigma_{j=1}^{k} \beta_j x_j \; ,$

της οποιας οι συντελεστες $\beta_0, \beta_1, \ldots, \beta_k$ ειναι αγνωστοι και χρειαζεται να προσδιορισθουν. Οπως και στην περιπτωση της (3.200) η γραμμικοτητα — ως προς τις παραμετρους $\underline{\beta}$ — της συναρτησης παλινδρομησης μπορει να προερχεται απο μια καλη προσεγγιση της αληθους $\mu(x)$ ή τα x_j να ειναι τα ιδια συναρτησεις αλλων μεταβλητων, π.χ.,

$$y = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 \text{συν}(t) + \beta_4 e^{-t^2} ,$$

οποτε θετοντες $x_1=t$, $x_2=t^2$, $x_3=\sigma\upsilon\nu t$, $x_4=e^{-t^2}$, επανερχομεθα στο αρχικο μοντελο (3.202).

Τα μεγεθη $\underline{x}$, οι <u>ανεξαρτητες μεταβλητες</u>, εδω, δεν ειναι στοχαστικα, ή μπορουν να θεωρηθουν ως μη στοχαστικα επειδη ειναι δυνατον να μετρηθουν με πολυ μεγαλη ακριβεια. <u>Η εξηρτημενη μεταβλητη</u> Y ομως ειναι στοχαστικη ή μαλλον υπαρχει λαθος $\varepsilon := Y-\mu(\underline{x})$ στη μετρηση Y της $\mu(\cdot)$ στη θεση $\underline{x}$. Εχουμε δηλαδη,

(3.203) $Y_i=\mu(\underline{x}_i)+\varepsilon_i = \beta_0+\beta_1 x_{i1}+\dots+\beta_k x_{ik}+\varepsilon_i$, $i=1,\dots,n$

οπου υποθετουμε οτι, $\varepsilon_1,\dots,\varepsilon_n$ α.ι. με $E\varepsilon_i=0$, $i=1,\dots,n$. Εχουμε, δηλαδη, και εδω μια περιπτωση αναλογη των (3.198) και (3.199) μονο που εδω οι ανεξαρτητες μεταβλητες $\underline{x}$ δεν ειναι στοχαστικες.

(3.204) <u>Παραδειγμα.</u> Μια κλωστοβιομηχανια εχει σκοπο να προμηθευθη μια ποσοτητα μηχανων για την κατασκευη νηματος. Στην αγορα υπαρχουν c ειδη -μαρκες- μηχανων που κανουν γι' αυτη τη δουλεια. Η κλωστοβιομηχανια θελει να μαθη ποσο αποδιδουν οι διαφορες μηχανες στην πραξη – μετα δε, λαμβανοντας υπ' οψιν και τις τιμες τους, θα διαλεξη τον τυπο που θα προμηθευθη. Αναλαμβανουν λοιπον r κλωστουφαντες της βιομηχανιας, να χειριστουν ο καθενας επι μια ωρα τον καθε τυπο μηχανης, και εστω Y_{ij} η ποσοτητα νηματος που παρηγαγε ο $i=1,\dots,r$ χειριστης με την $j=1,\dots,c$ μηχανη. Υποθετουμε οτι

(3.205) $Y_{ij} = \mu_0 + a_i + b_j + \varepsilon_{ij}$, $i=1,\dots,r$, $j=1,\dots,c$,

και ε_{ij} α.ι. με $E\varepsilon_{ij}=0$, $r \geqslant c \geqslant 2$,

(3.206) $\Sigma_{i=1}^{r} a_i = \Sigma_{j=1}^{c} b_j = 0$,

οπου η παραμετρος a_i μετρα την επιδραση του i χειριστου στην παραγωγη του νηματος και η παραμετρος b_j την επιδραση της μηχανης τυπου j στην παραγωγη του νηματος. (Για απλοποιηση του μοντελου δεν θεωρουμε εδω τυχον αλληλεπιδρασεις γ_{ij}.)

Η παραμετρος μ_0 ειναι μια γενικη μεση τιμη, ουτως ωστε οι παραμετροι a_i και b_j μετρουν διαφορες στην αποδοση χειριστων και μηχανων αντιστοιχα (ειναι δηλαδη θετικες ή αρνητικες αν η αντιστοιχη αποδοση ειναι πανω ή κατω της μεσης), και αρα πρεπει να ισχυουν και οι συνθηκες (3.206). Βασει, λοιπον, του μοντελου (3.205,6) μπορουμε να εκτιμησουμε τις αποδοσεις b_j, $j=1,\dots,c$ των διαφορων τυπων μηχανων, απαλλοιφοντας την επιδραση που εχει η ικανοτητα του χειριστου στην στην ποσοτητα του νηματος που παραγεται με την μηχανη $j=1,\dots,c$.

Το μοντελο (3.205,6) καλειται <u>μοντελο αναλυσης διασποράς</u>, διαφερει δε των αλλων, που συναντησαμε προηγουμενως, στο οτι οι ανεξαρτητες μεταβλητες x_{ij} παιρνουν, εδω, μονο τις τιμες 0 ή 1 — δηλαδη, κατατασουν απλως τις παρατηρησεις σε δυο κατηγοριες — και επισης στο οτι εχουμε να λαβουμε εδω υπ' οψιν μας τις συνθηκες (3.206). Παντως, και το μοντελο αναλυσης διασπορας μπορει να τεθη στη μορφη (3.199) (ή (3.198) και (3.203)), ως εξης: θετουμε,

$$\underline{Y}^T \equiv (Y_{11},\dots,Y_{1c},Y_{21},\dots,Y_{2c},\dots,Y_{r1},\dots,Y_{rc}) \in \mathbb{R}^{r.c},$$

$$(3.207)\qquad \underline{\varepsilon}^T \equiv (\varepsilon_{11},\dots,\varepsilon_{1c},\varepsilon_{21},\dots,\varepsilon_{2c},\dots,\varepsilon_{r1},\dots,\varepsilon_{rc}) \in \mathbb{R}^{r.c},$$

$$\underline{\beta}^T \equiv (\mu_0,a_1,\dots,a_r,b_1,\dots,b_c) \in \mathbb{R}^{r+c+1},$$

επειδη δε, $\forall i=1,\dots,r$ και $\forall j=1,\dots,c$ εχουμε:

$$\begin{aligned} Y_{ij} &= \mu_0 + 0.a_1 + \dots + 0.a_{i-1} + 1.a_i + 0.a_{i+1} + \dots + 0.a_r + \\ &\quad + 0.b_1 + \dots + 0.b_{j-1} + 1.b_j + 0.b_{j+1} + \dots + 0.b_c + \varepsilon_{ij} \\ &= (1,0,\dots,0,1,0,\dots,0,0,\dots,0,1,0,\dots,0)\underline{\beta} + \varepsilon_{ij}, \end{aligned}$$

οπου το διανυσμα που πολλαπλασιαζει το $\underline{\beta}$ εχει παντου 0 εκτος των θεσεων 1, $1+i$, $1+r+j$ οπου εχει την τιμη 1, ο πινακας σχεδιασμου $\underline{\underline{X}}$ του μοντελου αναλυσης διασπορας, εχει το διανυσμα αυτο, που πολλαπλασιαζει το $\underline{\beta}$, για $(i-1)c+j$ σειρα. Δηλαδη, ο πινακας $\underline{\underline{X}} \in \mathbb{R}^{r.c} \times \mathbb{R}^{r+c+1}$, εχει την μορφη:

$$(3.208)\qquad \underline{\underline{X}} \equiv \begin{pmatrix} \underline{1} & \underline{1} & \underline{0} & \underline{0} & \cdots & \underline{0} & \underline{\underline{I}} \\ \underline{1} & \underline{0} & \underline{1} & \underline{0} & \cdots & \underline{0} & \underline{\underline{I}} \\ \vdots & \vdots & & \ddots & & \vdots & \vdots \\ \vdots & \vdots & & & & \underline{0} & \vdots \\ \underline{1} & \underline{0} & \underline{0} & \cdots & \cdots & \underline{0}\,\underline{1} & \underline{\underline{I}} \end{pmatrix} \in \mathbb{R}^{r.c} \times \mathbb{R}^{r+c+1},$$

με rc γραμμες και $r+c+1$ στηλες,

οπου, τα διανυσματα-στηλες $\underline{0}$, $\underline{1} \in \mathbb{R}^c$ αποτελουνται απο μονο μηδενικα και μονο μοναδες αντιστοιχως, και οι πινακες $\underline{\underline{I}}$ ειναι οι μοναδιαιοι του $\mathbb{R}^c \times \mathbb{R}^c = \mathbb{R}^{2c}$. Αρα το μοντελο αναλυσης διασπορας παιρνει και αυτο την μορφη (βλ. (3.199)):

$$(3.209)\qquad \underline{Y} = \underline{\underline{X}}\,\underline{\beta} + \underline{\varepsilon}\ , \quad \text{με } \Sigma_{i=1}^{r} a_i = 0 \text{ και } \Sigma_{j=1}^{c} b_j = 0\ ,\ r \geq c \geq 2,$$

οπου, ο πινακας σχεδιασμου $\underline{\underline{X}}$ ειναι ο (3.208) και τα διανυσματα $\underline{Y}$, $\underline{\beta}$, $\underline{\varepsilon}$ διδονται απο τις (3.207), τα δε λαθη $\underline{\varepsilon}$ υποθετουμε οτι ειναι α.ι. και $E\,\underline{\varepsilon} = \underline{0} \in \mathbb{R}^{r.c}$. Οι παραμετροι που μας ενδιαφερουν ειναι οι επιδρασεις των γραμμων a_i, $i=1,\dots,r$, οι επιδρασεις των στηλων b_j, $j=1,\dots,c$ και δευτερευοντως η γενικη μεση τιμη μ_0, δηλαδη, η παραμετρος $\underline{\beta} \in \mathbb{R}^{r+c+1}$. Παρατηρουμε οτι λογω των (3.206), ο πινακας σχεδιασμου $\underline{\underline{X}}$ δεν εχει πληρη διασταση, δηλαδη,

$$(3.210)\qquad \dim(\underline{\underline{X}}) = (r.c) \wedge (r+c+1) - 2 \;\begin{cases} = r+c-1 & \text{αν } (r-1)(c-1) \geq 2 \\ = 2 & \text{αν } r=c=2. \end{cases}$$

Στα ανωτερω παραδειγματα (3.195), (3.201) και (3.204) περιγραψαμε τις βασικες περιπτωσεις στις οποιες εμφανιζεται η αναγκη στατιστικης αναλυσης του <u>κλασσικου γραμμικου μοντελου παλινδρομησης</u>, το οποιο χαρακτηριζεται ως εξης:

δεδομενων των <u>ανεξαρτητων μεταβλητων</u> $x_{ij} \in \mathbb{R}$ $i=1,\dots,n$, $j=1,\dots,k$, τις οποιες, συνηθως, συγκεντρωνουμε στον $n \times k$ <u>πινακα σχεδιασμου</u> $\underline{\underline{X}} \equiv [x_{ij}]$,

οι <u>εξηρτημενες μεταβλητες</u> Y_i, $i=1,\dots,n$ και τα ανεξαρτητα και ισονομα $F(\cdot|\underline{\eta})$, $\underline{\eta} \in \mathcal{H}$ λαθη των μετρησεων ε_i, $i=1,\dots,n$, συνδεονται με τη σχεση:

(3.211) $\quad Y_i = \mu(\underline{x}_i|\underline{\beta}) + \varepsilon_i = \Sigma_{j=1}^{k} x_{ij}\beta_j + \varepsilon_i \quad , i=1,\ldots,n$

ή

$$\underline{Y} = \underline{\underline{X}}\,\underline{\beta} + \underline{\varepsilon} \qquad ,$$

οπου η παραμετρος που μας ενδιαφερει ειναι η $\underline{\beta} \in \mathbb{R}^k$ και με την ενοχλητικη παραμετρο $\underline{\eta} \in \mathcal{H}$ ασχολουμεθα εφοσον και στο βαθμο που χρειαζεται για τον προσδιορισμο της $\underline{\beta}$. Σε πολλες εφαρμογες υποθετουμε οτι $\varepsilon_1,\ldots,\varepsilon_n$ α.ι. $N(0,\sigma^2)$, ή, ισοδυναμα, $\underline{\varepsilon} \sim N_n(\underline{0}, \sigma^2 \underline{\underline{I}})$, με $\sigma^2 \in \mathcal{H} = (0,+\infty)$ την ενοχλητικη παραμετρο του προβληματος.

Η μεθοδος της εκτιμητικης την οποια θα χρησιμοποιησουμε για την εκτιμηση της παραμετρου $\underline{\beta}$ ειναι ανεξαρτητη της κατανομης των λαθων και συνισταται στην ελαχιστοποιηση του δειγματικου αναλογου $\underline{\varepsilon}^T\underline{\varepsilon}/n = \frac{1}{n}\Sigma_{i=1}^{n}\varepsilon_i^2$ του μεσου τετραγωνικου σφαλματος (ΜΤΣ), η θεωρηση του οποιου μας εδωσε την συναρτηση παλινδρομησης της (3.196). Εκτιμουμε, δηλαδη, την παραμετρο που ενδιαφερει $\underline{\beta}$ με την εκτιμητρια $\hat{\underline{\beta}}$ η οποια ικανοποιει την συνθηκη των ελαχιστων τετραγωνων:

(3.212) $\quad \Sigma_{i=1}^{n}\left[Y_i - \mu(\underline{x}_i|\hat{\underline{\beta}})\right]^2 = \min\left\{\Sigma_{i=1}^{n}\left[Y_i - \mu(\underline{x}_i|\underline{\beta})\right]^2, \underline{\beta}\in\mathbb{R}^k\right\}.$

Η μεθοδος των ελαχιστων τετραγωνων, η οποια αναγεται στον Gauss, εχει παντοτε λυση $\hat{\underline{\beta}}$ – την εκτιμητρια ελαχιστων τετραγωνων (Ε.Ε.Τ.).

(3.213) Ασκηση. Δειξτε οτι στην περιπτωση που τα λαθη $\underline{\varepsilon}$ ακολουθουν την $N_n(\underline{0}, \sigma^2\underline{\underline{I}})$, η ΕΕΤ $\hat{\underline{\beta}}$ της παραμετρου $\underline{\beta}$ του μοντελου παλινδρομησης (3.211), ταυτιζεται με την ε.μ.π. της $\underline{\beta}$.

Παρατηρουμε οτι το αθροισμα των τετραγωνων της (3.212), ειναι μια διαφορισιμη κυρτη συναρτηση των $\underline{\beta} = (\beta_1,\ldots,\beta_k)^T \in \mathbb{R}^k$, και αρα μια ΕΕΤ $\hat{\beta} = (\hat{\beta}_1, \ldots, \hat{\beta}_k) \in \mathbb{R}^k$ ειναι μια λυση του συστηματος των <u>κανονικων</u> <u>εξισωσεων</u> :

$$-\frac{1}{2}\frac{\partial}{\partial \beta_{j_0}} \Sigma_{i=1}^{n}\left(Y_i - \Sigma_{j=1}^{k} x_{ij}\beta_j\right)^2 =$$

(3.214) $\Sigma_{i=1}^{n} x_{ij_0}\left(Y_i - \Sigma_{j=1}^{k} x_{ij}\beta_j\right) = 0 \quad , \quad j_0 = 1, \ldots, k \le n .$

(3.215) <u>Προταση.</u> Αν ο $n \times k$ πινακας σχεδιασμου $\underline{\underline{X}}$ εχει διασταση $k \le n$, τοτε οι κανονικες εξισωσεις (3.214) εχουν μοναδικη λυση $\hat{\beta}$ και
$$\hat{\beta} = (\underline{\underline{X}}^T\underline{\underline{X}})^{-1}\underline{\underline{X}}^T\underline{Y} .$$

<u>Αποδ.</u>

(α) $\Sigma_{i=1}^{n}[Y_i - \mu(x_i|\beta)]^2 = \varepsilon^T\varepsilon = (\underline{Y} - \underline{\underline{X}}\beta)^T(\underline{Y} - \underline{\underline{X}}\beta) =$

$= [(\underline{Y} - \underline{\underline{X}}\hat{\beta}) + \underline{\underline{X}}(\hat{\beta}-\beta)]^T[(\underline{Y} - \underline{\underline{X}}\hat{\beta}) + \underline{\underline{X}}(\hat{\beta}-\beta)] =$

$= [(\underline{Y} - \underline{\underline{X}}\hat{\beta})^T + (\hat{\beta}-\beta)^T\underline{\underline{X}}^T][(\underline{Y} - \underline{\underline{X}}\hat{\beta}) + \underline{\underline{X}}(\hat{\beta}-\beta)]$

$= (\underline{Y} - \underline{\underline{X}}\hat{\beta})^T(\underline{Y} - \underline{\underline{X}}\hat{\beta}) + (\underline{Y} - \underline{\underline{X}}\hat{\beta})^T\underline{\underline{X}}(\hat{\beta}-\beta) + (\hat{\beta}-\beta)^T\underline{\underline{X}}^T(\underline{Y} - \underline{\underline{X}}\hat{\beta}) +$

$\quad + (\hat{\beta}-\beta)^T\underline{\underline{X}}^T\underline{\underline{X}}(\hat{\beta}-\beta) =$

$= (\underline{Y} - \underline{\underline{X}}\hat{\beta})^T(\underline{Y} - \underline{\underline{X}}\hat{\beta}) + 2(\hat{\beta}-\beta)[\underline{\underline{X}}^T\underline{Y} - (\underline{\underline{X}}^T\underline{\underline{X}})\hat{\beta}] + [\underline{\underline{X}}(\hat{\beta}-\beta)]^T[\underline{\underline{X}}(\hat{\beta}-\beta)]$

$= (\underline{Y} - \underline{\underline{X}}\hat{\beta})^T(\underline{Y} - \underline{\underline{X}}\hat{\beta}) + [\underline{\underline{X}}(\hat{\beta}-\beta)]^T[\underline{\underline{X}}(\hat{\beta}-\beta)]$

$\geq (\underline{Y} - \underline{\underline{X}}\hat{\beta})^T(\underline{Y} - \underline{\underline{X}}\hat{\beta}) \quad \forall \beta \in \mathbb{R}^k .$

(β) $\nabla_\beta \Sigma_{i=1}^{n}[Y_i - \mu(x_i|\beta)]^2 = \nabla_\beta\{(\underline{Y} - \underline{\underline{X}}\beta)^T(\underline{Y} - \underline{\underline{X}}\beta)\} =$

$= \nabla_\beta\{\beta^T(\underline{\underline{X}}^T\underline{\underline{X}})\beta - 2\underline{Y}^T\underline{\underline{X}}\beta + \underline{Y}^T\underline{Y}\} =$

$= 2[(\underline{\underline{X}}^T\underline{\underline{X}})\beta - \underline{\underline{X}}^T\underline{Y}] = 0 \iff$

(3.216) $(\underline{\underline{X}}^T\underline{\underline{X}})\hat{\beta} = \underline{\underline{X}}^T\underline{Y} \quad \iff$

$\iff \hat{\beta} = (\underline{\underline{X}}^T\underline{\underline{X}})^{-1}\underline{\underline{X}}^T\underline{Y} .$

(3.217) <u>Σημ.</u> Το συστημα (3.216) ειναι ισοδυναμο με το (3.214) και μια λυση του ειναι μια ΕΕΤ ακομη και αν ο $n \times k$ πινακας $\underline{\underline{X}}$ δεν εχει (πληρη) διασταση $k (\le n)$. Ειναι, δηλαδη, οι (3.216) οι <u>κανονικες</u> <u>εξισωσεις</u> των ελαχιστων τετραγωνων. Για παραδειγμα στην περιπτωση του μοντελου αναλυσης διασποράς —

βλ. (3.204) — ο πίνακας σχεδιασμού $\underline{\underline{X}}$ έχει διάσταση $k-2$, όπου, $k = r+c+1$.

(3.218) <u>Πόρισμα</u>: Έστω ότι η διάσταση του $\underline{\underline{X}}$ είναι $k \le n$, τότε η ΕΕΤ $\hat{\underline{\beta}}$ — βλ. (3.215) — έχει:

(α) $E\hat{\underline{\beta}} = \underline{\beta}$,

(β) $\mathcal{D}(\hat{\underline{\beta}}) = \sigma^2 (\underline{\underline{X}}^T \underline{\underline{X}})^{-1}$.

<u>Απόδ.</u> (α) Εφόσον $\underline{Y} = \underline{\underline{X}}\underline{\beta} + \underline{\varepsilon}$ και $E\underline{\varepsilon} = \underline{0}$ έχουμε ότι

$$E\underline{Y} = \underline{\underline{X}}\underline{\beta} \Rightarrow E\hat{\underline{\beta}} = E\{(\underline{\underline{X}}^T\underline{\underline{X}})^{-1}\underline{\underline{X}}^T\underline{Y}\} =$$

$$= (\underline{\underline{X}}^T\underline{\underline{X}})^{-1}\underline{\underline{X}}^T E\underline{Y} = (\underline{\underline{X}}^T\underline{\underline{X}})^{-1}\underline{\underline{X}}^T\underline{\underline{X}}\underline{\beta} = \underline{\beta} .$$

(β) $\mathcal{D}(\hat{\underline{\beta}}) = E\{(\hat{\underline{\beta}} - E\hat{\underline{\beta}})(\hat{\underline{\beta}} - E\hat{\underline{\beta}})^T\} = E\{(\hat{\underline{\beta}} - \underline{\beta})(\hat{\underline{\beta}} - \underline{\beta})^T\}$.

Αλλά,

$$\hat{\underline{\beta}} - \underline{\beta} = (\underline{\underline{X}}^T\underline{\underline{X}})^{-1}\underline{\underline{X}}^T(\underline{Y} - \underline{\underline{X}}\underline{\beta}) \Rightarrow$$

$$\mathcal{D}(\hat{\underline{\beta}}) = E\{(\underline{\underline{X}}^T\underline{\underline{X}})^{-1}\underline{\underline{X}}^T(\underline{Y} - \underline{\underline{X}}\underline{\beta})(\underline{Y} - \underline{\underline{X}}\underline{\beta})^T\underline{\underline{X}}(\underline{\underline{X}}^T\underline{\underline{X}})^{-1}\}$$

$$= (\underline{\underline{X}}^T\underline{\underline{X}})^{-1}\underline{\underline{X}}^T\{E[(\underline{Y} - \underline{\underline{X}}\underline{\beta})(\underline{Y} - \underline{\underline{X}}\underline{\beta})^T]\}\underline{\underline{X}}(\underline{\underline{X}}^T\underline{\underline{X}})^{-1}$$

$$= (\underline{\underline{X}}^T\underline{\underline{X}})^{-1}\underline{\underline{X}}^T E(\underline{\varepsilon}\underline{\varepsilon}^T)\underline{\underline{X}}(\underline{\underline{X}}^T\underline{\underline{X}})^{-1} = (\underline{\underline{X}}^T\underline{\underline{X}})^{-1}\underline{\underline{X}}^T\sigma^2\underline{\underline{I}}\,\underline{\underline{X}}(\underline{\underline{X}}^T\underline{\underline{X}})^{-1}$$

$$= \sigma^2(\underline{\underline{X}}^T\underline{\underline{X}})^{-1} .$$

(3.219) <u>Άσκηση</u>: Έστω ότι η διάσταση του $n\times k$ πίνακα σχεδιασμού $\underline{\underline{X}}$ είναι $k \le n$. Δείξτε ότι, αν $\underline{\varepsilon} \sim N_n(\underline{0}, \sigma^2\underline{\underline{I}})$ τότε:

(α) $\hat{\underline{\beta}} = (\underline{\underline{X}}^T\underline{\underline{X}})^{-1}\underline{\underline{X}}^T Y \sim N_k(\underline{\beta}, \sigma^2(\underline{\underline{X}}^T\underline{\underline{X}})^{-1})$.

(β) $\underline{c}^T\hat{\underline{\beta}} \sim N_1(\underline{c}^T\underline{\beta}, \sigma^2\underline{c}^T(\underline{\underline{X}}^T\underline{\underline{X}})^{-1}\underline{c})$ $\forall\, \underline{c} \in \mathbb{R}^k$ σταθερά.

(3.220) <u>Άσκηση</u>. Έστω ότι η διάσταση του $n\times k$ πίνακα σχεδιασμού $\underline{\underline{X}}$, του μοντέλου (3.211), είναι $k \le n$, και ότι $\underline{\varepsilon} \sim N_n(\underline{0}, \sigma^2\underline{\underline{I}})$.

Δείξτε ότι:

(α) $I(\underline{\beta}) = (\underline{\underline{X}}^T\underline{\underline{X}})/\sigma^2$, ο πίνακας πληροφορίας κατά Fisher — βλ. (3.163) — και συμπεράνατε ότι για την $\hat{\underline{\beta}}$ η (3.164) ισχύει με ισότητα.

(β) $\mathcal{D}(\underline{c}^T\hat{\underline{\beta}}) = \underline{c}^T I(\underline{\beta})^{-1}\underline{c}$ και συμπεράνατε ότι

$\underline{c}^T\hat{\underline{\beta}}$ ειναι η ΟΑΕΕΔ της $\underline{c}^T\underline{\beta}$ $\forall \underline{c}\in\mathbb{R}^k$ σταθερο.

(γ) η $\hat{\beta}_i$ ειναι η ΟΑΕΕΔ της β_i , $i=1,\dots,k$

Σημ. Απο την (3.213) εχουμε οτι, εδω, η $\hat{\underline{\beta}}$ ειναι και η εμπ της $\underline{\beta}$.

Απο το γραμμικο συστημα (3.216) των κανονικων εξισωσεων — ισοδυναμο του (3.214) — βλεπουμε οτι η $\hat{\underline{\beta}}$, ακομη και οταν δεν ειναι μοναδικη — δηλαδη ο πινακας σχεδιασμου $\underline{\underline{X}}$ δεν εχει πληρη διασταση k — ειναι γραμμικη ως προς $\underline{Y}$, π.χ. βλ. (3.215). Εστω, λοιπον, οτι μας ενδιαφερει η εκτιμηση της παραμετρικης συναρτησεως $q(\underline{\beta})=\underline{c}^T\underline{\beta}$, π.χ., της $\beta_1=(1,0,\dots,0)\underline{\beta}$, για καποιο $\underline{c}\in\mathbb{R}^k$. Απο την (3.220), βλεπουμε οτι αν τα λαθη ακολουθουν την κανονικη κατανομη, η γραμμικη ως προς τα $\underline{Y}$ εκτιμητρια $q(\hat{\underline{\beta}})=\underline{c}^T\hat{\underline{\beta}}$, ειναι η ε.μ.π και ΟΑΕΕΔ της $q(\underline{\beta})=\underline{c}^T\underline{\beta}$. Αν ομως τα λαθη του γραμμικου μοντελου παλινδρομησης δεν ειναι κανονικα, τοτε, εν γενει, η $\underline{c}^T\hat{\underline{\beta}}$ δεν ειναι ΟΑΕΕΔ (ουτε, φυσικα, ε.μ.π.). Ισχυει, ομως, οτι η $\underline{c}^T\hat{\underline{\beta}}$ ειναι ΟΑΕΕΔ μεταξυ ολων των γραμμικων ως προς $\underline{Y}$ εκτιμητριων της $\underline{c}^T\underline{\beta}$, ΟΑΓΕΕΔ, οπως δειχνει το ακολουθο θεωρημα.

(3.221) Θεωρημα (Gauss – Markov) Εστω το γραμμικο μοντελο παλινδρομησης (3.211) και εστω $\hat{\underline{\beta}}$ μια ΕΕΤ της $\underline{\beta}$, δηλαδη, $\underline{\underline{X}}^T\underline{\underline{X}}\,\hat{\underline{\beta}}=\underline{\underline{X}}^T\underline{Y}$. Εστω δε $\underline{a}^T\underline{Y}$, $\underline{a}\in\mathbb{R}^n$, μια αμεροληπτη εκτιμητρια της παραμετρικης συναρτησης $\underline{c}^T\underline{\beta}$, $\underline{c}\in\mathbb{R}^k$ σταθερο. Τοτε, $\forall\underline{c}\in\mathbb{R}^k$ σταθερο,

$$\mathcal{D}(\underline{c}^T\hat{\underline{\beta}})\le\mathcal{D}(\underline{a}^T\underline{Y})\quad\forall\underline{a}\in\mathbb{R}^n:\ E(\underline{a}^T\underline{Y})=\underline{c}^T\underline{\beta}\ \forall\underline{\beta}\in\mathbb{R}^k,$$

ειναι, δηλαδη, η $\underline{c}^T\hat{\underline{\beta}}$ ΟΑΓΕΕΔ για την $\underline{c}^T\underline{\beta}$.

Αποδ. Παρατηρουμε οτι, απο τις κανονικες εξισωσεις:

$$\hat{\underline{\beta}}^T\underline{\underline{X}}^T(\underline{Y}-\underline{\underline{X}}\hat{\underline{\beta}})=\hat{\underline{\beta}}^T(\underline{\underline{X}}^T\underline{Y}-\underline{\underline{X}}^T\underline{\underline{X}}\hat{\underline{\beta}})=\hat{\underline{\beta}}^T\underline{0}=0,$$ δηλαδη,

$\underline{\underline{X}}\hat{\underline{\beta}}\perp\underline{Y}-\underline{\underline{X}}\hat{\underline{\beta}}$ και αρα,

$$(3.222)\qquad \underline{Y}=\underline{\underline{X}}\hat{\underline{\beta}}\oplus(\underline{Y}-\underline{\underline{X}}\hat{\underline{\beta}})\ ,$$

είναι, δηλαδή, η $\underline{\underline{X}}\hat{\underline{\beta}} \in \mathbb{R}^n$ η ορθογώνια προβολή του $\underline{Y} \in \mathbb{R}^n$ στον υπόχωρο $V_c(\underline{\underline{X}}) \subseteq \mathbb{R}^k \subseteq \mathbb{R}^n$ που δημιουργείται από τις στήλες του πίνακα σχεδιασμού $\underline{\underline{X}} = (\underline{b}_1, \ldots, \underline{b}_k)$, $\underline{\underline{X}}\hat{\underline{\beta}} = \Sigma_{i=1}^{k} \hat{\beta}_i \underline{b}_i$.

Τώρα, $\forall \underline{a} \in \mathbb{R}^n : E\,\underline{a}^T\underline{Y} = \underline{c}^T\underline{\beta} \;\; \forall \underline{\beta} \in \mathbb{R}^k$, έστω $\underline{a}_0$ η ορθογώνια προβολή του στον $V_c(\underline{\underline{X}})$, τότε, $\underline{a} = \underline{a}_0 \oplus (\underline{a} - \underline{a}_0)$ και $\underline{a} - \underline{a}_0 \perp V_c(\underline{\underline{X}})$, έχουμε δε ότι:

$$\forall \underline{\beta} \in \mathbb{R}^n \quad \underline{c}^T\underline{\beta} = E\,\underline{a}^T\underline{Y} = E\,\underline{a}_0^T\underline{Y} + E(\underline{a} - \underline{a}_0)^T\underline{Y} =$$
$$= E\,\underline{a}_0^T\underline{Y} + (\underline{a} - \underline{a}_0)^T\underline{\underline{X}}\,\underline{\beta} = E\,\underline{a}_0^T\underline{Y},$$
$$\mathcal{D}(\underline{a}^T\underline{Y}) = \underline{a}^T\mathcal{D}(\underline{Y})\underline{a} = \sigma^2\underline{a}^T\underline{\underline{I}}\,\underline{a} = \sigma^2\underline{a}^T\underline{a} \geqslant$$
$$\geqslant \sigma^2\underline{a}_0^T\underline{a}_0 = \mathcal{D}(\underline{a}_0^T\underline{Y}),$$

δηλαδή, αν η $\underline{a}^T\underline{Y}$ είναι αμερόληπτη εκτιμήτρια της $\underline{c}^T\underline{\beta}$, η $\underline{a}_0^T\underline{Y}$ είναι επίσης αμερόληπτη για την $\underline{c}^T\underline{\beta}$ και μάλιστα έχει μικρότερη διασπορά.

Θα δείξουμε τώρα ότι $\forall \underline{a} \in \mathbb{R}^n : E\,\underline{a}^T\underline{Y} = \underline{c}^T\underline{\beta} \;\; \forall \underline{\beta} \in \mathbb{R}^n$, $\underline{a}_0^T\underline{Y} = \underline{c}^T\hat{\underline{\beta}}$ — το οποίο και ολοκληρώνει την απόδειξη:

Παρατηρούμε ότι:

$\underline{c}^T\underline{\beta} = E\,\underline{a}_0^T\underline{Y} = \underline{a}_0^T\underline{\underline{X}}\,\underline{\beta} \;\; \forall \underline{\beta} \in \mathbb{R}^k \Leftrightarrow \underline{a}_0^T\underline{\underline{X}} = \underline{c}^T$, και άρα — χρησιμοποιώντας και την (3.222),

$$\underline{a}_0^T\underline{Y} - \underline{c}^T\hat{\underline{\beta}} = \underline{a}_0^T\left[\underline{\underline{X}}\hat{\underline{\beta}} \oplus (\underline{Y} - \underline{\underline{X}}\hat{\underline{\beta}})\right] - \underline{c}^T\hat{\underline{\beta}} =$$
$$= \underline{a}_0^T\underline{\underline{X}}\hat{\underline{\beta}} - \underline{c}^T\hat{\underline{\beta}} = (\underline{a}_0^T\underline{\underline{X}} - \underline{c}^T)\hat{\underline{\beta}} = 0 \quad .$$

Ένα άμεσο πόρισμα του Θεωρήματος των Gauss-Markov, είναι ότι:

(3.223) $\forall i = 1, \ldots, k$ η $\hat{\beta}_i$ είναι ΟΑΓΕΕΔ της β_i,

αν δε επί πλέον, $\underline{\varepsilon} \sim N_n(\underline{0}, \sigma^2\underline{\underline{I}})$, τότε από την (3.220), η $\hat{\beta}_i$ είναι ΟΑΕΕΔ και ε.μ.π. της β_i, $i = 1, \ldots, k$.

Θα αναλύσουμε τώρα το μοντέλο της <u>απλής γραμμικής παλινδρόμησης</u>, σαν ένα χαρακτηριστικό παράδειγμα γραμμικού μοντέλου και επίλυσης των κανονικών εξισώσεων (3.216) ή (3.214).

(3.224) <u>Παραδειγμα</u>. Θελουμε να διαπιστωσουμε την επιδραση της ποσοτητος x ενος τυπου λιπασματος στην αποδοση y μιας ποικιλιας δημητριακων. Λιπαινουμε λοιπον με αυτο το λιπασμα n ομοια χωραφια ισου εμβαδου, σε επιπεδα $x_1, \dots, x_n$ και εστω $Y_1, \dots, Y_n$ οι αντιστοιχες αποδοσεις αυτων των χωραφιων. Για τη σχεση των x, y υποθετουμε το μοντελο <u>απλης γραμμικης παλινδρομησης</u>:

$$Y_i = \beta_1 + \beta_2 x_i + \varepsilon_i \,, \quad i = 1, \dots, n \,,$$

με $\varepsilon_1, \dots, \varepsilon_n$ α.ι. .

Θα δειξουμε οτι οι ΕΕΤ των β_1, β_2 ειναι οι εξης:

(3.225) $\hat{\beta}_1 = \bar{Y} - \hat{\beta}_2 \bar{x} \,, \quad \hat{\beta}_2 = r\, \hat{\sigma}_y / \hat{\sigma}_x \,,$

οπου, $\hat{\sigma}_x^2 := \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \,, \quad \hat{\sigma}_y^2 := \frac{1}{n} \sum_{i=1}^n (Y_i - \bar{Y}) \,,$

$$r := \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(Y_i - \bar{Y}) \Big/ \sqrt{\hat{\sigma}_x^2 \hat{\sigma}_y^2} \,.$$

Θα υπολογισουμε τις εκτιμητριες μας με δυο τροπους:

(α) Μεσω των (3.214): $\quad \underline{\varepsilon}^T \underline{\varepsilon} \equiv \sum_{i=1}^n (Y_i - \beta_1 - \beta_2 x_i)^2 \,,$

$$\left.\begin{aligned} \frac{\partial}{\partial \beta_1} \underline{\varepsilon}^T \underline{\varepsilon} &= -2 \sum_{i=1}^n (Y_i - \beta_1 - \beta_2 x_i) = 0 \\ \frac{\partial}{\partial \beta_2} \underline{\varepsilon}^T \underline{\varepsilon} &= -2 \sum_{i=1}^n x_i (Y_i - \beta_1 - \beta_2 x_i) = 0 \end{aligned}\right\} \Leftrightarrow$$

$$\Leftrightarrow \left\{\begin{aligned} &\bar{Y} - \hat{\beta}_1 - \hat{\beta}_2 \bar{x} = 0 \\ &\textstyle\sum_{i=1}^n x_i Y_i - n \bar{x} \hat{\beta}_1 - \hat{\beta}_2 \sum_1^n x_i^2 = 0 \end{aligned}\right\} \Leftrightarrow \left\{\begin{aligned} &\hat{\beta}_1 + \bar{x} \hat{\beta}_2 = \bar{Y} \\ &n \bar{x} \hat{\beta}_1 + \left(\textstyle\sum_1^n x_i^2\right) \hat{\beta}_2 = \textstyle\sum_1^n x_i Y_i \end{aligned}\right\} \Leftrightarrow$$

$$\Leftrightarrow \left\{\begin{aligned} &\hat{\beta}_1 + \bar{x} \hat{\beta}_2 = \bar{Y} \\ &\left[\textstyle\sum_1^n (x_i - \bar{x})^2\right] \hat{\beta}_2 = \textstyle\sum_1^n (x_i - \bar{x})(Y_i - \bar{Y}) \end{aligned}\right\} \text{ οι οποιες διδουν τις (3.225).}$$

Σημειωνουμε οτι καναμε χρηση των εξης χρησιμων ταυτοτητων:

(3.226) $\quad \sum_1^n (x_i - \bar{x})^2 = \sum_1^n x_i^2 - n \bar{x}^2 \,,$ και την γενικοτερη,

(3.227) $\quad \sum_1^n (x_i - \bar{x})(y_i - \bar{y}) = \sum_1^n x_i y_i - n \bar{x} \bar{y} \,.$

(β) Μεσω των (3.216): ο $n \times 2$ πινακας σχεδιασμου $\underline{\underline{X}}$ του μοντελου μας εχει ως εξης: $\quad \underline{\underline{X}}^T = \begin{pmatrix} \underline{1}^T \\ \underline{x}^T \end{pmatrix} = \begin{pmatrix} 1 & \dots & 1 \\ x_1 & \dots & x_n \end{pmatrix} \in \mathbb{R}^2 \times \mathbb{R}^n \,.$ Αρα,

$$\underline{\underline{X}}^T\underline{\underline{X}} = \begin{pmatrix} \underline{1}^T \\ \underline{x}^T \end{pmatrix} (\underline{1}, \underline{x}) = \begin{pmatrix} n & n\bar{x} \\ n\bar{x} & \Sigma_1^n x_i^2 \end{pmatrix} , \quad |\underline{\underline{X}}^T\underline{\underline{X}}| = n^2 \hat{\sigma}_x^2 ,$$

και αρα,

$$(\underline{\underline{X}}^T\underline{\underline{X}})^{-1} = \frac{1}{n^2\hat{\sigma}_x^2} \begin{pmatrix} \Sigma_1^n x_i^2 & -n\bar{x} \\ -n\bar{x} & n \end{pmatrix} .$$ Εχουμε λοιπον,

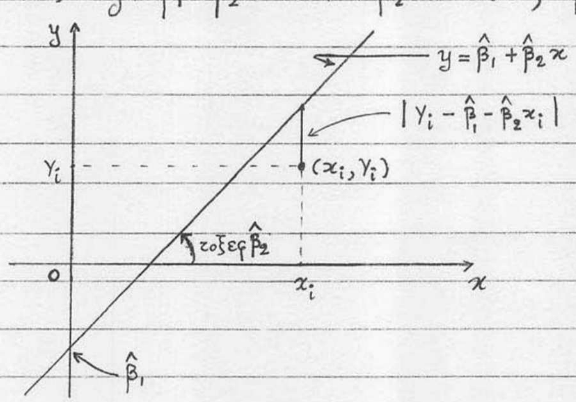
$$\hat{\underline{\beta}} = \begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} = (\underline{\underline{X}}^T\underline{\underline{X}})^{-1} \underline{\underline{X}}^T \underline{Y} = (\underline{\underline{X}}^T\underline{\underline{X}})^{-1} \begin{pmatrix} n\bar{Y} \\ \Sigma_1^n x_i Y_i \end{pmatrix} =$$

$$= \frac{1}{n^2\hat{\sigma}_x^2} \begin{pmatrix} n\bar{Y}\Sigma_1^n x_i^2 - n\bar{x}\Sigma_1^n x_i Y_i \\ -n^2\bar{x}\bar{Y} + n\Sigma_1^n x_i Y_i \end{pmatrix} = \frac{1}{n\hat{\sigma}_x^2} \begin{pmatrix} n\hat{\sigma}_x^2\bar{Y} - \bar{x}\Sigma_1^n (x_i-\bar{x})(Y_i-\bar{Y}) \\ \Sigma_1^n (x_i-\bar{x})(Y_i-\bar{Y}) \end{pmatrix} =$$

$$= \begin{pmatrix} \bar{Y} - \bar{x} r \hat{\sigma}_y/\hat{\sigma}_x \\ r\hat{\sigma}_y/\hat{\sigma}_x \end{pmatrix} .$$

Αρα, η ευθεια απλης παλινδρομησης εκτιμαται ως εξης:

(3.228) $\quad y = \hat{\beta}_1 + \hat{\beta}_2 x = \bar{Y} + \hat{\beta}_2 (x - \bar{x}) , \quad \hat{\beta}_2 = r\hat{\sigma}_y/\hat{\sigma}_x .$



(3.229) <u>Ασκηση</u>. Υπολογισε τις ΕΕΤ $\hat{\beta}_1$, $\hat{\beta}_2$ του Παραδ. (3.224), για

ποσοτητες λιπασματος x:	70	75	80	85	90	95	100	105	110	115	120	125	130
και αντιστοιχες αποδοσεις $Y=y$:	30	26	51	48	40	46	61	76	61	50	64	53	70

(3.230) <u>Ασκηση</u>. Εστω το γραμμικο μοντελο παλινδρομησης:

$$Y_i = \beta_1 + \beta_2 x_i + \beta_3 x_i^2 + \varepsilon_i , \quad i=1,\dots,n .$$

(α) Βρητε τις ΕΕΤ των παραμετρων $\underline{\beta}$, μεσω των (3.214) και των (3.216).

(β) Υπολογισε την παραβολη παλινδρομησης για τα δεδομενα της (3.229).

(γ) Υπολογισε την ΟΑΓΕΕΔ της $q(\underline{\beta}) = \beta_2 + 2\bar{x}\beta_3$.

Θα ασχοληθουμε τωρα με την εκτιμηση της διασπορας των λαθων ε_i, $i=1,\dots,n$.

(3.231) <u>Προταση.</u> Εστω το γραμμικο μοντελο παλινδρομησης $\underline{Y} = \underline{\underline{X}}\,\underline{\beta} + \underline{\varepsilon}$, με $E\underline{\varepsilon} = \underline{0}$, $\mathcal{D}(\underline{\varepsilon}) = \sigma^2 \underline{\underline{I}}$, και εστω οτι ο πινακας σχεδιασμου $\underline{\underline{X}}$ εχει πληρη διασταση **k**.
Τοτε, $E s^2 = \sigma^2 := \mathcal{D}(\varepsilon_1) = E\varepsilon_1^2$, οπου,
$$s^2 := \frac{1}{n-k}(\underline{Y} - \underline{\underline{X}}\hat{\underline{\beta}})^T(\underline{Y} - \underline{\underline{X}}\hat{\underline{\beta}}).$$
<u>Αποδ.</u> Απο την (3.222) εχουμε οτι
$$\underline{\varepsilon} = \underline{Y} - \underline{\underline{X}}\,\underline{\beta} = \underline{\underline{X}}(\hat{\underline{\beta}} - \underline{\beta}) \oplus (\underline{Y} - \underline{\underline{X}}\hat{\underline{\beta}}), \text{ και αρα,}$$
$$\Sigma_{i=1}^n \varepsilon_i^2 = \underline{\varepsilon}^T\underline{\varepsilon} = (\hat{\underline{\beta}} - \underline{\beta})^T \underline{\underline{X}}^T\underline{\underline{X}}(\hat{\underline{\beta}} - \underline{\beta}) + (n-k)s^2.$$
Απο την (3.218β), εχουμε οτι $(\underline{\underline{X}}^T\underline{\underline{X}})^{-1} = [c_{ij}] \in \mathbb{R}^k \times \mathbb{R}^k$, οπου $c_{ij} \equiv c_{ji} := \mathrm{cov}(\hat{\beta}_i, \hat{\beta}_j)/\sigma^2$, αν δε θεσουμε $\underline{\underline{X}}^T\underline{\underline{X}} \equiv [d_{ij}] \in \mathbb{R}^{2k}$, εχουμε οτι $\Sigma_{l=1}^k d_{il} c_{lj} = 1(i=j)$, $i,j=1,\dots,k$.
Αρα,
$$(n-k)Es^2 = E\{\Sigma_{i=1}^n \varepsilon_i^2 - \Sigma_{i=1}^k \Sigma_{j=1}^k d_{ij}(\hat{\beta}_i - \beta_i)(\hat{\beta}_j - \beta_j)\}$$
$$= n\sigma^2 - \sigma^2 \Sigma_{i=1}^k \Sigma_{j=1}^k d_{ij} c_{ij} =$$
$$= n\sigma^2 - \sigma^2 \Sigma_{i=1}^k \left(\Sigma_{j=1}^k d_{ij} c_{ji}\right) = n\sigma^2 - \sigma^2 \Sigma_{i=1}^k 1 = (n-k)\sigma^2,$$
δηλαδη, $Es^2 = \sigma^2$.

(3.232) <u>Ασκηση.</u> Εστω το γραμμικο μοντελο της (3.231) με την επι πλεον υποθεση οτι $\varepsilon_1, \dots, \varepsilon_n$ α.ι. $N(0, \sigma^2)$.
Δειξτε οτι:

(α) η s^2 ειναι η ΟΑΕΕΔ της σ^2.
(β) η εμπ της σ^2 ειναι η $\hat{\sigma}_n^2 := \frac{1}{n}(\underline{Y} - \underline{\underline{X}}\hat{\underline{\beta}})^T(\underline{Y} - \underline{\underline{X}}\hat{\underline{\beta}})$.
(γ) $(n-k)s^2/\sigma^2$, $n\hat{\sigma}_n^2/\sigma^2 \sim \chi^2_{n-k}$. (Β2, (1.62) για συγκριση.)
(δ) $\hat{\underline{\beta}}$, s^2 ανεξαρτητες.

(3.233) <u>Παρατηρηση.</u> Εστω το γραμμικο μοντελο παλινδρομησης:
$\underline{Y} = \underline{\underline{X}}\,\underline{\beta} + \underline{\varepsilon}$, $\underline{\varepsilon} \sim N_n(\underline{0}, \sigma^2\underline{\underline{I}})$, $\underline{\beta} \in \mathbb{R}^k$, $\sigma^2 > 0$ (η ενοχλητικη αγνωστη παραμετρος), ο πινακας σχεδιασμου $\underline{\underline{X}} \in \mathbb{R}^n \times \mathbb{R}^k$ εχει

πληρη διασταση k. Απο την (3.219β) εχουμε οτι $\forall \underline{c} \in \mathbb{R}^k$ σταθερο,

$$\frac{\underline{c}^T(\hat{\underline{\beta}}-\underline{\beta})}{\sigma\sqrt{\underline{c}^T(\underline{\underline{X}}^T\underline{\underline{X}})^{-1}\underline{c}}} \sim N_1(0,1) \quad ,$$

και επισης απο την (3.232γ και δ) αυτη ειναι ανεξαρτητη της $(n-k)S^2/\sigma^2 \sim \chi^2_{n-k}$. Αρα,

$$T := \frac{\underline{c}^T(\hat{\underline{\beta}}-\underline{\beta})}{S\sqrt{\underline{c}^T(\underline{\underline{X}}^T\underline{\underline{X}})^{-1}\underline{c}}} \sim t_{n-k} \quad ,$$

στην οποια μπορουμε να βασισουμε την κατασκευη διαστηματων εμπιστοσυνης για παραμετρικες συναρτησεις – βλ. (3.188). Εχουμε λοιπον οτι το στοχαστικο διαστημα

$$\left[\underline{c}^T\hat{\underline{\beta}} - S\sqrt{\underline{c}^T(\underline{\underline{X}}^T\underline{\underline{X}})^{-1}\underline{c}}\; t_{n-k}(1-\alpha/2)\, ,\ \underline{c}^T\underline{\beta} + S\sqrt{\underline{c}^T(\underline{\underline{X}}^T\underline{\underline{X}})^{-1}\underline{c}}\; t_{n-k}(1-\alpha/2)\right]$$

ειναι ενα συμμετρικο διαστημα εμπιστοσυνης, επιπεδου $1-\alpha$, για την παραμετρικη συναρτηση $q(\underline{\beta}) = \underline{c}^T\underline{\beta}$.

Επισης, οπως στο παραδειγμα (3.182), συμπεραινουμε οτι το στοχαστικο διαστημα:

$$\left[\frac{(n-k)S^2}{\chi^2_{n-k}(1-\alpha_2)}\, ,\ \frac{(n-k)S^2}{\chi^2_{n-k}(\alpha_1)}\right]$$ ειναι ενα, επιπεδου $1-\alpha_1-\alpha_2$,

διαστημα εμπιστοσυνης για την (ενοχλητικη) παραμετρο σ^2.

(3.234) <u>Ασκηση</u>. Εστω το μοντελο απλης γραμμικης παλινδρομησης της (3.224), με την επιπλεον υποθεση οτι $\underline{\varepsilon} \sim N_n(\underline{0}, \sigma^2\underline{\underline{I}})$. Κατασκευαστε συμμετρικα, επιπεδου $1-\alpha$, διαστηματα εμπιστοσυνης για τις παραμετρους β_1, β_2, σ^2. (Τα διαστηματα τα αυτα δεν ισχυουν συγχρονως, σ' αυτο το επιπεδο εμπιστοσυνης.). Για τα δεδομενα της (3.229) και $\alpha = 0{,}10$, υπολογιστε τα ανωτερω διαστηματα εμπιστοσυνης.

(3.235) Παραδειγμα. Εστω το μοντελο αναλυσης διασπορας του (3.204).

(α) Θα δειξουμε οτι οι ΕΕΤ των παραμετρων $\beta^T=(\mu, a_1,\dots,a_r, b_1,\dots,b_c)\in\mathbb{R}^{r+c+1}$, ειναι οι εξης:

$\hat\mu=\bar{Y}_{..}$, $\hat{a}_i:=\bar{Y}_{i.}-\bar{Y}_{..}$, $i=1,\dots,r$, $\hat{b}_j:=\bar{Y}_{.j}-\bar{Y}_{..}$, $j=1,\dots,c$,

οπου, $y_{i.}:=\Sigma_{j=1}^{c} y_{ij}$, $\bar{y}_{i.}:=y_{i.}/c$, $i=1,\dots,r$

$y_{.j}:=\Sigma_{i=1}^{r} y_{ij}$, $\bar{y}_{.j}:=y_{.j}/r$, $j=1,\dots,c$

$y_{..}:=\Sigma_{i=1}^{r}\Sigma_{j=1}^{c} y_{ij}$, $\bar{y}_{..}:=y_{..}/rc$.

(β) Εφοσον, $E\varepsilon_{ij}=0$ και $E(\varepsilon_{ij}\varepsilon_{i'j'})=\mathbb{1}(i=i', j=j')\sigma^2$, μια αμεροληπτη εκτιμητρια της σ^2 ειναι η

$$s^2:=\frac{1}{(r-1)(c-1)}\Sigma_{i=1}^{r}\Sigma_{j=1}^{c}(Y_{ij}-\hat\mu-\hat{a}_i-\hat{b}_j)^2 .$$

(γ) Η ονομασια "αναλυση διασπορας" που φερει αυτο μοντελο οφειλεται στην ακολουθη αναλυση της εκτιμητριας της σ^2:

$(r-1)(c-1)s^2 = SS_T - SS_r - SS_c$,

οπου, $SS_T:=\Sigma_{i=1}^{r}\Sigma_{j=1}^{c}(Y_{ij}-\bar{Y}_{..})^2$,

$SS_r:=c\,\Sigma_{i=1}^{r}(\bar{Y}_{i.}-\bar{Y}_{..})^2$,

$SS_c:=r\,\Sigma_{j=1}^{c}(\bar{Y}_{.j}-\bar{Y}_{..})^2$,

και μετρουν αντιστοιχα τη μεταβλητοτητα την ολικη του δειγματος, την οφειλομενη στη μεταβλητοτητα μεταξυ των σειρων και την οφειλομενη στη μεταβλητοτητα μεταξυ των στηλων του δειγματος. Η μεταβλητοτητα του δειγματος που δεν ερμηνευεται απο τις μεταβλητοτητες μεταξυ των σειρων και μεταξυ των στηλων του δειγματος, αθροιζεται στην $(r-1)(c-1)s^2$ η οποια καλειται υπολοιπομενη μεταβλητοτητα του δειγματος.

(δ) Αν επι πλεον $\varepsilon_{ij}\sim N(0,\sigma^2)$ $\forall i,j$ τοτε,

$SS_T/\sigma^2\sim\chi^2_{rc-1}$, $SS_r/\sigma^2\sim\chi^2_{r-1}$,

$SS_c/\sigma^2\sim\chi^2_{c-1}$, $(r-1)(c-1)s^2\sim\chi^2_{(r-1)(c-1)}$,

και ειναι ανεξαρτητες μεταξυ τους.

(ε) Με την επιπλέον υπόθεση του μέρους (δ), έχουμε ότι:

$$\frac{\sqrt{rc}\,(\hat{\mu}-\mu)}{S}\ ,\ \sqrt{\frac{rc}{r-1}}\ \frac{(\hat{a}_i-a_i)}{S}\ ,\ i=1,\dots,r\ ,$$

$$\sqrt{\frac{rc}{c-1}}\ \frac{(\hat{b}_j-b_j)}{S}\ ,\ j=1,\dots,c\ ,$$

είναι όλες τους $t_{(r-1)(c-1)}$ μεταβλητές.

Άρα, τα ακόλουθα στοχαστικά διαστήματα είναι συμμετρικά, επιπέδου $1-\alpha$, διαστήματα εμπιστοσύνης:

το $\hat{\mu} \pm t_{(r-1)(c-1)}^{(1-\alpha/2)}\ \frac{S}{\sqrt{rc}}$ για την μ,

το $\hat{a}_i \pm t_{(r-1)(c-1)}^{(1-\alpha/2)}\sqrt{\frac{r-1}{rc}}\ S$ για την a_i , $i=1,\dots,r$

και το $\hat{b}_j \pm t_{(r-1)(c-1)}^{(1-\alpha/2)}\sqrt{\frac{c-1}{rc}}\ S$ για την b_j , $j=1,\dots,c$.

Επίσης, το στοχαστικό διάστημα,

$$\left[\frac{(r-1)(c-1)S^2}{\chi^2_{(r-1)(c-1)}(1-\alpha_2)}\ ,\ \frac{(r-1)(c-1)S^2}{\chi^2_{(r-1)(c-1)}(\alpha_1)}\right]$$ είναι ένα, επιπέδου

$1-\alpha_1-\alpha_2$, διάστημα εμπιστοσύνης για την (ενοχλητική) παράμετρο σ^2.

Θα αποδείξουμε εδώ μόνο το μέρος (α). Η απόδειξη του (β) είναι ανάλογη της (3.231), το (γ) είναι ζήτημα πράξεων, τα δε (δ) και (ε) αφήνονται σαν άσκηση ανάλογη των (3.232) και (3.233).

Για την απόδειξη του (α), παρατηρούμε ότι πρέπει να βρούμε τα σημεία ελαχίστου της

$$\underline{\varepsilon}^T\underline{\varepsilon} = \Sigma_{i=1}^{r}\Sigma_{j=1}^{c}\,(Y_{ij}-\mu-a_i-b_j)^2\ ,$$

υπό τις συνθήκες: $\Sigma_{i=1}^{r} a_i = 0$,

και $\Sigma_{j=1}^{c} b_j = 0$.

Η αντικειμενικη συναρτηση $\underline{\varepsilon}^T\underline{\varepsilon}$ του προβληματος ειναι γνησια κυρτη και οι συνθηκες μας γραμμικες και αρα το προβλημα εχει μοναδικη λυση, την οποια και θα βρουμε με την χρηση των πολλαπλασιαστων Lagrange.

Η Lagrangian του προβληματος ειναι η:

$$L(\mu, \underline{a}, \underline{b}, \lambda_1, \lambda_2) := \Sigma_{i=1}^{r}\Sigma_{j=1}^{c}(Y_{ij}-\mu-a_i-b_j)^2+\lambda_1\Sigma_{i=1}^{n}a_i+\lambda_2\Sigma_{j=1}^{n}b_j,$$

με $\lambda_1, \lambda_2 \geq 0$.

Αρα,

$$\frac{\partial}{\partial\mu}L = -2\Sigma_{i=1}^{r}\Sigma_{j=1}^{c}(Y_{ij}-\mu-a_i-b_j) = $$
$$= -2(Y_{..}-rc\mu-ca_{.}-rb_{.}) = 0,$$

και,

$$\frac{\partial}{\partial a_l}L = -2\Sigma_{j=1}^{c}(Y_{lj}-\mu-a_l-b_j)+\lambda_1 = $$
$$= -2(Y_{l.}-c\mu-ca_l-b_{.})+\lambda_1 = 0, \quad l=1,\dots,r$$

και,

$$\frac{\partial}{\partial b_l}L = -2\Sigma_{i=1}^{r}(Y_{il}-\mu-a_i-b_l)+\lambda_2 = $$
$$= -2(Y_{.l}-r\mu-a_{.}-rb_l)+\lambda_2 = 0, \quad l=1,\dots,c,$$

Δηλαδη εχουμε το συστημα:

$$\begin{cases} Y_{..}-rc\mu-ca_{.}-rb_{.}=0 \\ Y_{l.}-c\mu-ca_l-b_{.}=\lambda_1/2, & l=1,\dots,r \\ Y_{.l}-r\mu-a_{.}-rb_l=\lambda_2/2, & l=1,\dots,c \end{cases}$$

και τις συνθηκες $a_{.}=b_{.}=0$. Αρα, το συστημα μας γινεται:

$$\begin{cases} Y_{..}-rc\hat{\mu}=0 \\ Y_{l.}-c\hat{\mu}-c\hat{a}_l=\lambda_1/2, & l=1,\dots,r \\ Y_{.l}-r\hat{\mu}-r\hat{b}_l=\lambda_2/2, & l=1,\dots,c. \end{cases}$$

Απο την πρωτη εξισωση εχουμε οτι: $\hat{\mu}=\overline{Y}_{..}$,

και αρα οι επομενες r εξισωσεις γινονται:

$$Y_{l.}-c\overline{Y}_{..}-c\hat{a}_l=\lambda_1/2 \quad l=1,\dots,r,$$

οι οποιες αθροιζομενες διδουν οτι:

$$\frac{n\lambda_1}{2} = Y_{..}-rc\overline{Y}_{..}-c\hat{a}_{.} = 0 \Rightarrow \lambda_1 = 0 \quad \text{και αρα}$$

$Y_{\ell\cdot} - c\bar{Y}_{\cdot\cdot} - c\hat{a}_\ell = 0$, $\ell = 1, \dots, r$, και αρα

$\hat{a}_\ell = \bar{Y}_{\ell\cdot} - \bar{Y}_{\cdot\cdot}$, $\ell = 1, \dots, r$.

Με ακριβως τον ιδιο τροπο οι υπολοιπες c εξισωσεις του συστηματος μας διδουν $\hat{b}_\ell = \bar{Y}_{\cdot\ell} - \bar{Y}_{\cdot\cdot}$, $\ell = 1, \dots, c$.

(3.236) Ασκηση. Εστω το μοντελο αναλυσης διασποράς του παραδειγματος (3.235) — βλ. και (3.204). Εστω δε οτι τα δεδομενα μας Y_{ij} ειναι τα εξης:

i\j	μηχανες			
χειριστες	6	15	10	12
	14	16	14	18
	8	14	10	15

(α) Υπολογισε τις ΕΕΤ των παραμετρων μ, a_i , $i=1,2,3$, b_j , $j=1,2,3,4$.

(β) Υπολογισε την αμεροληπτη εκτιμητρια s^2 της σ^2.

(γ) Υποθεσε οτι τα ε_{ij} ειναι α.ι. $N(0,\sigma^2)$ και κατασκευασε συμμετρικα, επιπεδου $1-\alpha = 0{,}95$, διαστηματα εμπιστοσυνης για τις παραμετρους μ, a_i, b_j , $i=1,2,3$ και $j=1,2,3,4$. Επισης, κατασκευασε ενα συμμετρικο διαστημα εμπιστοσυνης, επιπεδου $1-\alpha = 0{,}90$ για την (ενοχλητικη) παραμετρο σ^2 του προβληματος.

(δ) Υπολογισε τα SS_T, SS_r, SS_c του προβληματος

(ε) Με τις υποθεσεις του μερους (γ), κατασκευασε και υπολογισε την ε.μ.π. της σ^2.

4. ΕΛΕΓΧΟΙ ΥΠΟΘΕΣΕΩΝ

Εστω $X_1, \ldots, X_n$ α.ι. παρατηρησεις απο την κατανομη με πυκνοτητα $f(\cdot|\theta)$ ή συναρτηση μαζας πιθανοτητος $p(\cdot|\theta)$, $\theta \in \Theta \subseteq \mathbb{R}^k$. Το προβλημα του ελεγχου μιας υποθεσεως (Hypothesis) H, σχετικης με την αγνωστη παραμετρο θ, π.χ., οτι η αληθης τιμη της θ ανηκει στο υποσυνολο Θ_0 του παραμετρικου χωρου Θ, την οποια υποθεση συμβολιζουμε ως $H: \theta \in \Theta_0$, μπορει να διατυπωθη ως εξης:

εναντι της μονης εναλλακτικης δυνατοτητας $K: \theta \in \Theta_1$, με $\Theta_1 \subseteq \Theta$ και $\Theta_0 \cap \Theta_1 = \emptyset$, που δεχομαστε, ειναι η υποθεση $H: \theta \in \Theta_0$ αληθης;

Διατυπωνουμε αυτο συμβολικα ως εξης:

(4.1) $H: \theta \in \Theta_0$ vs $K: \theta \in \Theta_1$,

Ζητουμε δε την κατασκευη μιας στατιστικης συναρτησεως αποφασεως

(4.2) $d: \mathcal{X} \ni \underline{x} \longmapsto d(\underline{x}) \in [0,1]$,

την οποια καλουμε ελεγχοσυναρτηση ή απλως ελεγχο της υποθεσεως H εναντι της εναλλακτικης δυνατοτητας K, και εαν μεν $d(\underline{X}) = 0$ τοτε δεχομαστε την H ως αληθη, εαν δε $d(\underline{X}) = 1$ τοτε απορριπτουμε την υποθεση H, εναντι της εναλλακτικης δυνατοτητας K. Στην περιπτωση που $d(\underline{X}) \in (0,1)$, ο ελεγχος καλειται τυχαιοποιημενος, η ερμηνεια δε που διδεται στην τιμη $d(\underline{x})$, βασει του δειγματος $\underline{X}(\omega) = \underline{x} \in \mathcal{X}$, ειναι οτι η υποθεση μας H ειναι αληθης με πιθανοτητα $1 - d(\underline{x})$ και εσφαλμενη με πιθανοτητα $d(\underline{x})$. Για να καταληξουμε δε σε μια αποφαση, υπερ ή κατα της υποθεσεως H, κατασκευαζουμε μια τυχαια μεταβλητη

(4.3) $\xi \sim \mathcal{B}ernoulli\,(p = d(\underline{x}))$,

και εκτελουμε ενα τυχαιο πειραμα με βαση αυτην — εξ ου και η ονομασια του ελεγχου ως τυχαιοποιημενου — και εαν μεν η τ.μ. ξ λαβη την τιμη 0 τοτε δεχομεθα την υποθεση H ως αληθη, εαν ομως η ξ λαβη την τιμη 1, τοτε απορριπτουμε την H εναντι της εναλλακτικης δυνατοτητας K που εχουμε.

<u>(4.4) Σημ.</u> Μια ελεγχοσυναρτηση $d(\underline{X})$ ειναι η ιδια μια στοχαστικη μεταβλητη με τιμες στο κλειστο διαστημα $[0,1]$, που ειναι ο χωρος αποφασεων $\mathcal{A}$ του προβληματος μας — αν θελουμε να χρησιμοποιησουμε τη γλωσσα της θεωριας αποφασεων. Σ' αυτο το πλαισιο λοιπον, η $d(\underline{X})$ δεν ειναι η πιθανοτητα του να μην ειναι αληθης η H, αλλα η $d(\underline{X})$ ειναι ακριβως η αποφαση μας σχετικα με την αληθεια ή μη της H. Επειδη εχουμε συνηθισει να θεωρουμε ως αποφαση μονο την αποδοχη ή απορριψη μιας υποθεσης, δηλαδη, να εχουμε ως χωρο αποφασεων τον $\mathcal{A}_0 \equiv \{0,1\}$, μπορουμε να δουμε την $d(\underline{X})$, της (3.2), σαν μια <u>γενικευμενη αποφαση</u> στον χωρο αποφασεων $\mathcal{A} \equiv [0,1]$. Ειναι δε δυνατον, να ερμηνευσουμε την (γενικευμενη) αποφαση $d(\underline{x}) \in [0,1]$, που παρνουμε βασει του δειγματος μας $\underline{X}(\omega) = \underline{x} \in \mathcal{X}$ $(\omega \in \Omega)$, ως την πιθανοτητα του να μην ειναι αληθης η υποθεση μας H. Σ' αυτην την περιπτωση η τ.μ. της (4.3), μας παρεχει ενα μηχανισμο μεσω του οποιου ειναι δυνατον να φθασουμε σε μια αποφαση — με την κλασσικη εννοια — στον $\mathcal{A}_0 \equiv \{0,1\}$.

Φυσικα, ενας τυχαιοποιημενος ελεγχος <u>υστερει</u> ενος μη τυχαιοποιημενου, διοτι μεσω ενος τετοιου ή δεν καταληγουμε σε μια ξεκαθαρη — ναι ή οχι — απαντηση στον $\mathcal{A}_0$ ή χρειαζεται να την επιμοιεσουμε μεσω μιας ξενης προς το προβλημα τ.μ. ξ, της (4.3).

Θα χρησιμοποιησουμε τα δυο παραδειγματα που ακολουθουν για να ξεκαθαρισουμε τις ανωτερω εννοιες και για να εισαγουμε διαφορες νεες εννοιες.

(4.5) <u>Παραδειγμα</u>. Εστω οτι $Y_1, \dots, Y_n$ α.ι. $N(\mu, \tau^2)$ ειναι οι συστολικες πιεσεις $n \in \mathbb{N}$ ασθενων που υφιστανται την κλασσικη – δοκιμασμενη – αγωγη A για το ριξιμο της προβληματικα υψηλης πιεσης τους. Εστω δε οτι επιτυγχανεται απο ορισμενους ερευνητες η δημιουργια μιας νεας αγωγης B, για το ριξιμο, επισης, της συστολικης πιεσης. Τιθεται λοιπον τωρα το ερωτημα του κατα ποσον η αγωγη A παραμενει η προτιμητεα, δεδομενου οτι τωρα υπαρχει η εναλλακτικη αγωγη B. Για να μελετηθη αυτο το ερωτημα, οι ιδιοι η ασθενεις, εθελοντικα, υφιστανται τωρα την νεα αγωγη B επι καποιο χρονικο διαστημα και εστω οτι οι νεες συστολικες πιεσεις τους ειναι αντιστοιχως: $Y_1', \dots, Y_n'$ α.ι. $N(\mu', \tau^2)$ – για απλουστευση εχουμε δεχθη οτι οι δυο κατανομες ειναι κανονικες με την ιδια διασπορα τ^2.

Εστω οτι οι γιατροι εχουν συμφωνησει οτι αν $\mu \le \mu'$, δηλαδη, $\theta := \mu - \mu' \le 0$, τοτε θα δεχθουν την κλασσικη αγωγη A ως προτιμητεα, αν ομως $\theta > 0$ τοτε θα απορριψουν την επικρατουσα ως τωρα υποθεση οτι η αγωγη A ειναι η προτιμητεα.

Μια περεταιρω απλουστευση που δεχονται ειναι να βασισουν την αποφαση τους στις διαφορες $X_i := Y_i - Y_i'$, $i = 1, \dots, n$ των αντιστοιχων πιεσεων των ασθενων τους.

Εχουμε λοιπον τωρα το εξης στατιστικο προβλημα: δεδομενων των μετρησεων $X_1, \dots, X_n$ α.ι. $N(\theta, \sigma^2)$, με $\sigma^2 = 2\tau^2$ – το οποιο μας ειναι για καποιο λογο <u>γνωστο</u> – θελουμε να ελεγξουμε την υποθεση

(4.6) $\quad H: \theta \in \Theta_0 \equiv (-\infty, 0] \quad \text{vs} \quad K: \theta \in \Theta_1 \equiv (0, +\infty).$

Δεν θα επιχειρησουμε εδω την κατασκευη μιας ελεγχοσυναρτησης για το προβλημα (4.6) · για αυτην βλ. (4.   ). Θα χρησιμοποιησουμε τον διαισθητικα ικανοποιητικο ελεγχο — ειναι αλλωστε, οπως θα δουμε αργοτερα, βελτιστος βασει των κριτηριων που θα αναπτυξουμε:

$$(4.7)\quad d(\underline{X}) := 1\left(T(\underline{X})>c\right) \equiv \begin{cases} 1 & \text{αν } T(\underline{X})>c \\ 0 & \text{αν } T(\underline{X})\leq c \end{cases},$$

για την στατιστικη συναρτηση $T(\underline{X}) := \overline{X}_n$, η οποια καλειται *στατιστικη του ελεγχου* d και για καποια σταθερα $c \in \mathbb{R}$, η οποια καλειται *κρισιμη σταθερα* και την οποια εχουμε την ευχερεια να επιλεξουμε ετσι ωστε να σταθεροποιουμε καποια συναρτηση κινδυνου, π.χ., βλ. (2.20), στη ληψη αποφασεως μεσω του ελεγχου d.

Το μονο που μας υποδεικνυει ο ελεγχος (4.7) αυτη τη στιγμη ειναι να αποφασισουμε υπερ της H — δηλαδη, οτι η μεση τιμη θ της κατανομης μας ειναι μικρη και μαλιστα μη θετικη — αν και μονο αν ο μεσος ορος $\overline{X}_n$ των παρατηρησεων μας ειναι μικρος· το ποσο μικρος πρεπει να ειναι — δηλαδη, το ποια πρεπει να ειναι η κρισιμη σταθερα c του ελεγχου — θα το αποφασισουμε τωρα.

Παρατηρουμε, κατ' αρχην, οτι τα σφαλματα στα οποια μπορει να μας οδηγησει ενας ελεγχος d, οπως ο (4.7) ειναι:

(4.8.I). να μη δεχθουμε την H ως αληθη — $d(\underline{X})>0$ — ενω η H ειναι αληθης, και αυτο καλειται *σφαλμα τυπου I*,

και,

(4.8.II). να μην απορριψουμε την H ως εσφαλμενη — $d(\underline{X})<1$ — ενω η H ειναι εσφαλμενη, και αυτο καλειται *σφαλμα τυπου II*.

Στη γλωσσα της θεωριας αποφασεων, η απωλεια που ειναι δυνατον να υποστουμε οταν παιρνουμε την αποφαση $d(\underline{x})$, ενω η αγνωστη αληθης τιμη της παραμετρου ειναι θ, οριζεται ως :

(4.9) $L(\theta,d(\underline{x})) := d(\underline{x})\,1(\theta\in\Theta_0) + [1-d(\underline{x})]\,1(\theta\in\Theta_1)\ ,\ \underline{x}\in\mathcal{X}\ ,\ \theta\in\Theta.$

Οριζεται δηλαδη, η συναρτηση απωλειας, στο προβλημα ελεγχου υποθεσεως, ως η αποσταση της αποφασης $d(\underline{x})\in[0,1]$ — που παιρνουμε βασει του δειγματος $\underline{x}\in\mathcal{X}$ — απο την σωστη αποφαση 0 ή 1 που θα επρεπε να παρουμε εφοσον $\theta\in\Theta_0$ ή $\theta\in\Theta_1$ αντιστοιχα. Τοτε, η συναρτηση κινδυνου του ελεγχου d ειναι η εξης :

(4.10) $R(\theta,d) := E_\theta L(\theta,d(\underline{X})) = \beta_d(\theta)\,1(\theta\in\Theta_0) + [1-\beta_d(\theta)]\,1(\theta\in\Theta_1)$,

οπου, η συναρτηση

(4.11) $\beta_d(\theta) := E_\theta\{d(\underline{X})\}\ ,\ \theta\in\Theta$,

καλειται συναρτηση ισχυος του ελεγχου d, η δε τιμη της $\beta_d(\theta)$ καλειται ισχυς του ελεγχου d στη θεση $\theta\in\Theta$.

Παρατηρουμε οτι η $\beta_d(\cdot)$ τεινει να παιρνη μεγαλες τιμες — κοντα στο 1 — οταν η ελεγχοσυναρτηση $d(\underline{X})$ τεινει να παιρνη τιμες κοντα στο 1, δηλαδη τεινει να απορριπτη την H. Οπως, λοιπον, βλεπουμε και απο την (4.10), η συναρτηση ισχυος $\beta_d(\cdot)$ ένος "καλου" ελεγχου θα πρεπη να τεινη να παιρνη μικρες τιμες — κοντα στο 0 — πανω στο Θ_0 και μεγαλες τιμες — κοντα στο 1 — πανω στο Θ_1. Ενα τετοιος "καλος", λοιπον, ελεγχος d θα πρεπη να κρατα χαμηλο το μετρο μετρο του σφαλματος τυπου I :

(4.12) $\alpha_I(\theta) := \beta_d(\theta)\,1(\theta\in\Theta_0)$,

και επισης χαμηλο το μετρο του σφαλματος τυπου II :

(4.13) $\alpha_{II}(\theta) := [1-\beta_d(\theta)]\,1(\theta\in\Theta_1)$,

ή ισοδυναμα να εχη μεγαλη ισχυ πανω στο Θ_1.

Παρατηρουμε οτι δεν υπαρχει ελεγχος d^* που να ελαχιστοποιη την συναρτηση κινδυνου $R(\theta,d)$ ομοιομορφα ως προς $\theta\in\Theta$, διοτι η τετριμμενη ελεγχοσυναρτηση $d_1(\underline{x})\equiv 0\ \forall\, \underline{x}\in\mathcal{X}$,

που πάντα αποδέχεται την H έχει $R(\theta, d_1) = \alpha_I(\theta) = 0 \ \forall \theta \in \Theta_0$,
και η επίσης τετριμμένη ελεγχοσυνάρτηση $d_2(x) \equiv 1 \ \forall x \in \mathcal{X}$,
που πάντα απορρίπτει την H έχει $R(\theta, d_2) = \alpha_{II}(\theta) = 0 \ \forall \theta \in \Theta_1$.
Άρα, ένας d^* που θα ελαχιστοποιούσε την συνάρτηση κινδύνου ομοιόμορφα ως προς $\theta \in \Theta$, θα έπρεπε να έχει $R(\theta, d^*) = 0 \ \forall \theta \in \Theta$, και το οποίο απαιτεί γνώση της αληθούς θ.

Υπάρχουν διάφοροι τρόποι για να ξεπερασθεί αυτό το πρόβλημα — ανάλογοι προς τις μεθόδους minimax, Bayes και ΟΑΕΕΔ της εκτιμητικής. Η μέθοδος πάντως που εκεί επικράτησε στην πράξη της στατιστικής για την κατασκευή ελέγχων, είναι η εξής:

(4.14) <u>Μέθοδος των Neyman και Pearson</u>:

για κάποιο $\alpha \in [0,1]$, χρησιμοποιούμε για τον έλεγχο της $H: \theta \in \Theta_0$ vs $K: \theta \in \Theta_1$, εκείνη την ελεγχοσυνάρτηση d^*, αν υπάρχει, για την οποία

(4.15) $\beta_{d^*}(\theta) \geq \beta_d(\theta) \quad \forall \ \theta \in \Theta_1$,

για κάθε ελεγχοσυνάρτηση d τέτοια ώστε

(4.16) $\sup\{\beta_d(\theta), \theta \in \Theta_1\} \leq \sup\{\beta_{d^*}(\theta), \theta \in \Theta_1\} \leq \alpha$.

Το $\sup\{\beta_d(\theta), \theta \in \Theta_1\}$ καλείται <u>μέγεθος</u> του ελέγχου d, είναι δε συνήθως δυνατόν να έχουμε το μέγεθος του d^* ακριβώς ίσο με το προκαθορισμένο επίπεδο α, το οποίο και επιθυμούμε για να κερδίσουμε σε ισχύ του ελέγχου.
Ένας έλεγχος d^* για τον οποίον ισχύουν οι (4.15) και (4.16), καλείται <u>ομοιόμορφα πλέον ισχυρός</u> (Ο.Π.Ι.) για το μέγεθος του.
Ένας ΟΠΙ έλεγχος λοιπόν, κρατά το μέγεθος του σφάλματος τύπου I χαμηλό — δηλαδή, $\alpha_I(\theta) \leq \alpha$ — και ελαχιστοποιεί ομοιόμορφα ως προς $\theta \in \Theta$, το μέτρο $\alpha_{II}(\theta)$ του σφάλματος τύπου II, μεταξύ όλων των ελέγχων του ιδίου μεγέθους.

Μια άλλη επιθυμητή ιδιότητα των ελέγχων είναι η <u>αμεροληψία</u> τους, δηλαδή, (4.17) $\beta_d(\theta) \geq \sup\{\beta_d(\theta'), \theta' \in \Theta_0\} \quad \forall\, \theta \in \Theta_1$,

εξασφαλιζουν, δηλαδη, ισχυ — ικανοτητα διακρισεως της αληθειας ή μη της H — τουλαχιστον ιση προς το μεγεθος τους. Οι περισσοτεροι των ελεγχων που θα κατασκευασουμε στην επομενη ενοτητα θα ειναι ΟΠΙ ή τουλαχιστον ΟΠΙ μεταξυ των αμεροληπτων ελεγχων, για το μεγεθος τους.

Το υποσυνολο του δειγματικου χωρου $\mathcal{X}$,

(4.18) $C_d := \{\underline{x} \in \mathcal{X} : d(\underline{x}) = 1\}$,

καλειται κρισιμη περιοχη του ελεγχου d, διοτι για τα δειγματα που ανηκουν στο C_d απορριπτουμε την H, και το

(4.19) $A_d := \{\underline{x} \in \mathcal{X} : d(\underline{x}) = 0\}$,

καλειται περιοχη αποδοχης του ελεγχου d, διοτι για τα δειγματα που ανηκουν στο A_d αποδεχομεθα την H ως αληθη. Παρατηρουμε οτι $A_d \cap C_d = \emptyset$, το δε υποσυνολο του δειγματικου χωρου $\mathcal{X} \setminus (A_d \cup C_d)$ περιεχει εκεινα τα δειγματα για τα οποια δεν ειναι δυνατον να εχουμε απο τον ελεγχο d, ξεκαθαρη αποφαση για την αληθεια ή μη της υποθεσης H και αναγκαζομαστε να καταφυγουμε στην τυχαιοποιηση του ελεγχου μας, με σκοπο παντα να αυξησουμε την ισχυ του ελεγχου μας. Για εναν μη τυχαιοποιημενο ομως ελεγχο — και αυτοι ειναι που μας ενδιαφερουν εδω κυριως — εχουμε οτι $C_d = \mathcal{X} \setminus A_d$, τουλαχιστον, με πιθανοτητα 1.

Θα ειδικευσουμε τωρα τις ανωτερω εννοιες στην περιπτωση μη τυχαιοποιημενων ελεγχων d, θα τις εφαρμοσουμε δε στην περιπτωση του μη τυχαιοποιημενου ελεγχου (4.7) της υποθεσεως (4.6). Εστω λοιπον ο μη τυχαιοποιημενος ελεγχος d_0, με $d_0(\underline{x}) = 1(\underline{x} \in C_{d_0}) \in \mathcal{A}_0 \equiv \{0,1\} \ \forall\, x \in \mathcal{X}$, τοτε η συναρτηση απωλειας γινεται:

(4.9') $L(\theta, d_0(\underline{x})) = 1(\underline{x} \in C_{d_0})1(\theta \in \Theta_0) + 1(x \notin C_{d_0})1(\theta \in \Theta_1)$,

$\underline{x} \in \mathcal{X}$, $\theta \in \Theta$, και αρα η συναρτηση κινδυνου του d_0 ειναι η εξης:

$$(4.10') \quad R(\theta, d_0) = E_\theta L(\theta, d(\underline{X})) =$$
$$= P_\theta(\underline{X} \in C_{d_0})\, 1(\theta \in \Theta_0) + P_\theta(\underline{X} \notin C_{d_0})\, 1(\theta \in \Theta_1)$$
$$= P_\theta(d_0(\underline{X}) = 1)\, 1(\theta \in \Theta_0) + P_\theta(d_0(\underline{X}) = 0)\, 1(\theta \in \Theta_1)$$
$$= \beta_{d_0}(\theta)\, 1(\theta \in \Theta_0) + [1 - \beta_{d_0}(\theta)]\, 1(\theta \in \Theta_1),$$

και παλι, διοτι η ισχυς του ελεγχου ειναι τωρα η εξης:

$$(4.11') \quad \beta_{d_0}(\theta) = E_\theta\{d_0(\underline{X})\} = P_\theta(d_0(\underline{X}) = 1), \quad \theta \in \Theta.$$

Τα μετρα δε των σφαλματων τυπου I και II γινονται τωρα ακριβως οι πιθανοτητες αυτων των σφαλματων, δηλαδη,

$$(4.12') \quad \alpha_{I}(\theta) = \beta_{d_0}(\theta)\, 1(\theta \in \Theta_0) = P_\theta(d_0(\underline{X}) = 1)\, 1(\theta \in \Theta_0)$$
$$= P(\text{σφαλματος τυπου I}),$$

$$(4.13') \quad \alpha_{II}(\theta) = [1 - \beta_{d_0}(\theta)]\, 1(\theta \in \Theta_1) = P_\theta(d_0(\underline{X}) = 0)\, 1(\theta \in \Theta_1)$$
$$= P(\text{σφαλματος τυπου II}).$$

Αρα, και το μεγεθος του ελεγχου d_0 ειναι η μεγιστη πιθανοτητα του σφαλματος τυπου I, οπως το θ κινειται στον Θ_0:

$$(4.16') \quad \sup\{\beta_{d_0}(\theta),\ \theta \in \Theta_0\} = \sup\{P_\theta(d_0(\underline{X}) = 1),\ \theta \in \Theta_0\},$$

ενας μη τυχαιοποιημενος δε ΟΠΙ ελεγχος d_0^* ικανοποιει την σχεση

$$(4.15') \quad \beta_{d_0^*}(\theta) = P_\theta(d_0^*(\underline{X}) = 1) \geq \beta_d(\theta) \quad \forall \theta \in \Theta_1,$$

και για καθε ελεγχο (τυχαιοποιημενο ή μη) του ιδιου μεγεθους.

Για παραδειγμα, ο ελεγχος (4.7) της υποθεσεως (4.6), εχει ισχυ:

$$\beta(\theta) = E_\theta\{d(\underline{X})\} = P_\theta(d(\underline{X}) = 1) = P_\theta(T(\underline{X}) > c) =$$
$$= P_\theta(\bar{X}_n > c) = P_\theta\left(\frac{\sqrt{n}(\bar{X}_n - \theta)}{\sigma} > \frac{\sqrt{n}(c - \theta)}{\sigma}\right)$$
$$= 1 - \Phi\left(\frac{\sqrt{n}(c-\theta)}{\sigma}\right) = \Phi\left(\frac{\sqrt{n}(\theta - c)}{\sigma}\right), \quad \theta \in \Theta,$$

η οποια ειναι αυξουσα συναρτηση του θ και αρα το μεγεθος του ελεγχου ειναι η ακολουθη συναρτηση της κρισιμης σταθερας c:

$$\sup\{\beta(\theta),\ \theta \in \Theta_0 = (-\infty, 0]\} = \sup\left\{\Phi\left(\frac{\sqrt{n}(\theta - c)}{\sigma}\right),\ \theta \in \Theta_0 = (-\infty, 0]\right\} =$$
$$= \Phi\left(-\frac{\sqrt{n}c}{\sigma}\right) = 1 - \Phi\left(\frac{\sqrt{n}c}{\sigma}\right) = \alpha \iff c \equiv c_\alpha := \frac{\sigma z(1-\alpha)}{\sqrt{n}},$$

οπου $z(1-\alpha)$ διδεται απο την (3.176). Αρα, ο ελεγχος $d(\underline{X}) = 1\left(\bar{X}_n > c_\alpha = \frac{\sigma z(1-\alpha)}{\sqrt{n}}\right)$, με κρισιμη περιοχη $C_d = \{\underline{x} \in \mathcal{X} = \mathbb{R}^n : \bar{x}_n > c_\alpha\}$ ειναι μεγεθους α — το οτι ειναι και ΟΠΙ για το μεγεθος του εξασφαλιζεται απο την κατασκευη του οπως θα δουμε στην επομενη ενοτητα. Σαν εφαρμογη, εστω $\sigma = 1$, $n = 100$, $\alpha = 0.05$ και αρα $z(1-\alpha) = 1.645$, τοτε $c_\alpha = 0.1645$.

(4.20) Σημειωση. Η προταση (4.14) των Neyman και Pearson — να ελαχιστοποιηται το μετρο $\alpha_{II}(\theta)$ του σφαλματος τυπου ΙΙ, οσο γινεται, αλλα οπωσδηποτε να κρατιεται το μετρο $\alpha_I(\theta)$ του σφαλματος τυπου Ι, χαμηλο, συνηθως πολυ χαμηλο: $\alpha_I(\theta) \leq \alpha = 0{,}05$ ή $0{,}10$ ή και μικροτερο — δεν μεταχειριζεται την υποθεση Η και την εναλλακτικη της Κ συμμετρικα, αλλα μαλλον μεροληπτει υπερ της Η. Αλλες μεθοδοι ελεγχου, οπως η minimax ή η Bayes, δεν εχουν δομημενη μεσα τους μια τετοια μεροληπτικοτητα υπερ της Η ή της εναλλακτικης της. Στην πραξη παντως, η προταση των Neyman και Pearson ειναι δικαιολογημενη, διοτι συνηθως η υποθεση Η αντιπροσωπευει καποιο standard, και το κλασσικο και δοκιμασμενο, π.χ., μια κλασσικη θεωρια της φυσικης, μια εν χρησει θεραπευτικη αγωγη, μια δοκιμασμενη μεθοδο βιομηχανικης κατασκευης, κ.τ.τ.. Το να ανατραπη λοιπον η υποθεση Η δεν ειναι επιθυμητο, εκτος αν η εναλλακτικη Κ προσφερει κατι το σαφως καλυτερο. Γιαυτο λοιπον, κρατάμε την πιθανοτητα να απορριψουμε την Η, ενω αυτη ειναι αληθης, μικρη — το πολυ $\alpha \in [0,1]$ — συνηθως δε παιρνουμε $\alpha = 0{,}05$ ή $0{,}10$ ή και μικροτερο.

Παντως, η θεωρια των Neyman και Pearson ειναι απολυτα συμβιβαστη με το ακολουθο (καθως πιο μοντερνο) πλαισιο αντιμετωπισεως του προβληματος του ελεγχου μιας υποθεσεως H, αποτελει δε τη βαση της ερμηνειας αυτης της τακτικης:

Εστω οτι με καποια μεθοδο, π.χ., του λημματος των Neyman και Pearson, ή του πηλικου πιθανοφανειων, κατασκευαζουμε τον <u>μη τυχαιοποιημενο</u> ελεγχο $d(\underline{X}) = 1(T(\underline{X}) > c)$ για καποια σταθερα c — την οποια θα μπορουσαμε να καθορισουμε οπως πριν, ουτως ωστε να επιτυχουμε καποιο επιθυμητο μεγεθος α — την οποια δεν χρειαζεται να καθορισουμε πλεον, και καποια στατιστικη συναρτηση $T(\underline{X})$ — οπως θα δουμε στις επομενες ενοτητες ολοι οι (βελτιστοι) ελεγχοι μπορουν να εκφρασθουν κατ' αυτον τον τροπο. Τοτε για το συγκεκριμενο δειγμα $\underline{x}$ που πηραμε μπορουμε να υπολογισουμε την πιθανοτητα:

(4.21) $\alpha(T(\underline{x})) := \sup\{P_\theta(T(\underline{X}) > T(\underline{x})),\ \theta \in \Theta_0\}\ \forall \underline{x} \in \mathcal{X}$,

η οποια καλειται <u>p-τιμη</u> του ελεγχου d - βασισμενου στη στατιστικη T - στο δειγμα $\underline{x} \in \mathcal{X}$.

Παρατηρουμε οτι η p-τιμη του ελεγχου d στο δειγμα $\underline{x} \in \mathcal{X}$, ειναι ενα μετρο του ποσο συνηθες ειναι να παρατηρηθη τυχον δειγμα $\underline{X}$ πιο ακραιο απο αυτο, το $\underline{x}$, που παρατηρηθη — ακραιο, με την εννοια του $T(\underline{X}) > T(\underline{x})$ — <u>εφοσον</u> η $H: \theta \in \Theta_0$ ειναι αληθης. Αν λοιπον η p-τιμη του ελεγχου d στο συγκεκριμενο δειγμα $\underline{x} \in \mathcal{X}$ που παρατηρησαμε ειναι μεγαλη, αυτο σημαινει οτι αν η H ειναι αληθης, η παρατηρηση ενος δειγματος σαν το $\underline{x}$ δεν ειναι και το ασυνηθες, και αρα το οτι παρατηρησαμε το δειγμα $\underline{x}$, αυτο αποτελει στοιχειο που ενισχυει την αληθοφανεια της H.

Απο την αποψη τωρα της μεθοδου των Neyman και Pearson, $\forall \alpha < \alpha(T(\underline{x}))$, εστω c_α η αντιστοιχη κριτικη τιμη του ελεγχου $d(\underline{X}) = 1(T(\underline{X}) > c_\alpha)$ μεγεθους α, δηλαδη,

$\sup\{P_\theta(T(\underline{X})>c_\alpha),\ \theta\in\Theta_0\} \leq \alpha < \sup\{P_\theta(T(\underline{X})>T(\underline{x})),\ \theta\in\Theta_0\}$.

Τωρα,

αν $T(\underline{x}) \geq c_\alpha \Rightarrow P_\theta(T(\underline{X})>c_\alpha) \geq P_\theta(T(\underline{X})>T(\underline{x}))\ \forall\theta$
$\Rightarrow \sup\{P_\theta(T(\underline{X})>c_\alpha),\ \theta\in\Theta_0\} \geq \sup\{P_\theta(T(\underline{X})>T(\underline{x})),\ \theta\in\Theta_0\}$,
το οποιο ειναι ατοπο, και αρα $T(\underline{x})<c_\alpha$, δηλαδη, το δειγμα $\underline{x}$ που παρατηρησαμε μας οδηγει – βασει του ελεγχου d – στην αποδοχη της H, για καθε μεγεθος ελεγχου μικροτερου της p-τιμης (μια καθως πιο λειτη αποδειξη μας οδηγει στο οτι η H γινεται δεκτη $\forall\alpha\leq\alpha(T(\underline{x}))$, βασει του δειγματος $\underline{x}\in\mathcal{X}$). Και παλι, λοιπον, οσο πιο μεγαλη ειναι η p-τιμη του ελεγχου στο δειγμα $\underline{x}$, τοσο πιο αληθοφανης ειναι η H.

Σαν εφαρμογη, εστω οτι στο παραδειγμα (4.5) εχουμε $\sigma=1$, $n=100$ και εστω οτι το δειγμα του πειραματος μας ειναι $\underline{X}(\omega)=\underline{x}$ με $\bar{x}_n=\frac{1}{100}\sum_{i=1}^{100}x_i=0.1$. Τοτε για καθε θ
$P_\theta(T(\underline{X})>T(\underline{x})) = P_\theta(\bar{X}_n>\bar{x}_n) = P_\theta\left(\frac{\sqrt{n}(\bar{X}_n-\theta)}{\sigma} > \frac{\sqrt{n}(\bar{x}_n-\theta)}{\sigma}\right) = 1-\Phi\left(\frac{\sqrt{n}(\bar{x}_n-\theta)}{\sigma}\right) = 1-\Phi(10(0.1-\theta))$.
Αρα, η p-τιμη του ελεγχου (4.7) στο δειγμα $\underline{x}$ ειναι:

$$\alpha(T(\underline{x})) = \alpha(\bar{x}_n) = \alpha(0.1) = \sup\{1-\Phi(10(0.1-\theta)),\ \theta\leq 0\}$$
$$= 1-\Phi(1) = 1-0.8413 = 0.1587,$$

και αρα για καθε επιπεδο $\alpha\leq\alpha(0.1)=0.1587$, ο ελεγχος (4.7) μας οδηγει στην αποδοχη της $H:\theta\leq 0$, ενω για καθε $\alpha>0.1587$ οδηγουμεθα στην απορριψη της H εναντι της εναλλακτικης της $K:\theta>0$. Για παραδειγμα, στο συνηθες επιπεδο $\alpha=0.05$ πρεπει να δεχθουμε την H ως αληθη· αλλωστε, για $\alpha=0.05$, $c_\alpha=0.1645>0.1=\bar{x}_n$.

Στις επομενες δυο ενοτητες θα αναπτυξουμε τις δυο επικρατεστερες μεθοδους κατασκευης ελεγχων, αυτην των Neyman και Pearson και την μεθοδο του λογου πιθανοφανειων. Η πρωτη, οταν μπορει να χρησιμοποιηθη, εξασφαλιζει οτι οι ελεγχοι της ειναι ΟΠΙ, η δευτερη εχει ευρυτερη εφαρμογη και οι ελεγχοι της ειναι, εν γενει, ασυμπτωτικα ΟΠΙ. Συχνα και οι δυο μεθοδοι καταληγουν στην ιδια ελεγχοσυναρτηση.

4.1. ΕΛΕΓΧΟΙ NEYMAN - PEARSON.

Η κατασκευη των ελεγχων των Neyman και Pearson, βασιζεται στο ακολουθο λημμα των ιδιων :

(4.22) Λημμα (των Neyman και Pearson). Εστω δειγμα $X_1, \ldots, X_n$ α.ι. $f \in \mathcal{F}$. Για τον ελεγχο της υποθεσεως $H: f=f_0$ vs $K: f=f_1$, οπου $f_0, f_1 \in \mathcal{F}$ απολυτα καθορισμενες κατανομες, ο πλεον ισχυρος ελεγχος για το μεγεθος του ειναι ο ακολουθος :

$$d_c(\underline{X}) := \begin{cases} 1 & \text{αν} \quad \dfrac{\prod_{i=1}^n f_1(X_i)}{\prod_{i=1}^n f_0(X_i)} > c \\ \gamma & \text{αν} \quad \quad '' \quad \quad = c \\ 0 & \text{αν} \quad \quad '' \quad \quad < c \end{cases},$$

για καθοια κρισιμη σταθερα $c \geq 0$ και καθοιο $\gamma \in [0,1]$. Εχουμε, δηλαδη, οτι για καθε ελεγχο d μεγεθους $\beta_d(f_0) \leq \beta_{d_c}(f_0)$, ο ελεγχος d_c εχει μεγαλυτερη ισχυ:

$$\beta_d(f_1) \leq \beta_{d_c}(f_1).$$

Αποδ. Εχουμε οτι: $\beta_d(f_1) - c\beta_d(f_0) :=$

$$:= E_{f_1}\{d(\underline{X})\} - c\,E_{f_0}\{d(\underline{X})\} =$$

$$= \int_{\mathbb{R}^n} d(\underline{x})\Big\{\prod_{i=1}^n f_1(x_i) - c\prod_{i=1}^n f_0(x_i)\Big\}\,dx_1\ldots dx_n \leq$$

$$\leq \int_{\mathbb{R}^n} d_c(\underline{x})\Big\{\prod_{i=1}^n f_1(x_i) - c\prod_{i=1}^n f_0(x_i)\Big\}\,dx_1\ldots dx_n$$

$$= E_{f_1}\{d_c(\underline{X})\} - c\,E_{f_0}\{d_c(\underline{X})\} = \beta_{d_c}(f_1) - c\,\beta_{d_c}(f_0) \Rightarrow$$

$$\Rightarrow \beta_d(f_1) - \beta_{d_c}(f_1) \leq c\left[\beta_d(f_0) - \beta_{d_c}(f_0)\right] \leq 0$$

Εστω λοιπον, $X_1, \dots, X_n$ α.ι. $f(\cdot|\theta), \theta \in \Theta$, και εστω οτι ενδιαφερομεθα για τον ελεγχο της υποθεσεως $H: \theta = \theta_0$ vs $K: \theta = \theta_1$, οπου οι τιμες $\theta_0, \theta_1 \in \Theta$ ειναι συγκεκριμμενες και αρα οι $f_i \equiv f(\cdot|\theta_i)$, $i=1,2$ απολυτα καθορισμενες. Προκαθοριζουμε το μεγεθος του ελεγχου στο επιπεδο $\alpha \in [0,1]$. Τοτε, βασει του Λημματος των Neyman και Pearson ο πλεον ισχυρος ελεγχος μεγεθους α της $H: \theta = \theta_0$ vs $K: \theta = \theta_1$ ειναι ο εξης:

$$(4.23) \quad d_c(\underline{X}) \equiv \begin{cases} 1 & \text{αν } \dfrac{L_n(\theta_1|\underline{X})}{L_n(\theta_0|\underline{X})} > c \\ \gamma & \text{αν } \quad \text{''} \quad = c \\ 0 & \text{αν } \quad \text{''} \quad < c \end{cases},$$

και οι σταθερες $c \geq 0$, $\gamma \in [0,1]$ καθοριζονται απο την εξισωση,

$$(4.24) \quad \beta_{d_c}(\theta_0) = E_{\theta_0}\{d_c(\underline{X})\} = P_{\theta_0}\left\{\frac{L_n(\theta_1|\underline{X})}{L_n(\theta_0|\underline{X})} > c\right\} + \gamma P_{\theta_0}\left\{\frac{L_n(\theta_1|\underline{X})}{L_n(\theta_0|\underline{X})} = c\right\} = \alpha.$$

Θυμιζουμε οτι $L_n(\theta|\underline{X}) := \prod_{i=1}^{n} f(X_i|\theta)$, $\theta \in \Theta$ ειναι η πιθανοφανεια της $\theta \in \Theta$ στο δειγμα $\underline{X}$.

Οταν τα υποσυνολα Θ_0 ή Θ_1 των H, K ειναι μονοσυνολα οι αντιστοιχες υποθεσεις λεγονται <u>απλες</u>, αλλιως λεγονται <u>συνθετες</u>. Παρατηρουμε δε οτι το Λημμα των Neyman και Pearson αφορα στον ελεγχο <u>απλης</u> υποθεσεως H εναντι <u>απλης</u> εναλλακτικης K. Θα δουμε ομως αργοτερα οτι κατω απο ορισμενες προϋποθεσεις (π.χ., μονοτονια του πηλικου πιθανοφανειων) μπορει να επεκταθη η χρηση της ελεγχοσυναρτησης (4.23) στον ελεγχο συνθετων υποθεσεων, π.χ., $H: \theta \leq \theta_0$ vs $K: \theta > \theta_0$.

Θα παραθεσουμε τωρα ορισμενα βασικα παραδειγματα εφαρμογων του Λημματος των Neyman και Pearson.

(4.25) Παραδειγμα. Εστω $X_1, \ldots, X_n$ α.ι. $N(\theta, \sigma_0^2)$, $\theta \in \Theta = \mathbb{R}$, $\sigma_0^2 > 0$ - γνωστη σταθερα. Μας ενδιαφερει να κατασκευασουμε τον πλεον ισχυρο ελεγχο μεγεθους $\alpha = 0.05$ της υποθεσεως $H: \theta = \theta_0$ vs $K: \theta = \theta_1$, με $\theta_0 < \theta_1$ συγκεκριμμενες σταθερες.

Απαντ. Απο το Λημμα των Neyman & Pearson ο πλεον ισχυρος ελεγχος της H vs K βασιζεται στο πηλικο των πιθανοφανειων $L_n(\theta_1)$, $L_n(\theta_0)$, οπου,

$$L_n(\theta) = \prod_{i=1}^n f(X_i|\theta) = \prod_{i=1}^n \frac{1}{\sigma_0} \phi\left(\frac{X_i - \theta}{\sigma_0}\right) =$$

$$= (2\pi\sigma_0^2)^{-n/2} \exp\left\{-\frac{1}{2\sigma_0^2} \sum_{i=1}^n (X_i - \theta)^2\right\}.$$

Αρα,

$$\frac{L_n(\theta_1)}{L_n(\theta_0)} > c \iff \exp\left\{-\frac{1}{2\sigma_0^2} \sum_{i=1}^n \left[(X_i - \theta_1)^2 - (X_i - \theta_0)^2\right]\right\} > c$$

$$\iff -\sum_{i=1}^n \left[(X_i - \theta_1)^2 - (X_i - \theta_0)^2\right] > 2\sigma_0^2 \log c =: c'$$

$$\iff (\theta_1 - \theta_0) \sum_{i=1}^n (2X_i - \theta_0 - \theta_1) > c'$$

$$\iff \sum_{i=1}^n X_i > \frac{1}{2}\left[\frac{c'}{\theta_1 - \theta_0} + n(\theta_1^2 - \theta_0^2)\right] =: c''$$

$\iff \bar{X}_n > c^* := c''/n$. Στο εξης θα συμβολιζουμε ολες τις σταθερες που χρησιμοποιουμε διαδοχικα με c, δηλαδη,

$$\frac{L_n(\theta_1)}{L_n(\theta_0)} > c \iff \bar{X}_n > c \quad \text{για καποια (αλλη) σταθερα } c,$$

το ουσιαστικο ειναι η φορα της ανισοτητος και η στατιστικη $T(\underline{X}) = \bar{X}_n$ του ελεγχου.

Αρα, απο το Λημμα των Neyman-Pearson ο πλεον ισχυρος ελεγχος της $H: \theta = \theta_0$ vs $K: \theta = \theta_1$, με $\theta_0 < \theta_1$, ειναι ο εξης:

$$d_c(\underline{X}) = \begin{cases} 1 & \text{αν } \bar{X}_n > c \\ 0 & \text{αν } \bar{X}_n \leq c \end{cases}$$

Το να θεσουμε $d_c(\underline{X}) = \gamma \in [0,1)$ αν $\bar{X}_n = c$

δεν εχει εννοια διοτι $P_\theta(\bar{X}_n = c) = 0$, δεδομενου οτι η κατανομη της $\bar{X}_n$ ειναι συνεχης. Αποτελει δε πρακτικη συμβαση, σε μια τετοια περιπτωση, να θετουμε $\gamma = 0$. Η ευρεση της κριτιμης σταθερας c επιτυγχανεται μεσω της (4.24), η οποια εξασφαλιζει το μεγεθος του ελεγχου: εχουμε, δηλαδη,

$$\alpha = \beta_{d_c}(\theta_0) = P_{\theta_0}(\bar{X}_n > c) =$$

$$= P_{\theta_0}\left(\frac{\sqrt{n}(\bar{X}_n - \theta_0)}{\sigma_0} > \frac{\sqrt{n}(c - \theta_0)}{\sigma_0} \right) =$$

$$= 1 - \Phi\left(\frac{\sqrt{n}(c-\theta_0)}{\sigma_0} \right),$$

διοτι, υπο την υποθεση $H: \theta = \theta_0$, $\bar{X}_n \sim N(\theta_0, \frac{\sigma_0^2}{n})$.

Αρα,

$$\Phi\left(\frac{\sqrt{n}(c-\theta_0)}{\sigma_0} \right) = 1 - \alpha \iff \frac{\sqrt{n}(c-\theta_0)}{\sigma_0} = Z(1-\alpha)$$

$$\iff c \equiv c_\alpha = \theta_0 + \frac{\sigma_0 Z(1-\alpha)}{\sqrt{n}}$$

$$= \theta_0 + \frac{\sigma_0 \, 1.645}{\sqrt{n}} \quad \text{για } \alpha = 0.05.$$

Ο πλεον ισχυρος λοιπον ελεγχος μεγεθους α της $H: \theta = \theta_0$ vs $K: \theta = \theta_1$, ειναι ο εξης:

$$d(\underline{X}) = 1\!\!1\left(\bar{X}_n > \theta_0 + \frac{\sigma_0 Z(1-\alpha)}{\sqrt{n}}\right) = 1\!\!1\left(\frac{\sqrt{n}(\bar{X}_n - \theta_0)}{\sigma_0} > Z(1-\alpha) \right),$$

με ισχυ:

$$\beta_d(\theta_1) = P_{\theta_1}\left(\bar{X}_n > \theta_0 + \frac{\sigma_0 Z(1-\alpha)}{\sqrt{n}} \right) = 1 - \Phi\left(Z(1-\alpha) - \frac{\sqrt{n}(\theta_1 - \theta_0)}{\sigma_0} \right) \uparrow_{\theta_1}.$$

Η p-τιμη του ελεγχου στο δειγμα $\underline{x}$ ειναι η εξης:

$$\alpha(\bar{x}_n) = P_{\theta_0}(\bar{X}_n > \bar{x}_n) = P_{\theta_0}\left(\frac{\sqrt{n}(\bar{X}_n - \theta_0)}{\sigma_0} > \frac{\sqrt{n}(\bar{x}_n - \theta_0)}{\sigma_0} \right)$$

$$= 1 - \Phi\left(\frac{\sqrt{n}(\bar{x}_n - \theta_0)}{\sigma_0} \right).$$

<u>(4.26) Ασκηση.</u> Δειξτε οτι, αν στο παραδειγμα (4.25) $\theta_0 > \theta_1$, τοτε ο πλεον ισχυρος ελεγχος μεγεθους α ειναι ο ακολουθος:

$d(\underline{X}) = 1\!\!1\left(\bar{X}_n < \theta_0 - \frac{\sigma_0 Z(1-\alpha)}{\sqrt{n}}\right)$, και υπολογιστε την ισχυ του, και την p-τιμη του στο δειγμα $\underline{x}$.

(4.27) Παραδειγμα. Εστω $X_1, \dots, X_n$ α.ι. $\mathcal{P}(\lambda)$, $\lambda \in \Theta = (0, +\infty)$. Μας ενδιαφερει να κατασκευασουμε τον πλεον ισχυρο ελεγχο της $H: \lambda = \lambda_0$ vs $K: \lambda = \lambda_1$, οπου $\lambda_0 < \lambda_1$, μεγεθους α.

Απαντ. $L_n(\lambda) = \prod_{i=1}^{n} p(x_i \mid \lambda) = \prod_{i=1}^{n} \frac{e^{-\lambda} \lambda^{x_i}}{x_i!}$

$$= e^{-n\lambda} \lambda^{n\bar{x}_n} \Big/ \prod_{i=1}^{n}(x_i!) .$$

Αρα, $\frac{L_n(\lambda_1)}{L_n(\lambda_0)} > c \iff$

$$\iff \frac{e^{-n\lambda_1} \lambda_1^{n\bar{x}_n}}{e^{-n\lambda_0} \lambda_0^{n\bar{x}_n}} > c \iff \left(\frac{\lambda_1}{\lambda_0}\right)^{n\bar{x}_n} > c \text{, καθοια (αλλη) σταθερα,}$$

$$\iff n\bar{x}_n \log\left(\frac{\lambda_1}{\lambda_0}\right) > c \iff \Sigma_{i=1}^{n} X_i > c ,$$

Αρα, απο το Λημμα των Neyman-Pearson, ο πλεον ισχυρος ελεγχος μεγεθους α, ειναι ο ακολουθος:

$$d(\underline{x}) = \begin{cases} 1 & \text{αν } \Sigma_{i=1}^{n} X_i > c_\alpha \\ \gamma & \text{αν } \Sigma_{i=1}^{n} X_i = c_\alpha \\ 0 & \text{αν } \Sigma_{i=1}^{n} X_i < c_\alpha \end{cases} ,$$

οπου, οι c_α, γ υπολογιζονται απο την εξισωση (4.24), δηλαδη,

$$\alpha = \beta_d(\lambda_0) = P_{\lambda_0}\left(\Sigma_{i=1}^{n} X_i > c_\alpha \right) + \gamma P_{\lambda_0}\left(\Sigma_{i=1}^{n} X_i = c_\alpha \right),$$

οπου, υπο την $H: \lambda = \lambda_0$, $\Sigma_{i=1}^{n} X_i \sim \mathcal{P}(n\lambda_0)$.

Αρα,

$$\alpha = \Sigma_{k=c_\alpha+1}^{\infty} \frac{e^{-n\lambda_0} (n\lambda_0)^k}{k!} + \gamma \frac{e^{-n\lambda_0} (n\lambda_0)^{c_\alpha}}{c_\alpha!} ,$$

με $c_\alpha \in \mathbb{N}_0 \equiv \{0, 1, 2, \dots\}$, $\gamma \in [0, 1)$.

Εχουμε, λοιπον,

$$(4.28) \quad 0 \leq \gamma = \frac{\alpha - \sum_{k=c_\alpha+1}^{\infty} \frac{e^{-n\lambda_0} (n\lambda_0)^k}{k!}}{\frac{e^{-n\lambda_0} (n\lambda_0)^{c_\alpha}}{c_\alpha!}} < 1 . \iff$$

$$(4.29) \quad \sum_{k=c_\alpha+1}^{\infty} \frac{e^{-n\lambda_0} (n\lambda_0)^k}{k!} \leq \alpha < \sum_{k=c_\alpha}^{\infty} \frac{e^{-n\lambda_0} (n\lambda_0)^k}{k!} .$$

Αρκει, λοιπον, να βρουμε το <u>μοναδικο</u> c_α που ικανοποιει την (4.29), το δε γ υπολογιζεται απο την (4.28).
[Παρατηρουμε οτι το να εχουμε $\gamma=1$ αντιστοιχει σε αλλαγη της κρισιμης σταθερας απο c_α σε $c_\alpha - 1$, ενχερεια που ηδη εχουμε θεωροντας $\gamma \in [0,1)$.]

Για παραδειγμα, εστω $\lambda_0 = 0.1$, $n=10$, $\alpha=0{,}05$, τοτε

$$\sum_{k=4}^{\infty} \frac{e^{-n\lambda_0}(n\lambda_0)^k}{k!} = 0.01334 < \alpha = 0{,}05 < \sum_{k=3}^{\infty} \frac{e^{-n\lambda_0}(n\lambda_0)^k}{k!} = 0{,}075$$

και αρα $c_\alpha = 3$ για $\alpha = 0{,}05$ και

$$\gamma = \frac{\alpha - \sum_{k=4}^{\infty} \frac{e^{-1}1^k}{k!}}{\frac{e^{-1}1^3}{3!}} = \frac{0{,}05 - 0{,}01334}{0.061666} = 0.5945 .$$

Δηλαδη, ο ελεγχος της $H: \lambda = 0.1$ vs $K: \lambda = \lambda_1$, με $\lambda_1 > 0.1$, μεγεθους $\alpha = 0{,}05$, βασισμενος σε ενα δειγμα $X_1, \ldots, X_{10}$ ειναι ο εξης:

$$d(\underline{X}) = \begin{cases} 1 & \text{αν } \sum_{i=1}^{10} X_i > 3 \\ 0.5945 & \text{αν } \sum_{i=1}^{10} X_i = 3 \\ 0 & \text{αν } \sum_{i=1}^{10} X_i < 3 \end{cases},$$

δηλαδη, αν $T(\underline{x}) = \sum_{i=1}^{10} x_i \in \{0,1,2\} = A_d$ αποδεχομεθα την H ως αληθη, αν $T(\underline{x}) \in \{4,5,\ldots\}$ απορριπτουμε την H, αν ομως $T(\underline{x}) = 3$ τοτε παιρνουμε μια παρατηρηση απο μια τ.μ. $\xi \sim \text{Bernoulli}(p=0{,}5945)$ και αν ελθη $\xi=0$ δεχομεθα την H αν δε ελθη $\xi=1$ τοτε απορριπτουμε την H.

Η <u>ισχυς του τυχαιοποιημενου</u> ελεγχου d ειναι η ακολουθη:

$$\begin{aligned} \beta_d(\lambda_1) &= P_{\lambda_1}\left(\sum_{i=1}^{10} X_i > 3\right) + \gamma P_{\lambda_1}\left(\sum_{i=1}^{10} X_i = 3\right) \\ &= \sum_{k=4}^{\infty} e^{-10\lambda_1}(10\lambda_1)^k/k! + 0.5945\, e^{-10\lambda_1}(10\lambda_1)^3/3! \\ &= 1 - e^{-10\lambda_1}\left\{1 + 10\lambda_1 + (10\lambda_1)^2/2 + 0.4055\,(10\lambda_1)^3/6\right\} \uparrow_{\lambda_1} \end{aligned}$$

Παρατηρουμε οτι οι μη τυχαιοποιημενοι ελεγχοι $d_1(\underline{X}) := 1(\sum_{i=1}^{10} X_i > 3)$ και $d_2(\underline{X}) := 1(\sum_{i=1}^{10} X_i > 4)$ εχουν μεγεθη 0.01334 και 0.075 αντιστοιχως, και ισχυ $\beta_{d_1}(\lambda_1) < \beta_d(\lambda_1) < \beta_{d_2}(\lambda_1)$. Στην επιθυμια μας λοιπον να πετυχουμε μεγιστη ισχυ χωρις να ξεπερασουμε σε μεγεθος το επιπεδο $\alpha = 0.05$ οδηγηθηκαμε στον τυχαιοποιημενο ελεγχο d

Η p-τιμη του ελεγχου d στο δειγμα $\underline{x}$, με $\Sigma_1^{10} x_i = 3$ - ας πουμε -, ειναι

$$P_{\lambda_0}\left(\Sigma_{i=1}^{n} X_i > n\bar{x}_n \right) = P_{\lambda_0}\left(\Sigma_{i=1}^{10} X_i > 3 \right) =$$

$$= P_{\lambda_0}\left(\Sigma_{i=1}^{10} X_i \geq 4 \right) = \sum_{k=4}^{\infty} \frac{e^{-1} 1^k}{k!} = 0.01334 ,$$

και αρα με το δειγμα που πηραμε, ο ελεγχος d απορριπτει την H για καθε επιπεδο $\alpha > 0.01334$ και την δεχεται ως αληθη $\forall \alpha \leq 0.01334$.

(4.30) <u>Ασκηση</u>. Επαναλαβατε το παραδειγμα (4.27) για $\lambda_1 < \lambda_0$.

(4.31) <u>Ασκηση</u>. Εστω $X_1, \ldots, X_n$ α.ι. απο την κατανομη:

(α) $\Gamma(\alpha_0, \theta)$, $\theta \in \Theta = (0, +\infty)$, $\alpha_0 > 0$ γνωστο,

(β) Weibull(α_0, θ), $\theta \in \Theta = (0, +\infty)$, $\alpha_0 > 0$ γνωστο,

(γ) $\mathcal{U}(0, \theta)$, $\theta \in \Theta = (0, +\infty)$,

(δ) $\mathcal{U}(\theta, 1)$, $\theta \in \Theta = (-\infty, 1)$,

(ε) Bernoulli(θ), $\theta \in \Theta = (0, 1)$,

(ς) Γεωμετρικη(θ), $\theta \in \Theta = (0, 1)$.

(i) Κατασκευασατε τον πλεον ισχυρο ελεγχο μεγεθους $\alpha \in (0,1)$, της υποθεσεως $H: \theta = \theta_0$ vs $K: \theta = \theta_1$, με $\theta_0 < \theta_1$.

(ii) Επαναλαβατε το μερος (i) για $\theta_0 > \theta_1$.

(iii) Υπολογιστε την ισχυ των ελεγχων των μερων (i) και (ii).

(iv) Ειδικευστε τα αποτελεσματα των μερων (i), (ii) κ (iii) στην περιπτωση που $n = 64$, $\alpha = 0.05$, $\theta_0 = \frac{1}{2}$. Στην περιπτωση της κατανομης (α) θεστε $\alpha_0 = 3$ και στην περιπτωση της (β) θεστε $\alpha_0 = 1$.

(v) Βρητε τις p-τιμες των ελεγχων των μερων (i) και (ii), με n, α, θ_0 και α_0 - οπου χρειαζεται - οπως στο μερος (iv), αν το δειγμα που πηρατε, στην περιπτωση της (α) εχει $\bar{x}=4$, των (β) και (ζ) $\bar{x}=1.5$ της (ε) $\bar{x}=0.75$, στις δε περι των κατανομων (γ) και (δ) πηρατε $x_{(n)} \equiv x_{n1} = 0,4$ και $x_{(1)} \equiv x_{n1} = 0,6$ αντιστοιχως.

Θα ασχοληθουμε τωρα με την κατασκευη ομοιομορφα πλεον ισχυρων (ΟΠΙ) ελεγχων για συνθετες μονοπλευρες υποθεσεις της μορφης $H: \theta \le \theta_0$ vs $K: \theta > \theta_0$. Εχουμε, δηλαδη, $\Theta \subseteq \mathbb{R}$, $\Theta_0 = (-\infty, \theta_0]$, $\Theta_1 = (\theta_0, +\infty)$, ζητουμε δε να κατασκευασουμε ελεγχο d, μεγεθους $\alpha \in (0,1)$ — δηλαδη, $\sup\{\beta_d(\theta), \theta \in \Theta_0\} = \alpha$ — τετοιον ωστε $\beta_d(\theta) \geq \beta_{d'}(\theta)$ $\forall \theta \in \Theta_1$ για καθε ελεγχο d' μεγεθους μικροτερου ή ισου του α.

Θα εξηγησουμε, αρχικα, την κλασσικη μεθοδο κατασκευης τετοιων ελεγχων με ενα παραδειγμα το οποιο επεκτεινει την μεθοδολογια του (4.25) — και (4.27) — στην κατασκευη ελεγχων συνθετων υποθεσεων.

(4.32) Παραδειγμα. Εστω $X_1, \ldots, X_n$ α.ι. $N(\theta, \sigma_0^2)$, $\theta \in \Theta = \mathbb{R}$, $\sigma_0^2 > 0$ - γνωστο. Δειξτε οτι ο ελεγχος

$$d(\underline{X}) := 1\left(\bar{X}_n > c_\alpha \equiv \theta_0 + \frac{\sigma_0 z(1-\alpha)}{\sqrt{n}}\right)$$ ειναι ΟΠΙ

μεγεθους $\alpha \in (0,1)$ για τον ελεγχο της $H: \theta \le \theta_0$ vs $K: \theta > \theta_0$. (Παρατ. Η ελεγχοσυναρτηση d ειναι η αυτη με την του (4.25).).

Απαντ. Θα βρουμε κατ' αρχας την μορφη της ελεγχοσυναρτησης: Παρατηρουμε οτι $\forall$ $\underline{\theta} \in \Theta_0$ και $\forall$ $\bar{\theta} \in \Theta_1$, ο πλεον ισχυρος ελεγχος της $H': \theta = \underline{\theta}$ vs $K': \theta = \bar{\theta}$, ειναι της μορφης $d_c(\underline{X}) = 1(\bar{X}_n > c)$, επειδη και μονο $\underline{\theta} < \bar{\theta}$ και ανεξαρτητα απο τις συγκεκριμενες τιμες των $\underline{\theta}, \bar{\theta}$ — βλ. (4.25) — εχει δε μεγεθος $\beta_{d_c}(\underline{\theta}) =$

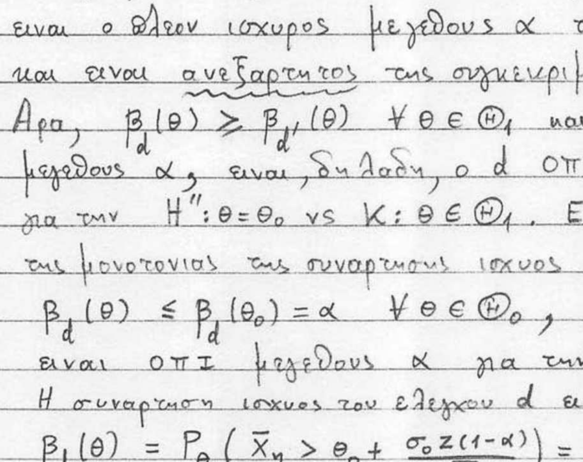
$P_{\underline{\theta}}(\bar{X}_n > c) = 1 - \Phi\left(\frac{\sqrt{n}(c-\underline{\theta})}{\sigma_0}\right)$ αυξουσα συναρτηση του $\underline{\theta}$.

Αρα, $\sup\{\beta_{d_c}(\theta), \theta \in \Theta_0\} = P_{\theta_0}(\bar{X}_n > c) =$

$= 1 - \Phi\left(\frac{\sqrt{n}(c-\theta_0)}{\sigma_0}\right) = \alpha \Leftrightarrow c = c_\alpha = \theta_0 + \frac{\sigma_0 z(1-\alpha)}{\sqrt{n}}$.

Εχουμε, λοιπον, οτι $\forall \bar{\theta} \in \Theta_1$, ο ελεγχος d ειναι ο Β.Ι.ΕΟΝ ισχυρος μεγεθους α της $H'': \theta = \theta_0$ vs $K': \theta = \bar{\theta}$ και ειναι ανεξαρτητος της συγκεκριμμενης τιμης του $\bar{\theta} \in \Theta_1$. Αρα, $\beta_d(\theta) \geq \beta_{d'}(\theta)$ $\forall \theta \in \Theta_1$ και καθε ελεγχο d' μεγεθους α, ειναι, δηλαδη, ο d ΟΠΙ μεγεθους α για την $H'': \theta = \theta_0$ vs $K: \theta \in \Theta_1$. Εφοσον δε — λογω της μονοτονιας της συναρτησης ισχυος του d — εχουμε:
$\beta_d(\theta) \leq \beta_d(\theta_0) = \alpha$ $\forall \theta \in \Theta_0$, ο ελεγχος d ειναι ΟΠΙ μεγεθους α για την $H: \theta \in \Theta_0$ vs $K: \theta \in \Theta_1$.
Η συναρτηση ισχυος του ελεγχου d ειναι η ακολουθη:

$$\beta_d(\theta) = P_\theta\left(\bar{X}_n > \theta_0 + \frac{\sigma_0 z(1-\alpha)}{\sqrt{n}}\right) =$$
$$= 1 - \Phi\left(z(1-\alpha) - \frac{\sqrt{n}(\theta-\theta_0)}{\sigma_0}\right), \quad \theta \in \Theta = \mathbb{R}.$$



(4.33) Ασκηση. Εστω $X_1, \ldots, X_n$ α.ι. $N(\theta, \sigma_0^2)$, $\sigma_0^2 > 0$ - γνωστο. Δειξτε οτι ο ελεγχος $d(\underline{X}) = 1\left(\bar{X}_n < \theta_0 - \frac{\sigma_0 z(1-\alpha)}{\sqrt{n}}\right)$, ειναι ΟΠΙ μεγεθους α για την $H: \theta \geq \theta_0$ vs $K: \theta < \theta_0$, εχει δε συναρτηση ισχυος:

$$\beta_d(\theta) = 1 - \Phi\left(z(1-\alpha) - \frac{\sqrt{n}(\theta_0-\theta)}{\sigma_0}\right), \quad \theta \in \Theta = \mathbb{R}.$$

(4.34) <u>Ορισμος</u>. Εστω $X_1, \dots, X_n$ α.ι. με πυκνοτητα ή σ.μ.π. $f(\cdot|\theta)$, $\theta \in \Theta \subseteq \mathbb{R}$, και εστω οτι η παραμετριση της οικογενειας αυτων των κατανομων ειναι προσδιορισιμη – βλ.(2.2). Λεμε οτι η οικογενεια $\{f(\cdot|\theta), \theta \in \Theta\}$ αυτων των κατανομων εχει <u>μονοτονο πηλικο πιθανοφανειων</u> (ΜΠΠ), αν υπαρχει στατιστικη $T(\underline{X})$, τετοια ωστε, αν $\theta_0 < \theta_1$ τοτε το πηλικο $\dfrac{L(\theta_1|\underline{X})}{L(\theta_0|\underline{X})}$ ειναι μονοτονη συναρτηση της $T(\underline{X})$ – αυξουσα ή φθινουσα. Εχουμε θεσει $L(\theta|\underline{X}) := \prod_{i=1}^{n} f(X_i|\theta)$, $\theta \in \Theta$.

Αυτη ακριβως η ιδιοτητα της ΜΠΠ, σε συνδιασμο με το Λημμα των Neyman και Pearson, εκανε δυνατη την κατασκευη του ΟΠΙ ελεγχου του (4.32), καθως και της (4.33). Γενικοτερα, εχουμε το ακολουθο αποτελεσμα:

(4.35) <u>Προταση</u>. Εστω $X_1, \dots, X_n$ α.ι. $f(\cdot|\theta)$, $\theta \in \Theta \subseteq \mathbb{R}$, με ΜΠΠ ως προς τη στατιστικη $T(\underline{X})$. Για τον ελεγχο της υποθεσεως $H: \theta \le \theta_0$ vs $K: \theta > \theta_0$, υπαρχει ΟΠΙ μεγεθου $\alpha \in (0,1)$ και ειναι ο εξης:

(i) $d(\underline{x}) = 1(T(\underline{x}) > c_\alpha) + \gamma\, 1(T(\underline{x}) = c_\alpha)$,
  αν το ΜΠΠ ειναι αυξουσα συναρτηση της $T(\underline{x})$,

(ii) $d(\underline{x}) = 1(T(\underline{x}) < c_\alpha) + \gamma\, 1(T(\underline{X}) = c_\alpha)$,
  αν το ΜΠΠ ειναι φθινουσα συναρτηση της $T(\underline{x})$,

εφοσον υπαρχουν: $\theta_* \in \Theta_0 = (-\infty, \theta_0]$ και σταθερα c_α, τετοια ωστε

(4.36) $\beta_d(\theta) \equiv E_\theta\{d(\underline{X})\} \le \beta_d(\theta_*) = \alpha \quad \forall \theta \in \Theta_0$.

<u>Σημ.1</u>. Συνηθως η συναρτηση $\beta_d(\cdot)$ ειναι αυξουσα και η (3.36) επαληθευεται με $\theta_* = \theta_0$ – βλ. (4.32).

<u>Σημ.2</u>. Για τον ελεγχο της $H: \theta \ge \theta_0$ vs $K: \theta < \theta_0$, οι ρολοι των ελεγχων (i), (ii) αντιστρεφονται, η δε $\beta_d(\cdot)$ ειναι τωρα φθινουσα

και επαληθευεται παλι για $\theta_* = \theta_0$ – βλ. (4.33).

<u>Αποδ.</u> Θα κατασκευασουμε τον ελεγχο για την περιπτωση (i), στις περιπτωσεις της (ii) και της Σημ. 2 τα πραγματα ειναι αναλογα.

Εστω, λοιπον, οτι αν $\theta_0 < \theta_1$ τοτε $\dfrac{L(\theta_1|\underline{x})}{L(\theta_0|\underline{x})} = g(T(\underline{x}))$ – μια <u>αυξουσα</u> συναρτηση της $T(\underline{x})$.

Τοτε, $\forall\, \bar{\theta} \in \Theta_1 = (\theta_0, +\infty)$ και για το θ_* της (4.36), ο <u>πλεον ισχυρος</u> ελεγχος μεγεθους α της απλης vs απλης : $H': \theta = \theta_*$ vs $K': \theta = \bar{\theta}$ διδεται απο το Λημμα των Neyman και Pearson και εχει ως εξης:

$$d_c(\underline{x}) = 1\left(\frac{L(\bar{\theta}|\underline{x})}{L(\theta_*|\underline{x})} > c'\right) + \gamma 1\left(\frac{L(\bar{\theta}|\underline{x})}{L(\theta_*|\underline{x})} = c'\right)$$

$$= 1\big(g(T(\underline{x})) > c'\big) + \gamma 1\big(g(T(\underline{x})) = c'\big)$$

$$= 1\big(T(\underline{x}) > c\big) + \gamma 1\big(T(\underline{x}) = c\big),$$

για καποια σταθερα c, εφοσον η $g(\cdot)$ ειναι αυξουσα συναρτηση, ειναι δε <u>ανεξαρτητος</u> της συγκεκριμενης τιμης της $\bar{\theta} \in \Theta_1$ – αρκει που $\bar{\theta} > \theta_*$. Η συγκεκριμενη τιμη c_α της c και η τιμη της γ υπολογιζονται, βασει της (4.36) (και εξαρτωνται απο την θ_*), ως εξης:

(4.37) $\beta_d(\theta_*) = P_{\theta_*}(T(\underline{x}) > c_\alpha) + \gamma P_{\theta_*}(T(\underline{x}) = c_\alpha) = \alpha$

με,

(4.38) $$0 \le \gamma = \frac{\alpha - P_{\theta_*}(T(\underline{x}) > c_\alpha)}{P_{\theta_*}(T(\underline{x}) = c_\alpha)} < 1,$$

εφοσον $P_{\theta_*}(T(\underline{x}) = c_\alpha) \neq 0$, αλλως $\gamma \equiv 0$.

Εχουμε λοιπον δειξει οτι ο ελεγχος

$$d(\underline{x}) = 1\big(T(\underline{x}) > c_\alpha\big) + \gamma 1\big(T(\underline{x}) = c_\alpha\big),$$

με c_α, γ υπολογιζομενα μονοσημαντα μεσω των (4.37) & (4.38), ειναι ο πλεον ισχυρος μεγεθους α για την $H': \theta = \theta_*$ vs $K': \theta = \bar{\theta}$, δηλαδη, $\beta_d(\bar{\theta}) \geq \beta_{d'}(\bar{\theta})$ για καθε ελεγχο d' με $\beta_{d'}(\theta_*) \le \beta_d(\theta_*) = \alpha$.

Αλλα, ο d δεν εξαρταται απο την συγκεκριμενη

τιμη της $\bar{\theta} \in \Theta_1$, αρα $\beta_d(\theta) \geq \beta_{d'}(\theta)$ $\forall \theta \in \Theta_1$
για καθε ελεγχο d' με $\beta_{d'}(\theta_*) \leq \beta_d(\theta_*) = \alpha$,
ειναι, δηλαδη, ο ελεγχος d ΟΠΙ μεγεθους α
για την $H': \theta = \theta_*$ vs $K: \theta \in \Theta_1 = (\theta_0, +\infty)$.
Τωρα, απο την (4.36) $\beta_d(\theta) \leq \alpha$ $\forall$ $\theta \in \Theta_0 = (-\infty, \theta_0]$
και αρα ο ελεγχος d ειναι ΟΠΙ μεγεθους α
για την $H: \theta \in \Theta_0$ vs $K: \theta \in \Theta_1$.

(4.39) <u>Παρατηρηση</u>. Η μονοπαραμετρικη εκθετικη οικογενεια
$f(x|\theta) = \exp\{c(\theta)T(x) + d(\theta) + S(x)\}\, 1(x \in A)$, $\theta \in \Theta \subseteq \mathbb{R}$,
με $c(\cdot)$ μονοτονη συναρτηση εχει ΜΠΠ ως προς την
στατιστικη συναρτηση $T_n(\underline{X}) = \sum_{i=1}^n T(X_i)$, και αρα οι ελεγχοι
(i), (ii) της (4.35) ειναι ΟΠΙ μεγεθους $\alpha \in (0,1)$ για
την $H: \theta \leq \theta_0$ vs $K: \theta > \theta_0$ (ή τις αντιστοιχες της Συμ.2),
εφοσον μπορει να επαληθευθη και η (4.36).

Για να διαπιστωσουμε την ΜΠΠ της οικογενεια αυτης,
παρατηρουμε οτι, για $\theta_0 < \theta_1$, εχουμε:

$$\frac{L(\theta_1|\underline{X})}{L(\theta_0|\underline{X})} = \exp\{[c(\theta_1) - c(\theta_0)]\, T_n(\underline{X}) + d(\theta_1) - d(\theta_0)\}\, 1(\underline{X} \in A^n),$$

το οποιο ειναι αυξουσα (φθινουσα) συναρτηση της $T_n(\underline{X})$,
αν η $c(\theta)$ ειναι αυξουσα (φθινουσα) συναρτηση της θ,
διοτι τοτε $c(\theta_1) - c(\theta_0) \geq 0$ (≤ 0).

(4.40) <u>Παραδειγμα</u>. Εστω $X_1, \ldots, X_n$ α.ι. Poisson(θ), $\theta > 0$.
Δειξτε οτι για τον ελεγχο της $H: \theta \leq \theta_0$ vs $K: \theta > \theta_0$,
ο ελεγχος $d(\underline{x}) = 1(\sum_{i=1}^n x_i > c_\alpha) + \gamma\, 1(\sum_{i=1}^n x_i = c_\alpha)$,
οπου οι σταθερες $c_\alpha \in \mathbb{N}_0$ και $\gamma \in [0,1)$ υπολογιζονται
(μονοσημαντα) απο τις σχεσεις: $0 \leq \gamma < 1$,
$P_{\theta_0}(\sum_{i=1}^n X_i > c_\alpha) + \gamma\, P_{\theta_0}(\sum_{i=1}^n X_i = c_\alpha) = \alpha$,
ειναι ΟΠΙ μεγεθους α.

Απαντ. Θα δειξουμε οτι η οικογενεια των κατανομων Poisson εχει αυξον ΜΠΠ: εστω $\theta_0 < \theta_1$, τοτε

$$\frac{L(\theta_1|\underline{x})}{L(\theta_0|\underline{x})} = \frac{e^{-n\theta_1}\,\theta_1^{\Sigma_1^n x_i}/\prod_1^n x_i!}{e^{-n\theta_0}\,\theta_0^{\Sigma_1^n x_i}/\prod_1^n x_i!} =$$

$$= \exp\{[\log\theta_1 - \log\theta_0]\Sigma_1^n x_i + n\theta_0 - n\theta_1\},$$

το οποιο ειναι αυξουσα συναρτηση της $T(\underline{x}) = \Sigma_{i=1}^n x_i$.

Αρα, απο την (4.35) η μορφη του ΟΠΙ ελεγχου ειναι:

$d_c(\underline{x}) = 1(T(\underline{x}) > c) + \gamma\, 1(T(\underline{x}) = c)$, οπου οι $c \in \mathbb{N}_0$, $\gamma \in [0,1)$ υπολογιζονται απο τις (4.37) και (4.38).

Αρκει λοιπον να βρουμε το θ_* της (4.36): εχουμε,

$$\beta_{d_c}(\theta) = E_\theta\{d_c(\underline{x})\} = P_\theta(T(\underline{x}) > c) + \gamma P_\theta(T(\underline{x}) = c)$$

$$= \Sigma_{k=c+1}^{\infty} \frac{e^{-n\theta}(n\theta)^k}{k!} + \gamma \frac{e^{-n\theta}(n\theta)^c}{c!}$$

$$= 1 - \Sigma_{k=0}^{c-1} \frac{e^{-n\theta}(n\theta)^k}{k!} - (1-\gamma)\frac{e^{-n\theta}(n\theta)^c}{c!},$$

η οποια ειναι αυξουσα συναρτηση της θ, μια και

$$\beta'_{d_c}(\theta) = n e^{-n\theta}\left\{(1-\gamma)\frac{(n\theta)^c}{c!} + \gamma\frac{(n\theta)^{c-1}}{(c-1)!}\right\} > 0 \quad \forall\, \theta > 0.$$

Αρα, $\sup\{\beta_{d_c}(\theta),\ \theta \in \Theta_0 = (0,\theta_0]\} = \beta_{d_c}(\theta_0)$, δηλαδη, $\theta_* = \theta_0$, και οι (4.37) και (4.38) ειναι εδω οι εξης: $0 \le \gamma < 1$,

$$P_{\theta_0}(T(\underline{x}) > c_\alpha) + \gamma P_{\theta_0}(T(\underline{x}) = c_\alpha) = \alpha.$$

Για ενα παραδειγμα υπολογισμου συγκεκριμενων c_α, γ, βλεπε (4.27).

(4.41) Ασκηση. Δειξτε οτι, αν στην περιπτωση του παραδειγματος (4.40) μας ενδιαφερει ο ελεγχος της $H: \theta \ge \theta_0$ vs $K: \theta < \theta_0$, ο ΟΠΙ ελεγχος ειναι της μορφης $d(\underline{x}) = 1(\Sigma_1^n x_i < c_\alpha) + \gamma\, 1(\Sigma_1^n x_i = c_\alpha)$, οπου $0 \le \gamma < 1$, $P_{\theta_0}(\Sigma_1^n x_i < c_\alpha) + \gamma P_{\theta_0}(\Sigma_1^n x_i = c_\alpha) = \alpha$.

(4.42) <u>Παραδειγμα</u>: Εστω $X_1, \dots, X_n$ α.ι. $\mathcal{E}(\lambda)$, $\lambda > 0$.
Δειξτε οτι για τον ελεγχο της $H: \lambda \le \lambda_0$ vs $K: \lambda > \lambda_0$,
ο ελεγχος $d(\underline{X}) = 1(\Sigma_{i=1}^{n} X_i < c_\alpha)$ με $c_\alpha = \dfrac{n}{\lambda_0 f_{\infty, 2n}(1-\alpha)} = \dfrac{\chi^2_{2n}(\alpha)}{2\lambda_0}$
ειναι ΟΠΙ μεγεθους α.

<u>Απαντ.</u> Θα δειξουμε η οικογενεια κατανομων $\{\mathcal{E}(\lambda), \lambda > 0\}$
εχει φθινον ΜΠΠ : εστω $\lambda_0 < \lambda_1$, τοτε

$$\frac{L(\lambda_1|\underline{X})}{L(\lambda_0|\underline{X})} = \exp\{-(\lambda_1-\lambda_0)\Sigma_{i=1}^{n} X_i + n\log(\lambda_1/\lambda_0)\} 1(\underline{x} \in \mathbb{R}_+^n),$$

το οποιο ειναι <u>φθινουσα</u> συναρτηση της $T(\underline{X}) \equiv \Sigma_{i=1}^{n} X_i$,
και αρα απο την (4.35) ο ΟΠΙ ελεγχος μεγεθους α ειναι ο:

$$d(\underline{x}) = 1(T(\underline{x}) < c_\alpha),$$

($\gamma \equiv 0$, διοτι $T \sim \mathcal{G}(n, \lambda)$ ειναι συνεχης τ.μ.)
οπου η σταθερα c_α υπολογιζεται απο την (4.36), δηλαδη:

$$\alpha = \sup\{P_\lambda(T(\underline{X}) < c_\alpha), \lambda \le \lambda_0\} =$$
$$= \sup\{P(\chi^2_{2n} < 2\lambda c_\alpha), \lambda \le \lambda_0\}$$

(διοτι, $2\lambda T \sim \mathcal{G}(n, 1/2)$ – βλ. (1.35), (1.38β))

$$= P(\chi^2_{2n} < 2\lambda_0 c_\alpha) \Rightarrow 2\lambda_0 c_\alpha = \chi^2_{2n}(\alpha) = \frac{2n}{f_{\infty, 2n}(1-\alpha)}$$

(βλ. (3.184)) $\Rightarrow c_\alpha = \dfrac{n}{\lambda_0 f_{\infty, 2n}(1-\alpha)}$.

(4.43) <u>Ασκηση</u>: Εστω $X_1, \dots, X_n$ α.ι. $\mathcal{E}(\lambda)$, $\lambda > 0$. Δειξτε οτι
για τον ελεγχο της $H: \lambda \ge \lambda_0$ vs $K: \lambda < \lambda_0$,
ο ελεγχος $d(\underline{X}) = 1(\Sigma_{i=1}^{n} X_i > c_\alpha)$ με $c_\alpha = \chi^2_{2n}(1-\alpha)/2\lambda_0$.

(4.44) <u>Ασκηση</u>: Επαναλαβατε το (4.42) και την (4.43)
στην περιπτωση που οι $X_1, \dots, X_n$ ειναι α.ι. $\mathcal{G}(\alpha_0, \lambda)$,
$\lambda > 0$ και η παραμετρος $\alpha_0 > 0$ ειναι γνωστη.

(4.45) <u>Ασκηση</u>: Επαναλαβατε την (4.44) στην περιπτωση που
$X_1, \dots, X_n$ α.ι. Weibull(α_0, λ), $\lambda > 0$ και $\alpha_0 > 0$ γνωστο.

· (4.46) <u>Ασκηση</u>. Εστω $X_1, \dots, X_n$ α.ι. $\mathcal{E}(\frac{1}{\theta})$, $\theta>0$.
Βρητε το ΟΠΙ ελεγχο μεγεθους α για την $H:\theta\le\theta_0$ vs $K:\theta>\theta_0$.
Υπολογισε την συναρτηση ισχυος αυτου του ελεγχου.

(4.47) <u>Ασκηση</u>. Εστω $X_1, \dots, X_n$ α.ι. $\mathcal{U}(0,\theta)$, $\theta>0$.
Βρητε τον ΟΠΙ ελεγχο μεγεθους α για την
$H:\theta\le\theta_0$ vs $K:\theta>\theta_0$, και υπολογισε την ισχυ του.

· (4.48) <u>Ασκηση</u>. Εστω $X_1, \dots, X_n$ α.ι. Bernoulli(θ), $\theta\in[0,1]$.
Βρητε τον ΟΠΙ ελεγχο μεγεθους α για την
$H:\theta\ge\theta_0$ vs $K:\theta<\theta_0$, και υπολογισε την
συναρτηση ισχυος του.

Θα ασχοληθουμε τωρα με την κατασκευη ελεγχων για απλες υποθεσεις εναντι <u>αμφιπλευρων</u> εναλλακτικων, δηλαδη, $H:\theta=\theta_0$ vs $K:\theta\ne\theta_0$, με $\theta\in\Theta\subseteq\mathbb{R}$.
Εν γενει, οπως θα δουμε, δεν υπαρχουν ΟΠΙ ελεγχοι για αυτο το προβλημα, υπαρχουν ομως σε μερικες περιπτωσεις ΟΠΙ συμμετρικοι ή αμεροληπτοι ($\beta(\theta)\ge\beta(\theta_0)=\alpha\ \ \forall\,\theta\ne\theta_0$) ελεγχοι – βασικα για την εκθετικη οικογενεια – με τους οποιους ενδεικτικα μονο θα ασχοληθουμε εδω.

Συγκεκριμενα εστω $X_1, \dots, X_n$ α.ι. $f(\cdot|\theta)$, $\theta\in\Theta\subseteq\mathbb{R}$, και εστω οτι η οικογενεια αυτη των κατανομων εχει Μ.Π.Π. ως προς την στατιστικη $T(\underline{X})$. Μας ενδιαφερει ο ελεγχος της

(4.49) $\quad H:\theta=\theta_0$ vs $K:\theta\ne\theta_0$.

Θα υποθεσουμε οτι εχουμε <u>αυξον</u> ΜΠΠ (η αλλη περιπτωση αντιμετωπιζεται κατα αναλογο τροπο). Τοτε, ο ελεγχος

(4.50) $d_1(\underline{X}) := 1(T(\underline{X})>c_1) + \gamma_1\, 1(T(\underline{X})=c_1)$, οπου

$0\le\gamma_1<1$ και $\beta_{d_1}(\theta_0)=P_{\theta_0}(T(\underline{X})>c_1)+\gamma_1 P_{\theta_0}(T(\underline{X})=c_1)=\alpha_1$, ειναι – βασει της (4.35) – ΟΠΙ μεγεθους $\alpha_1 \in (0,1)$ για την $H: \theta=\theta_0$ vs $K_1: \theta>\theta_0$.

Επισης, ο ελεγχος

(4.51) $d_2(\underline{X}) := 1(T(\underline{X}) < c_2) + \gamma_2 1(T(\underline{X}) = c_2)$, οπου $0 \le \gamma_2 < 1$, $\beta_{d_2}(\theta_0) = P_{\theta_0}(T(\underline{X}) < c_2) + \gamma_2 P_{\theta_0}(T(\underline{X}) = c_2) = \alpha_2$,

ειναι ΟΠΙ μεγεθους $\alpha_2 \in (0,1)$ για την $H: \theta=\theta_0$ vs $K_2: \theta<\theta_0$.

Παρατηρουμε δε, οτι

(4.52) $c_2 < c_1$, εφοσον

(4.53) $\alpha_1+\alpha_2 = \alpha \in (0,1)$,

διοτι, αν ηταν $c_1 \le c_2$, τοτε, επειδη οτι $\beta_{d_2}(\theta_0) = \alpha_2 \Rightarrow P_{\theta_0}(T(\underline{X}) < c_2) \le \alpha_2 \Rightarrow$
$\Rightarrow P_{\theta_0}(T(\underline{X}) < c_1) \le \alpha_2 \Rightarrow P_{\theta_0}(T(\underline{X}) \ge c_1) \ge$
$\ge 1-\alpha_2 > \alpha-\alpha_2 = \alpha_1$, το οποιο ειναι ατοπο, δεδομενου οτι $\beta_{d_1}(\theta_0) = \alpha_1$.

Αρα, η στατιστικη συναρτηση,

(4.54) $d(\underline{X}) = d_1(\underline{X}) + d_2(\underline{X})$
$= 1(T>c_1) + 1(T<c_2) + \gamma_1 1(T=c_1) + \gamma_2 1(T=c_2)$
$= 1 - 1(c_2 \le T \le c_1) + \gamma_1 1(T=c_1) + \gamma_2 1(T=c_2)$,

ειναι ενας ελεγχος μεγεθους $\alpha = \alpha_1+\alpha_2 \in (0,1)$ για την (4.49).

Το οτι ο (4.54) δεν ειναι, εν γενει, ΟΠΙ μεγεθους α, ειναι προφανες δεδομενου οτι ο $d'(\underline{X}) = 1(T>c'_\alpha) + \gamma' 1(T=c'_\alpha)$, με $\beta_{d'}(\theta_0)=\alpha$ ειναι ισχυροτερος του d για $\theta>\theta_0$ (αν και ειναι "κακος" ελεγχος για $\theta<\theta_0$) εφοσον το $\alpha_2>0$.

Παντως, εφοσον το μοντελο μας $f(\cdot|\theta)$, $\theta\in\Theta$, εχει ΜΠΠ ως προς την στατιστικη T, δειξαμε οτι για την $H: \theta=\theta_0$ vs $K: \theta\ne\theta_0$, η στατιστικη

(4.55) $d(\underline{X}) = 1 - 1(\, c_2 \le T(\underline{X}) \le c_1) + \Sigma_{i=1}^{2} \gamma_i 1(T(\underline{X}) = c_i)$

με $P_{\theta_0}(\, T(\underline{X}) > c_1\,) + \gamma_1 P_{\theta_0}(\, T(\underline{X}) = c_1) = \alpha_1$, $0 \le \gamma_1 < 1$

και $P_{\theta_0}(\, T(\underline{X}) < c_2) + \gamma_2 P_{\theta_0}(\, T(\underline{X}) = c_2) = \alpha_2$, $0 \le \gamma_2 < 1$,

οπου $\alpha_1 + \alpha_2 = \alpha \in (0,1)$,

ειναι ενας ελεγχος μεγεθους α, αλλ' οχι εν γενει ΟΠΙ.

Οι ελεγχοι της μορφης (4.55) ειναι, εν γενει, καλοι ελεγχοι, και οπως θα δουμε οι ελεγχοι πηλικου πιθανοφανειων εχουν συνηθως αυτη την μορφη, με ελευθερια στην επιλογη των $\alpha_1, \alpha_2 \in (0, \alpha)$, π.χ., $\alpha_1 = \alpha_2 = \alpha/2$.
Αυτη η τελευταια επιλογη των α_1, α_2 εχει εννοια κυριως οταν η κατανομη της $T(\underline{X})$ ειναι, υπο την H, συμμετρικη γυρω απ' την θ_0 , δηλαδη, $(T-\theta_0)$, $-(T-\theta_0)$ εχουν την ιδια κατανομη. Τοτε, ειναι απλο να δη κανεις οτι τα c_1, c_2 ικανοποιουν τη σχεση $(c_1 + c_2)/2 = \theta_0$ & $\gamma_1 = \gamma_2$.
Μεταξυ των συμμετρικων ελεγχων ειναι, μερικες φορες, εφικτο να βρεθη ΟΠΙ συμμετρικος ελεγχος:

(4.56) $d(\underline{X}) = 1\,(\,|T(\underline{X}) - \theta_0| > c_\alpha) + \gamma_\alpha P(|T(\underline{X}) - \theta_0| = c_\alpha)$

με $E_{\theta_0} d(\underline{X}) = \alpha$.

(4.57) <u>Παραδειγμα:</u> Εστω $X_1, \dots, X_n$ α.i. $N(\theta, \sigma_0^2)$, $\theta \in \Theta \equiv \mathbb{R}$, $\sigma_0^2 > 0$ γνωστο. Μας ενδιαφερει ο ελεγχος της $H: \theta = \theta_0$ vs $K: \theta \ne \theta_0$.

$$\underline{X} \sim \left(\frac{1}{2\pi\sigma_0^2}\right)^{n/2} \exp\left\{-\frac{1}{2\sigma_0^2}\Sigma_{i=1}^n x_i^2\right\} \exp\left\{-\frac{n\theta^2}{2\sigma_0^2} + \frac{n\theta}{\sigma_0^2}\bar{x}_n\right\},$$

και αρα, για $\underline{\theta} < \bar{\theta}$,

$$\frac{f_{\underline{X}}(\underline{x}|\bar{\theta})}{f_{\underline{X}}(\underline{x}|\underline{\theta})} = \exp\left\{-\frac{n(\bar{\theta}^2 - \underline{\theta}^2)}{2\sigma_0^2}\right\} \exp\left\{\frac{n(\bar{\theta} - \underline{\theta})}{\sigma_0^2} T(\underline{x})\right\} \uparrow_{T(\underline{x})}$$

οπου, $T(\underline{X}) := \bar{X}_n \sim N\left(\theta, \frac{\sigma_0^2}{n}\right)$ <u>συμμετρικη</u> γυρω απ' το $\theta \in \Theta$.

Τα οσα αναπτυξαμε προηγουμενως μας οδηγουν στον ελεγχο:

$$d^*(\underline{X}) = 1\left(|\bar{X}_n - \theta_0| > c_\alpha \right),$$

με $E_{\theta_0} d(\underline{X}) = P_{\theta_0}\left(|\bar{X}_n - \theta_0| > c_\alpha \right) = \alpha$

$$\Leftrightarrow \quad P_{\theta_0}\left(|N(0,1)| \leq \frac{\sqrt{n}\, c_\alpha}{\sigma_0} \right) = 1-\alpha$$

$$\Leftrightarrow \quad \Phi\left(\frac{\sqrt{n}\, c_\alpha}{\sigma_0}\right) - \Phi\left(-\frac{\sqrt{n}\, c_\alpha}{\sigma_0}\right) = 1-\alpha$$

$$\Leftrightarrow \quad \Phi\left(\frac{\sqrt{n}\, c_\alpha}{\sigma_0}\right) = 1 - \alpha/2 \quad \Leftrightarrow \quad c_\alpha = \frac{\sigma_0}{\sqrt{n}} \Phi^{-1}(1-\alpha/2) = Z(1-\alpha/2)\frac{\sigma_0}{\sqrt{n}}.$$

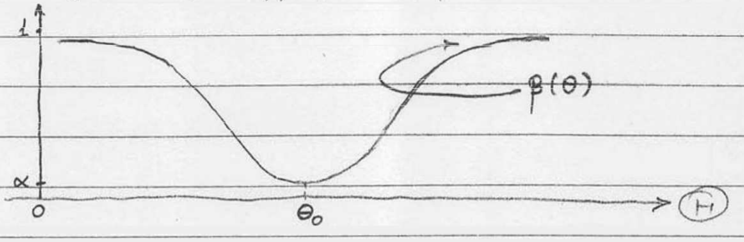
Ο ελεγχος αυτος εχει συναρτηση ισχυος:

$$\beta(\theta|d^*) = E_\theta d^*(\underline{X}) = P_\theta\left(|\bar{X}_n - \theta_0| > z(1-\alpha/2)\frac{\sigma_0}{\sqrt{n}} \right)$$

$$= \Phi\left(-z(1-\alpha/2) + \frac{\sqrt{n}(\theta-\theta_0)}{\sigma_0}\right) + \Phi\left(-z(1-\alpha/2) + \frac{\sqrt{n}(\theta_0-\theta)}{\sigma_0}\right)$$

$$= \Phi\left(-z(1-\alpha/2) + \frac{\sqrt{n}|\theta-\theta_0|}{\sigma_0}\right) + \Phi\left(-z(1-\alpha/2) - \frac{\sqrt{n}|\theta-\theta_0|}{\sigma_0}\right), \theta \in \Theta$$

ειναι, δηλαδη, συμμετρικη γυρω απ' το θ_0:



,

εξ ου και η ονομασια του ελεγχου ως συμμετρικου. (μεταχειριζεται εξ ισου την θ και την $2\theta_0 - \theta$).

Ο ελεγχος αυτος ειναι ΟΠΙ μεταξυ ολων των συμμετρικων ελεγχων που βασιζονται στην επαρκη (και πληρη), για την θ, στατιστικη $T(\underline{X}) = \bar{X}_n$ ή ισοδυναμα στην $\zeta := \frac{\sqrt{n}(\bar{X}_n - \theta_0)}{\sigma_0} \sim N(\Delta, 1)$ οπου $\Delta := \theta - \theta_0$, εφοσον $\theta \in \Theta$ ειναι η αληθης τιμη της παραμετρου, για τον ελεγχο της $H: \theta = \theta_0$ vs $K: \theta \neq \theta_0$. (Το να βασιζεται ενας ελεγχος σε επαρκη στατιστικη φαινεται – διαισθητικα τουλαχιστον – επιθυμητο.) Θα δειξουμε τωρα τον παραπανω ισχυρισμο:

Εφοσον ενδιαφερομαστε για συμμετρικους – γυρω απ' το θ_0 –

ελεγχους, αυτοι εχουν συναρτηση ισχυος $\beta(\theta|d) = g_d(\delta)$,
οπου $\delta \equiv |\Delta|$ (οπως: $\beta(\theta|d^*) = g_{d^*}(\delta) = \Phi\left(z(\alpha/2) + \frac{\sqrt{n}\,\delta}{\sigma_0}\right) + \Phi\left(z(\alpha/2) - \frac{\sqrt{n}\,\delta}{\sigma_0}\right)$).
Επισης, εφοσον δεχθηκαμε να βασισθουμε σε ελεγχους βασισμενους στην επαρκη στατιστικη $T \equiv \bar{X}_n$ ή ισοδυναμα στη ζ, μπορουμε να προσποιηθουμε οτι εχουμε μια παρατηρηση

$$\xi := |\zeta| \equiv \frac{\sqrt{n}\,|\bar{X}_n - \theta_0|}{\sigma_0} \sim \{\phi(z+\delta) + \phi(z-\delta)\}1(z>0) \quad (\forall \theta \in \Theta),$$

με βαση την οποια θελουμε να ελεγξουμε την

$$H: \delta = 0 \ \text{ vs } \ K: \delta > 0 \quad (\Leftrightarrow H: \theta = \theta_0 \ \text{ vs } \ K: \theta \neq \theta_0), \text{ σε επιπεδο } \alpha.$$

Αλλα, $\dfrac{f(\xi|\delta)}{f(\xi|0)} = \dfrac{\phi(\xi+\delta) + \phi(\xi-\delta)}{2\phi(\xi)} = \dfrac{e^{-\delta^2/2}}{2}\left[e^{-\delta\xi} + e^{\delta\xi}\right] \uparrow \xi$

$\left(\frac{\partial}{\partial\xi}\left(e^{-\delta\xi} + e^{\delta\xi}\right) = \delta\left(e^{\delta\xi} - e^{-\delta\xi}\right) > 0\right)$.

Αρα, ο ΟΠΙ της $\delta = 0$ vs $\delta > 0$, ειναι ο εξης:

$$d(\xi) = 1(\xi > c_\alpha), \text{ με } E_{\delta=0}\, d(\xi) = P_{\delta=0}(\xi > c_\alpha) =$$
$$= \int_{c_\alpha}^{\infty} 2\phi(z)\,dz = 2\left[1 - \Phi(c_\alpha)\right] = \alpha \Leftrightarrow c_\alpha = z(1-\alpha/2),$$

δηλαδη,

$$d(\xi) = 1\left(\frac{\sqrt{n}\,|\bar{X}_n - \theta_0|}{\sigma_0} > z(1-\alpha/2)\right) = d^*(\underline{X}).$$

Δειξαμε, λοιπον, οτι ο d^* ειναι ΟΠΙ, μεγεθους α, για την $H: \theta = \theta_0$ vs $K: \theta \neq \theta_0$, μεταξυ ολων των συμμετρικων ελεγχων που βασιζονται στην επαρκη στατιστικη $T \equiv \bar{X}_n$ για την $\theta \in \Theta$.

Ξανακοιτωντας την ισχυ του d^*, $\beta(\theta|d^*) = g_{d^*}(\delta)$, παρατηρουμε οτι $g'_{d^*}(\delta) = \frac{\sqrt{n}}{\sigma_0}\left[\phi\left(z(\alpha/2) + \frac{\sqrt{n}\,\delta}{\sigma_0}\right) - \phi\left(z(\alpha/2) - \frac{\sqrt{n}\,\delta}{\sigma_0}\right)\right]$
≥ 0, με "=" αν και μονο αν $\delta = 0$ ή $\theta = \theta_0$, ειναι, δηλαδη, η $\beta(\theta|d^*)$ αυξουσα συναρτηση της $|\theta - \theta_0|$,
και,

$$\beta(\theta|d^*) \geq \beta(\theta^0|d^*) = \alpha \quad \forall \theta \in \Theta,$$

ειναι, δηλαδη, ο d^* <u>αμεροληπτος</u> ελεγχος:

<u>Ορισμος:</u> Ενας ελεγχος d, της $H: \theta \in \Theta_0$ vs $K: \theta \in \Theta_1$, με $\Theta_0 \cap \Theta_1 = \phi$, $\Theta_0, \Theta_1 \subseteq \Theta$, καλειται <u>αμεροληπτος</u>, μεγεθους α, ανν: $\beta(\theta|d) \leq \alpha \quad \forall\, \theta \in \Theta_0$

και $\beta(\theta|d) \geq \alpha \quad \forall\, \theta \in \Theta_1$.

Μπορει να δειχθη οτι ο d^* του προηγουμενου παραδειγματος ειναι ΟΠΙ, μεγεθους α, μεταξυ ολων των αμεροληπτων ελεγχων. Επισης, γενικα, ενας ΟΠΙ ελεγχος, μεγεθους α, ειναι αμεροληπτος δεδομενου οτι ειναι ισχυροτερος του ελεγχου $d \equiv \alpha$.

Η αμεροληψια ενος ελεγχου ειναι μια επιθυμητη ιδιοτητα, διοτι εξασφαλιζει οτι η πιθανοτητα να απορριφθη η Η οταν πρεπει ($\theta \in \Theta_1$), δεν ειναι μικροτερη της πιθανοτητος να απορριφθη η Η οταν δεν πρεπει ($\theta \in \Theta_0$).

Για την <u>μονοπαραμετρικη εκθετικη οικογενεια κατανομων</u> (η οποια εχει ΜΠΠ ως προς την επαρκη στατιστικη T, για την θ), μπορει να αποδειχθη οτι υπαρχει ΟΠΙ αμεροληπτος ελεγχος, μεγεθους α, για την $H: \theta \in \Theta_0 \equiv [\theta_1, \theta_2]$ vs $K: \theta \in \Theta \setminus \Theta_0$, και ειναι ο εξης:

$$(4.58)\quad d(\underline{X}) = 1 - 1\left(c_1 < T(\underline{X}) < c_2\right) + \Sigma_{i=1}^{2} \gamma_i 1\left(T(\underline{X}) = c_i\right),$$

οπου οι σταθερες $c_1 < c_2$ και $\gamma_1, \gamma_2 \in [0,1)$, υπολογιζονται απο τις εξισωσεις: $E_{\theta_i} d(\underline{X}) = \alpha$, $i = 1, 2$.

Στην ειδικη περιπτωση που $\theta_1 = \theta_2 \equiv \theta_0$, οι εξισωσεις που διδουν τις $c_1 < c_2$ και $\gamma_1, \gamma_2 \in [0,1)$ γινονται:

$$E_{\theta_0} d(\underline{X}) = \alpha \quad \underline{\text{και}} \quad E_{\theta_0}\{T(\underline{X}) d(\underline{X})\} = \alpha E_{\theta_0}\{T(\underline{X})\}.$$

Αν επι πλεον η επαρκης στατιστικη T κατανεμεται, υπο την $H: \theta = \theta_0$, συμμετρικα γυρω απο καποιο σημειο m και περιοριζουμε σε συμμετρικους ελεγχους, βασισμενους στην T, $d(\underline{X}) = \psi(|T(\underline{X}) - m|)$, τοτε

$$E_{\theta_0}(T d(\underline{X})) = E_{\theta_0}\{(T-m)\,\psi(|T-m|)\} + m\, E_{\theta_0}\psi(|T-m|)$$
$$= m\alpha = \alpha E_{\theta_0} T,$$

δηλαδη, η δευτερη των σχεσεων που δίδουν τις σταθερές, επεται της πρωτης ($E_{\theta_0} d(\underline{X}) = \alpha$). Σ' αυτη την περιπτωση ο ΟΠΙ αμεροληπτος συμμετρικος ελεγχος, επιπεδου α, της $H:\theta=\theta_0$ vs $K:\theta\neq\theta_0$, ειναι ο εξης:

$$d(\underline{X}) = 1(|T(\underline{X})-m| > c_\alpha) + \gamma\, 1(|T(\underline{X})-m| = c_\alpha)$$

με $P_{\theta_0}(|T-m| > c_\alpha) + \gamma P(|T-m| = c_\alpha) = \alpha$.

Ο ελεγχος του Παραδ. (4.57) ειναι μια τετοια ακριβως περιπτωση με $m=\theta_0$. Δηλαδη, ο ΟΠΙ αμεροληπτος, βγαινει και συμμετρικος.

Αν, συνεχιζοντας το Παραδ. (4.57), μας ενδιαφερει ο ελεγχος της $H:\theta\in\Theta_0=[\theta_1,\theta_2]$ vs $K:\theta\in\Theta=\mathbb{R}\setminus\Theta_0$, τοτε απ' την (4.58) εχουμε οτι ο ΟΠΙ αμεροληπτος, επιπεδου α, ειναι ο εξης:

$$d(\underline{X}) = 1(\bar{X}_n < c_1) + 1(\bar{X}_n > c_2),$$

με $c_1 < c_2$, που υπολογιζονται απο τις ακολουθες εξισωσεις:

$$E_{\theta_i}\{d(\underline{X})\} = P_{\theta_i}(\bar{X}_n < c_1) + P_{\theta_i}(\bar{X}_n > c_2) = \alpha, \quad i=1,2,$$
$$\Leftrightarrow \quad \Phi\left(\frac{\sqrt{n}(c_2-\theta_i)}{\sigma_0}\right) - \Phi\left(\frac{\sqrt{n}(c_1-\theta_i)}{\sigma_0}\right) = 1-\alpha, \quad i=1,2,$$

οι οποιες μπορουν να λυθουν αριθμητικα.

(4.59) Παραδ. Εστω $X_1,\dots,X_n$ α.ι. Εκθ.(λ), $\lambda\in\Theta\equiv(0,+\infty)$. Μας ενδιαφερει ο ελεγχος της $H:\lambda=\lambda_0$ vs $K:\lambda\neq\lambda_0$.

Εφοσον η κατανομη αυτη εχει Μ.Π.Π. ως προς $T\equiv\sum_{i=1}^n X_i$, ειδαμε οτι ο ελεγχος : $d_0(\underline{X}) = 1(T>c_2) + 1(T<c_1)$, οπου : $P_{\lambda_0}(T>c_2) = P_{\lambda_0}(T<c_1) = \alpha/2$, δηλαδη, $c_1 = \chi^2_{2n}(\alpha/2)/(2\lambda_0)$ και $c_2 = \chi^2_{2n}(1-\alpha/2)/(2\lambda_0)$, ειναι ενας καλος (λογικος) ελεγχος, επιπεδου α, αλλα οχι ΟΠΙ. Μηπως ειναι ΟΠΙ αμεροληπτος; Η απαντηση ειναι ΟΧΙ,

διοτι, εφοσον η εκθετικη ανηκει στην μονοπαραμετρικη εκθετικη οικογενεια κατανομων, απο την (4.58), εχουμε οτι υπαρχει ΟΠΙ αμεροληπτος ελεγχος $d(\underline{X}) = 1(T<c_1) + 1(T>c_2)$,

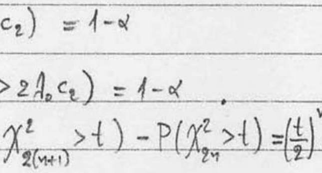
οπου: $$\left.\begin{aligned} & P_{\lambda_0}(T<c_1) + P_{\lambda_0}(T>c_2) = \alpha \\ & \int_0^{c_1} t\,\frac{\lambda_0^n}{(n-1)!}\,t^{n-1} e^{-\lambda_0 t}\,dt + \int_{c_2}^{\infty} t\,\frac{\lambda_0^n}{(n-1)!}\,t^{n-1} e^{-\lambda_0 t}\,dt = \alpha\,\frac{n}{\lambda_0} \end{aligned}\right\} \Leftrightarrow$$

$$\Leftrightarrow \left\{\begin{aligned} & P(\chi^2_{2n} < 2\lambda_0 c_1) + P(\chi^2_{2n} > 2\lambda_0 c_2) = \alpha \\ & P(\chi^2_{2(n+1)} < 2\lambda_0 c_1) + P(\chi^2_{2(n+1)} > 2\lambda_0 c_2) = \alpha \end{aligned}\right.$$

$$\Leftrightarrow \left\{\begin{aligned} & P(\chi^2_{2n} > 2\lambda_0 c_1) - P(\chi^2_{2n} > 2\lambda_0 c_2) = 1-\alpha \\ & P(\chi^2_{2(n+1)} > 2\lambda_0 c_1) - P(\chi^2_{2(n+1)} > 2\lambda_0 c_2) = 1-\alpha \,. \end{aligned}\right.$$

Ειναι απλο να ελεγξει κανεις οτι: $P(\chi^2_{2(n+1)} > t) - P(\chi^2_{2n} > t) = \left(\frac{t}{2}\right)^n e^{-t/2}$ και το συστημα μας γινεται:

$$\left\{\begin{aligned} & P(\chi^2_{2n} > 2\lambda_0 c_1) - P(\chi^2_{2n} > 2\lambda_0 c_2) = 1-\alpha \\ & g(\lambda_0 c_1/n) = g(\lambda_0 c_2/n)\,, \end{aligned}\right.$$



οπου, $g(x) := x - \log x$,

και $P(\chi^2_{2n} > t) = \sum_{k=0}^{n-1} \frac{(t/2)^k}{k!} e^{-t/2}$.

Η λυση πρεπει παλι να γινει με αριθμητικες μεθοδους.

Παντως, για "μεγαλο" n η κατανομη της T, υπο την $H: \lambda = \lambda_0$, ειναι "σχεδον" συμμετρικη γυρω απο την $m = \frac{n}{\lambda_0}$, οποτε ο ελεγχος (4.55) με $(\gamma_1 = \gamma_2 = 0)$ $\alpha_1 = \alpha_2 = \alpha/2$, δεν θα "απεχει" πολυ απο τον ΟΠΙ αμεροληπτο (συμμετρικο) ελεγχο που μολις κατασκευασαμε.

(4.60) <u>Ασκηση</u>: Κατασκευασατε αμφιπλευρους ελεγχους της μορφης (4.55) με $\alpha_1 = \alpha_2 = \alpha/2$, για την Poisson (θ), $\theta \in \Theta \equiv (0, +\infty)$, καθως και για τις κατανομες των ασκησεων (4.44), (4.45), (4.46) & (4.48).

Εστω $\underline{X} \equiv (X_1, \ldots, X_n)$ α.τ. $f(\cdot|\theta)$, , και εστω $d_{\theta_0}(\underline{X})$ ενας (μη τυχαιοποιημενος) ελεγχος, μεγεθους α, της $H: \theta = \theta_0$ vs $K: \theta \neq \theta_0$, για οποιο $\theta_0 \in \Theta$.

Το συνολο:

(4.61) $$I(\underline{X}) := \{ \theta_0 \in \Theta : d_{\theta_0}(\underline{X}) = 0 \},$$

ειναι μια περιοχη εμπιστοσυνης για την παραμετρο θ, με συντελεστη εμπιστοσυνης $1-\alpha$, διοτι:

$$P_\vartheta(I(\underline{X}) \ni \vartheta) = P_\vartheta(d_\vartheta(\underline{X}) = 0) = 1 - P_\vartheta(d_\vartheta(\underline{X}) = 1) = 1 - \alpha \quad \forall \vartheta \in \Theta.$$

Για παραδειγμα, εστω ο, μεγεθους α, ελεγχος

$$d^*(\underline{X}) = 1\left(|\bar{X}_n - \theta_0| > z(1-\tfrac{\alpha}{2}) \tfrac{\sigma_0}{\sqrt{n}} \right) =: d^*_{\theta_0}(\underline{X}),$$

της $H: \theta = \theta_0$ vs $K: \theta \neq \theta_0$, για οποιο $\theta_0 \in \Theta \equiv \mathbb{R}$, του παραδ. 4.57.

Η αντιστοιχη, του ελεγχου αυτου, περιοχη εμπιστοσυνης ειναι η:

$$\begin{aligned} I^*(\underline{X}) &= \{ \vartheta \in \Theta \equiv \mathbb{R} : d^*_\vartheta(\underline{X}) = 0 \} = \\ &= \{ \vartheta \in \Theta : |\bar{X}_n - \vartheta| \leq z(1-\alpha/2) \tfrac{\sigma_0}{\sqrt{n}} \} \\ &= \{ \vartheta \in \mathbb{R} : \bar{X}_n - z(1-\alpha/2)\tfrac{\sigma_0}{\sqrt{n}} \leq \vartheta \leq \bar{X}_n + z(1-\alpha/2)\tfrac{\sigma_0}{\sqrt{n}} \} \\ &= \left[\bar{X}_n - z(1-\alpha/2)\tfrac{\sigma_0}{\sqrt{n}}, \ \bar{X}_n + z(1-\alpha/2)\tfrac{\sigma_0}{\sqrt{n}} \right]. \end{aligned}$$

Οπως και στη γενικη περιπτωση, βλεπουμε οτι:

$$\begin{aligned} P_\vartheta(I^*(\underline{X}) \ni \vartheta) &= P_\vartheta\left(\bar{X}_n - z(1-\tfrac{\alpha}{2})\tfrac{\sigma_0}{\sqrt{n}} \leq \vartheta \leq \bar{X}_n + z(1-\alpha/2)\tfrac{\sigma_0}{\sqrt{n}} \right) = \\ &= P_\vartheta\left(|\bar{X}_n - \vartheta| \leq z(1-\alpha/2)\tfrac{\sigma_0}{\sqrt{n}} \right) = 1 - P_\vartheta\left(|\bar{X}_n - \vartheta| > z(1-\alpha/2)\tfrac{\sigma_0}{\sqrt{n}} \right) = \\ &= 1 - \alpha \quad \forall \vartheta \in \Theta. \end{aligned}$$

Αν ενας ελεγχος d ειναι ομοιομορφα ισχυροτερος ενος ελεγχου d', δηλ., $\beta(\theta|d_\vartheta) \geq \beta(\theta|d'_\vartheta)$ $\forall \theta \in \Theta \setminus \{\vartheta\}$, $\vartheta \in \Theta$, τοτε η αντιστοιχη περιοχη εμπιστοσυνης $I(\underline{X})$ ειναι ακριβεστερη της $I'(\underline{X})$:

(4.62) $$\begin{aligned} P_\theta(I(\underline{X}) \ni \vartheta) &= P_\theta(d_\vartheta(\underline{X}) = 0) = 1 - P_\theta(d_\vartheta(\underline{X}) = 1) = \\ &= 1 - \beta(\theta|d_\vartheta) \leq 1 - \beta(\theta|d'_\vartheta) = P_\theta(I'(\underline{X}) \ni \vartheta) \end{aligned}$$

$\forall \vartheta, \theta \in \Theta$ με $\vartheta \neq \theta$.

Επισης, ενας αμεροληπτος ελεγχος d οδηγει σε αμεροληπτη περιοχη εμπιστοσυνης I : $\forall \theta, \vartheta \in \Theta$

(4.63) $P_\theta(I(\underline{X}) \ni \vartheta) = 1 - \beta(\theta | d_\vartheta) \leq 1 - \beta(\theta | d_\theta) = P_\theta(I(\underline{X}) \ni \theta)$.

(4.64) <u>Ασκηση</u>: Κατασκευαστε τα επιπεδου $1-\alpha$, διαστηματα εμπιστοσυνης που αντιστοιχουν στους ελεγχους της Ασκ. 4.60.

(4.65) <u>Παραδ.</u>: Θα κατασκευασουμε το επιπεδου $1-\alpha$, διαστημα εμπιστοσυνης για την παραμετρο $\theta \in \Theta = (0,+\infty)$ της $\mathcal{E}(\lambda \equiv \theta)$, που αντιστοιχει στον ελεγχο d_ϑ του Παραδ. (4.59):

Ειδαμε, οτι $\forall \vartheta \in \Theta$, ο ελεγχος:

$$d_\vartheta(\underline{X}) := 1(T > \chi^2_{2n}(1-\alpha/2)/(2\vartheta)) + 1(T < \chi^2_{2n}(\alpha/2)/(2\vartheta)),$$

οπου $T := \Sigma_{i=1}^n X_i$, ειναι ενας, μεγεθους α ($E_\vartheta d_\vartheta(\underline{X}) = \alpha$), ελεγχος της $H: \theta = \vartheta$ vs $K: \theta \neq \vartheta$.

Το αντιστοιχο, λοιπον, διαστημα εμπιστοσυνης, επιπεδου $1-\alpha$, για την θ ειναι το εξης:

$$I_d(\underline{X}) := \{\vartheta \in \Theta : d_\vartheta(\underline{X}) = 0\} =$$
$$= \{\vartheta \in \Theta : \chi^2_{2n}(\alpha/2) \leq 2\vartheta T \leq \chi^2_{2n}(1-\alpha/2)\}$$
$$= \left[\frac{\chi^2_{2n}(\alpha/2)}{2n\,\overline{X}_n} , \frac{\chi^2_{2n}(1-\alpha/2)}{2n\,\overline{X}_n} \right].$$

4.2. ΕΛΕΓΧΟΙ ΠΗΛΙΚΟΥ ΠΙΘΑΝΟΦΑΝΕΙΩΝ.

Το κυριο προβλημα στην κατασκευη των ελεγχων που βασιζονται στο Λημμα των Neyman και Pearson ειναι το οτι ακομη και για τον ελεγχο απλης εναντι απλης, η H και η K πρεπει να αφορουν ολες τις αγνωστες παραμετρους του μοντελου. Για παραδειγμα αν το δειγμα μας ειναι απο την $N(\mu,\sigma^2)$, $\underline{\theta}=(\mu,\sigma^2) \in \Theta = \mathbb{R}\times(0,+\infty)$, με μ,σ^2 αγνωστα, μπορουμε μεν να ελεγξουμε την $H: \underline{\theta}=\underline{\theta}_0$ vs $K: \underline{\theta}=\underline{\theta}_1$, για συγκεκριμενες τιμες $\underline{\theta}_0, \underline{\theta}_1$, μεσω του Λημματος των Neyman & Pearson, αν ομως μας ενδιαφερει μονο η παραμετρος μ, και αντιμετωπιζουμε την σ^2 σαν την ενοχλητικη παραμετρο του προβληματος, τοτε δεν μπορουμε να χρησιμοποιησουμε το Λημμα των Neyman και Pearson για τον ελεγχο της $H: \mu=\mu_0$ vs $K: \mu=\mu_1$, εφοσον η σ^2 ειναι αγνωστη.

Επισης, αν $\dim\Theta > 1$, η ιδιοτητα του ΜΠΠ δεν εχει ικανοποιητικη γενικευση στις περισσοτερες της μιας διαστασεις και ετσι δεν μπορουμε να μιλαμε για ΟΠΙ ελεγχους, οι οποιοι, ουτως ή αλλως, συχνα δεν υπαρχουν ακομη και αν $\dim\Theta = 1$ (υπαρχουν μονο υπο ορισμενες συνθηκες, οπως το ΜΠΠ).

Οι ελεγχοι που θα κατασκευασουμε σ' αυτη την ενοτητα δεν εχουν πλουσια θεωρια για μικρα μεγεθη n δειγματος, μπορει ομως να δειχθη οτι ειναι ασυμπτωτικα (δηλαδη, για $n\to\infty$) βελτιστοι, βασει των κριτηριων που εχουμε ηδη δοσει. Επισης, ακομη και για μικρα n, οι ελεγχοι αυτοι ειναι εξαιρετικα καλοι στην πραξη — σχεδον βελτιστοι, ταυτιζονται δε, εν γενει, με τους ΟΠΙ ελεγχους της προηγουμενης ενοτητος, εφοσον αυτοι οι τελευταιοι υπαρχουν. Εξ αλλου, οι ελεγχοι πηλικου πιθανοφανειων (ΕΠΠ), εν γενει υπαρχουν!

Εστω, λοιπον, $X_1,\dots,X_n$ α.τ. $f(\cdot|\theta)$, $\theta\in\Theta\subseteq\mathbb{R}^k$.
Μας ενδιαφερει ο ελεγχος της

(4.65) $H:\theta\in\Theta_0$ vs $K:\theta\in\Theta_1$,

$\Theta_0,\Theta_1\subseteq\Theta$, με $\Theta_0\cap\Theta_1=\phi$.

Οι ελεγχοι ΕΠΠ, που θα κατασκευασουμε εδω, βασιζονται στη σύγκριση των πιθανοφανειων

$$L(\Theta_0|\underline{X}) := \sup\{L(\theta|\underline{X}),\ \theta\in\Theta_0\} =: L(\hat{\theta}_0|\underline{X}),$$

και,

$$L(\Theta_1|\underline{X}) := \sup\{L(\theta|\underline{X}),\ \theta\in\Theta_1\} =: L(\hat{\theta}_1|\underline{X}),$$

οπου,

$L(\theta|\underline{X}) := \prod_{i=1}^{n} f(X_i|\theta)$, ειναι η συναρτηση πιθανοφανειας, και η $\hat{\theta}_i$ ειναι η εκτιμητρια μεγιστης πιθανοφανειας της θ, υπο τον περιορισμο οτι $\theta\in\Theta_i$, $i=0,1$.

Για την κατασκευη λοιπον ενος ΕΠΠ, πρεπει πρωτα να λυσουμε τα ακολουθα δυο προβληματα εκτιμησης:

$$(4.66)\ \begin{cases}\max L(\theta|X)\\ \text{οταν } \theta\in\Theta_i, & i=0,1,\end{cases}$$

και να υπολογισουμε, μετα, το πηλικο των μεγιστων πιθανοφανειων:

$$(4.67)\ LR(\underline{X}) := \frac{L(\Theta_1|\underline{X})}{L(\Theta_0|\underline{X})} = \frac{L(\hat{\theta}_1|\underline{X})}{L(\hat{\theta}_0|\underline{X})}.$$

Τοτε, ο ΕΠΠ της (4.65), οριζεται ως εξης:

$$(4.68)\quad d(\underline{X}) = \mathbb{1}(LR(\underline{X})>c_\alpha) + \gamma\,\mathbb{1}(LR(\underline{X})=c_\alpha)$$

οπου, οι σταθερες c_α, $\gamma\in[0,1)$, ικανοποιουν την:

$$(4.69)\quad \sup\{P_\theta(LR(\underline{X})>c_\alpha)+\gamma P_\theta(LR(\underline{X})=c_\alpha),\ \theta\in\Theta_0\}=\alpha,$$

οπως και στην προηγουμενη ενοτητα.

Εν γενει, τα προβληματα (4.66) δεν ειναι απλο να λυθουν — χρειαζονται μεθοδους βελτιστοποιησης υπο περιορισμους, π.χ., πολλαπλασιαστες Lagrange. Θα δουμε ομως, αμμεσως, οτι — εφοσον $\exists \hat{\theta} \in \mathring{\Theta}$ — για τον ελεγχο της

(4.70) $H: \theta \in \Theta_0$ vs $K: \theta \in \Theta_1$,

με $\Theta_1 = \Theta \setminus \Theta_0$, $\Theta_0 \subseteq \Theta \subseteq \mathbb{R}^k$, και $\dim \Theta_0 < \dim \Theta$, αν η $f(\cdot|\theta)$, $\theta \in \Theta$ ειναι συνεχης συναρτηση της θ, τοτε μπορουμε να αποφυγουμε τη λυση τουλαχιστον του ενος των προβληματων (4.66) — αυτου με $i=1$.

Οριζουμε, λοιπον, τη στατιστικη:

(4.71) $$\lambda(\underline{X}) := \frac{L(\Theta|\underline{X})}{L(\Theta_0|\underline{X})} = \frac{L(\hat{\theta}|\underline{X})}{L(\hat{\theta}_0|\underline{X})},$$

οπου, $L(\Theta|\underline{X}) := \sup\{L(\theta|\underline{X}), \theta \in \Theta\}$, ειναι, δηλαδη, η $\hat{\theta}$ η — χωρις περιορισμους — εκτιμητρια μεγιστης πιθανοφανειας (εμπ) της θ, η οποια ειναι συχνα ευκολο να υπολογισθη, οπως ειδαμε στο προηγουμενο κεφαλαιο.

(4.72) Λημμα. Υπο τις προϋποθεσεις (4.70),
$$\lambda(\underline{X}) = LR(\underline{X}).$$

Αποδ. Παρατηρουμε οτι $\lambda(\underline{X}) \neq LR(\underline{X})$ ειναι δυνατον αν και μονο αν το σημειο ολικου μεγιστου $\hat{\theta} \in \Theta_0$, ειναι, δηλαδη, $\hat{\theta} = \hat{\theta}_0$, στην οποια περιπτωση $\lambda(\underline{X}) = 1 \geq LR(\underline{X})$.

Θα δειξουμε δε τωρα οτι, εφοσον $\dim \Theta_0 < \dim \Theta$ και η $L(\theta|\underline{X})$ ειναι συνεχης συναρτηση της $\theta \in \Theta$, εχουμε $L(\Theta_1|\underline{X}) = L(\hat{\theta}_0|\underline{X})$ και αρα και σ' αυτη την περιπτωση $\lambda(\underline{X}) = LR(\underline{X})$:

Αρκει να δειξουμε οτι $\forall \delta > 0$ $S(\hat{\theta}_0, \delta) \cap \Theta_1 \neq \emptyset$, οπου, $S(\hat{\theta}_0, \delta) := \{z \in \Theta : \|z - \hat{\theta}_0\| < \delta\}$, $\|z\|^2 := z^T z$.

Τοτε, αϖο τη συνεχεια της $L(\theta|\underline{x})$, $\theta\in\Theta$, θα εχουμε οτι $\forall\varepsilon>0\ \exists\delta>0$: $|L(\theta|\underline{x}) - L(\hat{\theta}_0|\underline{x})|<\varepsilon \quad \forall\,\theta\in S(\hat{\theta}_0,\delta)$ και αρα και $\forall\,\theta\in\Theta_1\cap S(\hat{\theta}_0,\delta)$, δηλαδη, $L(\Theta_1|\underline{x}) = L(\hat{\theta}_0|\underline{x})$.
Τωρα, για να δειξουμε οτι $\forall\delta>0 \quad S(\hat{\theta}_0,\delta)\cap\Theta_1\neq\emptyset$ αρκει να δειξουμε οτι υϖαρχει διανυσμα $y\in S(\hat{\theta}_0,\delta)$ του οϖοιου η διασταση ειναι ιση με $k=\dim\Theta$. Ενα τετοιο διανυσμα ειναι, π.χ., το $\hat{\theta}_0+\delta'\frac{1}{\sqrt{k}}\sum_{i=1}^k e_i$, για καϖοιο $\delta'<\delta$, και $e_1,\dots,e_k$ ειναι τα μοναδιαια διανυσματα του $\mathbb{R}^k$.

Αρα, υϖο τις συνθηκες (4.70) ο ΕΠΠ (4.68) γινεται:

(4.73) $d(\underline{x}) = 1(\lambda(\underline{x})>c_\alpha) + \gamma\, 1(\lambda(\underline{x})=c_\alpha)$

οϖου, οι σταθερες c_α, $\gamma\in[0,1)$, ικανοϖοιουν την:

(4.74) $\sup\{P_\theta(\lambda(\underline{x})>c_\alpha)+\gamma P_\theta(\lambda(\underline{x})=c_\alpha),\ \theta\in\Theta_0\}=\alpha$.

(4.75) Παραδειγμα: Εστω $X_1,\dots,X_n$ α.ι. $N(\mu,\sigma^2)$, $\theta=(\mu,\sigma^2)\in\Theta=\mathbb{R}\times(0,+\infty)$. Δειξτε οτι ο ΕΠΠ μεγεθους α

(a) της $H:\mu=\mu_0$ vs $K:\mu\neq\mu_0$ ειναι ο εξης:

$$d(\underline{x}) = 1(|T_n|>t_{n-1}(1-\alpha/2)),$$

οϖου, $T_n := \dfrac{\sqrt{n}(\bar{X}_n-\mu_0)}{S_n}\sim t_{n-1}$, $S_n^2=\frac{1}{n-1}\sum_{i=1}^n(X_i-\bar{X}_n)^2$.

(β) της $H:\sigma^2=\sigma_0^2$ vs $K:\sigma^2\neq\sigma_0^2$ ειναι ο εξης:

$$d(\underline{x}) = 1\Big(S_n^2<\frac{\sigma_0^2}{n-1}\chi^2_{n-1}(\alpha/2)\Big)+1\Big(S_n^2>\frac{\sigma_0^2}{n-1}\chi^2_{n-1}(1-\alpha/2)\Big)$$

Σημ. $\chi^2_{n-1}(\alpha/2) = (n-1)/f_{\infty,n-1}(1-\alpha/2)$, βλ. (3.184).

Αϖαντ. (a) $\Theta_0=\{\mu_0\}\times(0,+\infty)$ και $\dim\Theta_0=1<\dim\Theta=2$, $\Theta_1=\Theta\setminus\Theta_0$. Παρατηρουμε οτι, εφοσον το σ^2 ειναι αγνωστο, δεν μϖορει να κατασκευασθη ελεγχος του τυϖου Neyman-Pearson. Θα κατασκευασουμε ομως τον ΕΠΠ ο οϖοιος μϖορει να βασισθη στο $\lambda(\underline{x})$ διοτι $\dim\Theta_0<$

$< \dim \Theta$ και η $f(\cdot|\theta) = \frac{1}{\sigma}\phi\left(\frac{\cdot - \mu}{\sigma}\right)$ ειναι συνεχης ως προς $\theta = (\mu, \sigma^2) \in \Theta$.

(i) Για τον υπολογισμο του $L(\Theta_0|\underline{X}) = L(\hat{\theta}_0|\underline{X})$, παρατηρουμε οτι $\hat{\theta}_0 = (\hat{\mu}_0, \hat{\sigma}_0^2) = (\mu_0, \frac{1}{n}\Sigma_{i=1}^{n}(X_i - \mu_0)^2)$, και αρα,

$$L(\Theta_0|\underline{X}) = (2\pi\hat{\sigma}_0^2)^{-n/2} \exp\left\{-\frac{1}{2\hat{\sigma}_0^2}\Sigma_{i=1}^{n}(X_i - \hat{\mu}_0)^2\right\}$$

$$= (2\pi\hat{\sigma}_0^2)^{-n/2} \exp\left\{-\frac{n}{2}\right\}$$

(ii) Για τον υπολογισμο του $L(\Theta|\underline{X}) = L(\hat{\theta}|\underline{X})$, παρατηρουμε οτι $\hat{\theta} = (\hat{\mu}, \hat{\sigma}^2) = (\bar{X}_n, \frac{1}{n}\Sigma_{i=1}^{n}(X_i - \bar{X}_n)^2)$

$$= (\bar{X}_n, \hat{\sigma}_n^2) = (\bar{X}_n, \frac{n-1}{n}S_n^2),$$

και αρα,

$$L(\Theta|\underline{X}) = (2\pi\hat{\sigma}_n^2)^{-n/2} \exp\left\{-\frac{1}{2\hat{\sigma}_n^2}\Sigma_{i=1}^{n}(X_i - \bar{X}_n)^2\right\}$$

$$= (2\pi\hat{\sigma}_n^2)^{-n/2} \exp\left\{-\frac{n}{2}\right\}.$$

(iii) Τοτε, $\lambda(\underline{X}) = \frac{L(\Theta|\underline{X})}{L(\Theta_0|\underline{X})} = \left(\frac{\hat{\sigma}_0^2}{\hat{\sigma}_n^2}\right)^{n/2} =$

$$= \left(\frac{\Sigma_{i=1}^{n}(X_i - \mu_0)^2}{\Sigma_{i=1}^{n}(X_i - \bar{X}_n)}\right)^{n/2} = \left\{\frac{\Sigma_{i=1}^{n}(X_i - \bar{X}_n)^2 + n(\bar{X}_n - \mu_0)^2}{\Sigma_{i=1}^{n}(X_i - \bar{X}_n)^2}\right\}^{n/2} =$$

$= \left\{1 + \frac{1}{n-1}T_n^2\right\}^{n/2}$, μια αυξουσα συναρτηση της $|T_n|$.

Αρα, $\lambda(\underline{X}) > c \iff |T_n| > c$,
και, $P_{\mu_0}(|T_n| > c) = \alpha \iff P_{\mu_0}(|T_n| \leq c) = 1 - \alpha$

$$\iff F_{t_{n-1}}(c) - F_{t_{n-1}}(-c) = 1 - \alpha \iff 2F_{t_{n-1}}(c) - 1 = 1 - \alpha$$

$$\iff F_{t_{n-1}}(c) = 1 - \alpha/2 \iff c = c_\alpha = t_{n-1}(1 - \alpha/2).$$

Σχετικα, βλ. (1.59) και (1.62), καθως και (1.40).

Σημειωνουμε οτι $S_n^2 \xrightarrow[n\to\infty]{P} \sigma^2$ – βλ. (1.97),

και $\sqrt{n}\,(\bar{X}_n - \mu_0) \xrightarrow[n\to\infty]{d} N(0,\sigma^2)$ υπο την H, απο το ΚΟΘ.

Αρα, $T_n \xrightarrow[n\to\infty]{d} N(0,1)$, απο το Θεωρημα του Slutsky, και αρα ο ελεγχος-t, που μολις κατασκευασαμε, για μεγαλα μεγεθη η δειγματος, ειναι πολυ κοντα στον ΟΠΙ ελεγχο (4.57) – οπου το σ^2 εθεωρηθη γνωστο.

Σημειωνουμε ομως επισης, οτι αν $n = 20$, $\alpha = 0.05$, $z(1-\alpha/2) = 1.96$ και $t_{n-1}(1-\alpha/2) = 2.4334$, και αρα ακομη και αν η εκτιμητρια S_n^2 της σ^2 ηταν πολυ κοντα στη σ^2 οι δυο ελεγχοι διαφερουν σημαντικα.

(β) $\Theta_0 = \mathbb{R}\times\{\sigma_0^2\}$, $\dim \Theta_0 = 1 < \dim \Theta = 2$,

$L(\theta|\underline{X})$, $\theta\in\Theta$ ειναι συνεχης, και αρα ο ΕΠΠ μπορει να βασισθη στο $\lambda(\underline{X})$.

(i) Για το υπολογισμο του $L(\Theta_0|\underline{X}) = L(\hat{\theta}_0|\underline{X})$, εχουμε οτι $\hat{\theta}_0 = (\hat{\mu}_0, \hat{\sigma}_0^2) = (\bar{X}_n, \sigma_0^2)$ και αρα,

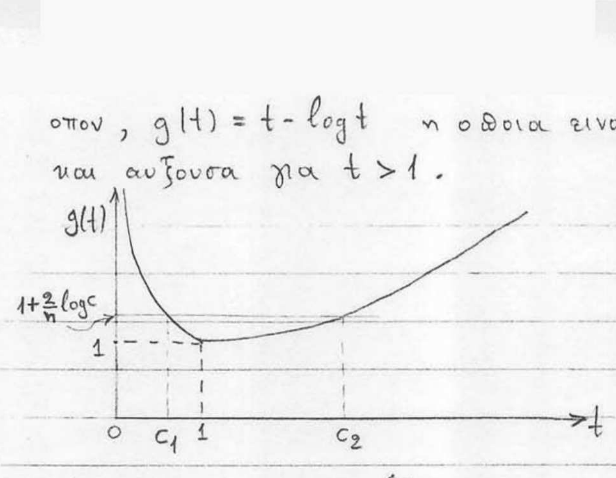
$$L(\hat{\theta}_0|\underline{X}) = (2\pi\sigma_0^2)^{-n/2}\exp\Big\{-\frac{1}{2\sigma_0^2}\Sigma_{i=1}^n (X_i-\bar{X}_n)^2\Big\}$$
$$= (2\pi\sigma_0^2)^{-n/2}\exp\Big\{-\frac{n\hat{\sigma}_n^2}{2\sigma_0^2}\Big\}.$$

(ii) Για τον υπολογισμο του $L(\Theta|\underline{X}) = L(\hat{\theta}|\underline{X})$, εχουμε οτι $\hat{\theta} = (\hat{\mu}, \hat{\sigma}^2) = (\bar{X}_n, \hat{\sigma}_n^2) = (\bar{X}_n, \frac{n-1}{n}S_n^2)$, και αρα,

$$L(\hat{\theta}|\underline{X}) = (2\pi\hat{\sigma}_n^2)^{-n/2}\exp\Big\{-\frac{1}{2\hat{\sigma}_n^2}\Sigma_{i=1}^n (X_i-\bar{X}_n)^2\Big\}$$
$$= (2\pi\hat{\sigma}_n^2)^{-n/2}\exp\Big\{-\frac{n}{2}\Big\},$$

(iii) Τοτε, $$\lambda(\underline{X}) = \frac{L(\Theta|\underline{X})}{L(\Theta_0|\underline{X})} = \left[\Big(\frac{\hat{\sigma}_n^2}{\sigma_0^2}\Big)\exp\Big\{1-\frac{\hat{\sigma}_n^2}{\sigma_0^2}\Big\}\right]^{-n/2}$$
$$= e^{-n/2}\left[\frac{\hat{\sigma}_n^2}{\sigma_0^2}\exp\Big\{-\frac{\hat{\sigma}_n^2}{\sigma_0^2}\Big\}\right]^{-n/2}$$
$$= e^{-n/2}\exp\Big\{\frac{n}{2}\,g\Big(\frac{\hat{\sigma}_n^2}{\sigma_0^2}\Big)\Big\}$$

οπου, $g(t) = t - \log t$ η οποια ειναι φθινουσα για $t \in (0,1)$ και αυξουσα για $t > 1$.



Αρα, $\lambda > c \iff g\left(\frac{\hat{\sigma}_n^2}{\sigma_0^2}\right) > c^* = 1 + \frac{2}{n}\log c$

$\iff \frac{\hat{\sigma}_n^2}{\sigma_0^2} < c_1$ ή $\frac{\hat{\sigma}_n^2}{\sigma_0^2} > c_2$

$\iff \Sigma_{i=1}^n (x_i - \bar{x}_n)^2 < c_1$ ή $\Sigma_{i=1}^n (x_i - \bar{x}_n)^2 > c_2$,

δηλαδη,

$$d(\underline{X}) = 1(S_n^2 < c_1) + 1(S_n^2 > c_2).$$

Οι κρισιμες σταθερες c_1, c_2 του ελεγχου, υπολογιζονται ουτως ωστε το μεγεθος του ελεγχου να ειναι α, δηλαδη,

$$P_{\sigma_0^2}(S_n^2 < c_1) + P_{\sigma_0^2}(S_n^2 > c_2) = \alpha$$

ή, δεδομενου οτι $(n-1)S_n^2/\sigma_0^2 \sim \chi^2_{n-1}$ υπο την H,

$$P(\chi^2_{n-1} < (n-1)c_1/\sigma_0^2) + P(\chi^2_{n-1} > (n-1)c_2/\sigma_0^2) = \alpha,$$

δηλαδη,

$$\frac{(n-1)c_1}{\sigma_0^2} = \chi^2_{n-1}(\alpha_1) \text{ και } \frac{(n-1)c_2}{\sigma_0^2} = \chi^2_{n-1}(1-\alpha_2), \ \alpha_1 + \alpha_2 = \alpha.$$

Για λογους συμμετριας του ελεγχου, συνηθως θετουμε $\alpha_1 = \alpha_2 = \alpha/2$ και καταληγουμε στον ελεγχο του μερους (β) της εκφωνησεως.

(4.76) Ασκηση. Δειξτε οτι οι ΕΜΠ μεγεθους α, ταυτιζονται με τους ελεγχους, μεγεθους α, της μορφης (4.45), στις ακολουθες περιπτωσεις ελεγχων της $H: \theta = \theta_0$ vs $K: \theta \neq \theta_0$:

(α) $\underline{X}$ a.i. $N(\theta, 1)$, $\theta \in \mathbb{R}$, (β) $\underline{X}$ α.ι. $N(0,\theta)$, $\theta>0$,

(γ) $\underline{X}$ α.ι. $\mathcal{E}(\theta)$, $\theta>0$,

(δ) $\underline{X}$ α.ι. Bernoulli(θ) , $\theta\in[0,1]$

(ε) $\underline{X}$ α.ι. $\mathcal{U}(0,\theta)$, $\theta>0$.

(4.77) <u>Παραδειγμα</u>. Εστω $X_1,\dots,X_{n_1}$ α.ι. $N(\mu_1,\sigma_1^2)$, και $Y_1,\dots,Y_{n_2}$ α.ι. $N(\mu_2,\sigma_2^2)$, και εστω οτι οι $\underline{X},\underline{Y}$ ειναι και μεταξυ τους ανεξαρτητες. Για παραδειγμα, οι X_i ειναι μετρησεις της συστολικης πιεσης n_1 ασθενων που εχουν υποβληθη στη θεραπεια A και οι Y_i οι αντιστοιχες μετρησεις ~~σε~~ n_2 ασθενων που εχουν υποβληθη στη θεραπεια B.

(α) Αν $\sigma_1^2=\sigma_2^2=\sigma^2$ – αγνωστο – δειξτε οτι ο ΕΠΠ μεγεθους α της $H:\mu_1=\mu_2$ vs $K:\mu_1\neq\mu_2$, ειναι ο ακολουθος <u>δι-δειγματικος ελεγχος – t</u> :

$$d(\underline{X},\underline{Y})=1\left(|T(\underline{X},\underline{Y})|>t_{n_1+n_2-2}^{(1-\alpha/2)}\right),$$

οπου,

$$T(\underline{X},\underline{Y})=\frac{\sqrt{\frac{n_1n_2}{n_1+n_2}}\,(\overline{X}-\overline{Y})}{S_p},$$

οπου, $S_p^2:=\dfrac{(n_1-1)S_X^2+(n_2-1)S_Y^2}{n_1+n_2-2}$, η <u>συγκερασμενη</u> εκτιμητρια της κοινης διασπορας σ^2, και $S_X^2=\frac{1}{n_1-1}\sum_{i=1}^{n_1}(X_i-\overline{X}_{n_1})^2$, το S_Y^2 οριζεται αναλογως.

<u>Σημ.</u> Χωρις την υποθεση : $\sigma_1^2=\sigma_2^2$, το προβλημα δυσκολευει πολυ, ειναι δε γνωστο ως το προβλημα των Behrens και Fisher.

(β) Δειξτε οτι ο ΕΠΠ , μεγεθους α, της

$H: \sigma_1^2 = \sigma_2^2$ vs $K: \sigma_1^2 \neq \sigma_2^2$, ειναι ο ακολουθος ελεγχος - F :

$$d(\underline{X}, \underline{Y}) = 1\left(\frac{S_X^2}{S_Y^2} < F_{n_1-1, n_2-2}(\alpha/2) \right)$$

$$+ 1\left(\frac{S_X^2}{S_Y^2} > F_{n_1-1, n_2-2}(1-\alpha/2) \right).$$

Απαντ. (α) $\Theta = \mathbb{R} \times \mathbb{R} \times (0, +\infty)$, $\Theta_0 = \mathbb{R} \times (0, +\infty)$, και αρα, $\dim \Theta_0 = 2 < \dim \Theta = 3$. Επισης, η $L(\theta \mid \underline{X}, \underline{Y})$, $\theta \in \Theta$ ειναι συνεχης.

(i) Υπο την $H: \mu_1 = \mu_2$ τα δυο δειγματα $\underline{X}, \underline{Y}$ προερχονται απο την ιδια κατανομη και αρα,

$$\hat{\mu}_1 = \hat{\mu}_2 = \hat{\mu}_0 = \frac{\Sigma_{i=1}^{n_1} X_i + \Sigma_i^{n_2} Y_i}{n_1 + n_2} = \frac{n_1}{n_1+n_2} \overline{X}_{n_1} + \frac{n_2}{n_1+n_2} \overline{Y}_{n_2},$$

$$\hat{\sigma}_0^2 = \frac{1}{n_1+n_2} \left\{ \Sigma_{i=1}^{n_1} (X_i - \hat{\mu}_0)^2 + \Sigma_{i=1}^{n_2} (Y_i - \hat{\mu}_0)^2 \right\}.$$

Εχουμε, λοιπον,

$$L(\Theta_0 \mid \underline{X}, \underline{Y}) = L(\hat{\mu}_0, \hat{\sigma}_0^2 \mid \underline{X}, \underline{Y}) = (2\pi \hat{\sigma}_0^2)^{-\frac{n_1+n_2}{2}} \exp\left\{-\frac{n_1+n_2}{2}\right\}.$$

(ii) Για τον υπολογισμο της $L(\Theta \mid \underline{X}, \underline{Y})$ εχουμε:

$$\hat{\mu}_1 = \overline{X}_{n_1}, \quad \hat{\mu}_2 = \overline{X}_{n_2}, \quad \hat{\sigma}^2 = \frac{1}{n_1+n_2} \left\{ \Sigma_{i=1}^{n_1} (X_i - \overline{X}_{n_1})^2 + \Sigma_{i=1}^{n_2} (Y_i - \overline{Y}_{n_2})^2 \right\},$$

και αρα,

$$L(\hat{\mu}_1, \hat{\mu}_2, \hat{\sigma}^2 \mid \underline{X}, \underline{Y}) = (2\pi \hat{\sigma}^2)^{-\frac{n_1+n_2}{2}} \exp\left\{-\frac{n_1+n_2}{2}\right\}.$$

(iii) Τοτε, $\lambda(\underline{X}, \underline{Y}) = \dfrac{L(\Theta \mid \underline{X}, \underline{Y})}{L(\Theta_0 \mid \underline{X}, \underline{Y})} = \left(\dfrac{\hat{\sigma}_0^2}{\hat{\sigma}^2} \right)^{\frac{n_1+n_2}{2}} =$

$$= \left(\frac{\Sigma_{i=1}^{n_1} (X_i - \hat{\mu}_0)^2 + \Sigma_{i=1}^{n_2} (Y_i - \hat{\mu}_0)^2}{\Sigma_{i=1}^{n_1} (X_i - \overline{X})^2 + \Sigma_{i=1}^{n_2} (Y_i - \overline{Y})^2} \right)^{\frac{n_1+n_2}{2}} =$$

$$= \left\{ \frac{\Sigma_{i=1}^{n_1} (X_i - \overline{X})^2 + n_1 (\overline{X} - \hat{\mu}_0)^2 + \Sigma_{i=1}^{n_2} (Y_i - \overline{Y})^2 + n_2 (\overline{Y} - \hat{\mu}_0)^2}{\Sigma_{i=1}^{n_1} (X_i - \overline{X})^2 + \Sigma_{i=1}^{n_2} (Y_i - \overline{Y})^2} \right\}^{\frac{n_1+n_2}{2}} =$$

$$= \left\{ 1 + \frac{n_1+n_2-2}{S_p^2} \left[\frac{n_1 n_2^2}{(n_1+n_2)^2} (\bar{X}-\bar{Y})^2 + \frac{n_2 n_1^2}{(n_1+n_2)^2} (\bar{X}-\bar{Y})^2 \right] \right\}^{\frac{n_1+n_2}{2}}$$

$$= \left\{ 1 + (n_1+n_2-2) \frac{n_1 n_2}{n_1+n_2} \frac{(\bar{X}-\bar{Y})^2}{S_p^2} \right\}^{\frac{n_1+n_2}{2}} =$$

$$= \left\{ 1 + (n_1+n_2-2) \left[T(\underline{X}, \underline{Y}) \right]^2 \right\}^{\frac{n_1+n_2}{2}}$$ αυξουσα συναρτηση

της στατιστικης $|T(\underline{X}, Y)|$. Αρα,

$$\lambda(\underline{X}, \underline{Y}) > c \iff |T(\underline{X}, \underline{Y})| > c$$, δηλαδη,

$$d(\underline{X}, \underline{Y}) = 1\left(|T(\underline{X}, \underline{Y})| > c \right)$$, οπου,

$$\alpha = P_H\left(|T(\underline{X}, \underline{Y})| > c \right) = 1 - P_H\left(|T| \leq c \right)$$
$$= 1 - F_{t_{n_1+n_2-2}}(c) + F_{t_{n_1+n_2-2}}(-c) \iff$$

$$\iff F_{t_{n_1+n_2-2}}(c) = 1 - \alpha/2 \iff c = c_\alpha = t_{n_1+n_2-2}(1-\alpha/2),$$

διοτι $$T(\underline{X}, \underline{Y}) = \frac{\left[(\bar{X}-\mu_0) - (\bar{Y}-\mu_0) \right] \Big/ \sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}}{\sqrt{\dfrac{\Sigma_{i=1}^{n_1} (X_i-\bar{X})^2/\sigma^2 + \Sigma_{i=1}^{n_2} (Y_i-\bar{Y})^2/\sigma^2}{n_1+n_2-2}}},$$

και, υπο την H, ο μεν αριθμητης ακολουθει την $N(0,1)$, ο δε αριθμητης του παρονομαστου την $\chi^2_{n_1-1} + \chi^2_{n_2-1} \stackrel{d}{=} \chi^2_{n_1+n_2-2}$ (εφοσον οι δυο χ^2-μεταβλητες ειναι ανεξαρτητες) και οι δυο τους ειναι ανεξαρτητοι – βλ. (1.62), και αρα η $T \sim t_{n_1+n_2-2}$ – βλ. (1.59).

(β) $\Theta_0 = \mathbb{R} \times \mathbb{R} \times (0, +\infty)$, $\Theta = \mathbb{R} \times \mathbb{R} \times (0, +\infty) \times (0, +\infty)$, και αρα $\dim \Theta_0 = 3 < \dim \Theta = 4$ και $L(\theta \mid \underline{X}, \underline{Y})$, $\theta \in \Theta$ συνεχης συναρτηση. Ο ΕΠΠ βασιζεται λοιπον στο $\lambda(\underline{X}, \underline{Y})$.

(i) Υπο την H, εχουμε,

$$\hat{\mu}_1 = \bar{X}, \quad \hat{\mu}_2 = \bar{Y}, \quad \hat{\sigma}_1^2 = \hat{\sigma}_2^2 \equiv \hat{\sigma}_0^2 = \frac{1}{n_1+n_2}\left\{\Sigma_{i=1}^{n_1}(X_i-\bar{X})^2 + \Sigma_{i=1}^{n_2}(Y_i-\bar{Y})^2\right\},$$

και αρα,

$$L(\Theta_0 \mid \underline{X}, \underline{Y}) = (2\pi\hat{\sigma}_0^2)^{-\frac{n_1+n_2}{2}} \exp\left\{-\frac{n_1+n_2}{2}\right\}.$$

(ii) Για τον υπολογισμό της $L(\Theta \mid \underline{X}, \underline{Y})$, εχουμε,

$$\hat{\mu}_1 = \bar{X}, \quad \hat{\mu}_2 = \bar{Y}, \quad \hat{\sigma}_1^2 = \frac{1}{n_1}\Sigma_{i=1}^{n_1}(X_i-\bar{X})^2, \quad \hat{\sigma}_2^2 = \frac{1}{n_2}\Sigma_{i=1}^{n_2}(Y_i-\bar{Y})^2,$$

και αρα,

$$L(\Theta \mid \underline{X}, \underline{Y}) = L(\hat{\mu}_1, \hat{\mu}_2, \hat{\sigma}_1^2, \hat{\sigma}_2^2 \mid \underline{X}, \underline{Y}) =$$

$$= (2\pi\hat{\sigma}_1^2)^{-\frac{n_1}{2}}(2\pi\hat{\sigma}_2^2)^{-\frac{n_2}{2}} \exp\left\{-\frac{n_1+n_2}{2}\right\}.$$

(iii) Εχουμε, λοιπον,

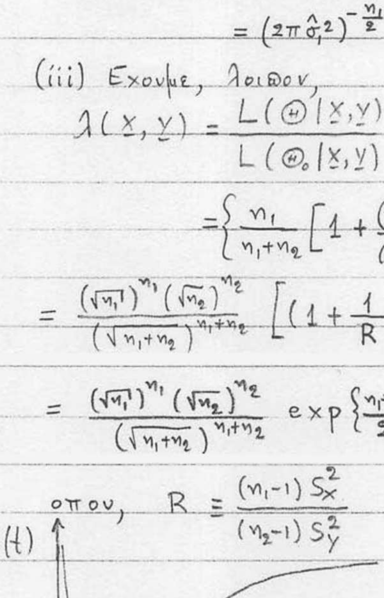
$$\lambda(\underline{X}, \underline{Y}) = \frac{L(\Theta \mid \underline{X}, \underline{Y})}{L(\Theta_0 \mid \underline{X}, \underline{Y})} = \left(\frac{\hat{\sigma}_0^2}{\hat{\sigma}_1^2}\right)^{\frac{n_1}{2}}\left(\frac{\hat{\sigma}_0^2}{\hat{\sigma}_2^2}\right)^{\frac{n_2}{2}} =$$

$$= \left\{\frac{n_1}{n_1+n_2}\left[1 + \frac{(n_2-1)S_Y^2}{(n_1-1)S_X^2}\right]\right\}^{\frac{n_1}{2}}\left\{\frac{n_2}{n_1+n_2}\left[1+\frac{(n_1-1)S_X^2}{(n_2-1)S_Y^2}\right]\right\}^{\frac{n_2}{2}}$$

$$= \frac{(\sqrt{n_1})^{n_1}(\sqrt{n_2})^{n_2}}{(\sqrt{n_1+n_2})^{n_1+n_2}}\left[\left(1+\frac{1}{R}\right)^{n_1}(1+R)^{n_2}\right]^{1/2} =$$

$$= \frac{(\sqrt{n_1})^{n_1}(\sqrt{n_2})^{n_2}}{(\sqrt{n_1+n_2})^{n_1+n_2}} \exp\left\{\frac{n_1+n_2}{2} g(R)\right\},$$

οπου, $R \equiv \dfrac{(n_1-1)S_X^2}{(n_2-1)S_Y^2}$, $g(t) \equiv \log(1+t) - \dfrac{n_1}{n_1+n_2}\log t$.



Αρα,

$$\lambda(\underline{X}, \underline{Y}) > c \Leftrightarrow g(R) > c$$

$$\Leftrightarrow R < c_1 \text{ ή } R > c_2$$

$$\Leftrightarrow T < c_1 \text{ ή } T > c_2,$$

οπου,

$$T := \frac{S_X^2}{S_Y^2} = \frac{\dfrac{\Sigma_{i=1}^{n_1}(X_i-\bar{X})^2/\sigma_0^2}{n_1-1}}{\dfrac{\Sigma_{i=1}^{n_2}(Y_i-\bar{Y})^2/\sigma_0^2}{n_2-1}} \overset{H}{\sim} \frac{\chi^2_{n_1-1}/(n_1-1)}{\chi^2_{n_2-1}/(n_2-1)} \overset{d}{=}$$

$$\overset{d}{=} \mathcal{F}_{n_1-1, n_2-1} \quad \underline{\text{υπο την } H}.$$

Για τον υπολογισμό των κρίσιμων σταθερών του ΕΠΠ :

$$d(\underline{X}, \underline{Y}) = 1(T < c_1) + 1(T > c_2),$$

έχουμε, $P_H(T < c_1) + P(T > c_2) = \alpha$,

δηλαδή, $c_1 = \mathcal{F}_{n_1-1, n_2-1}(\alpha_1)$, $c_2 = \mathcal{F}_{n_1-1, n_2-2}(1-\alpha_2)$, $\alpha_1 + \alpha_2 = \alpha$.

Άρα, θέτοντας – για τη συμμετρία του ελέγχου –
$\alpha_1 = \alpha_2 = \alpha/2$, καταλήγουμε στον συμμετρικό ΕΠΠ μεγέθους α, του μέρους (β) της εκφωνήσεως.

(4.78) Άσκηση. Έστω $X_1, \dots, X_{n_1}$ α.ι. $N(0, \sigma_1^2)$ και $Y_1, \dots, Y_{n_2}$ α.ι. $N(0, \sigma_2^2)$. Δείξτε ότι ο συμμετρικός ΕΠΠ μεγέθους α, της $H: \sigma_1^2 = \sigma_2^2$ vs $K: \sigma_1^2 \neq \sigma_2^2$, είναι ο ακόλουθος:

$$d(\underline{x}, \underline{y}) = 1\left(T < \mathcal{F}_{n_1, n_2}(\alpha/2)\right) + 1\left(T > \mathcal{F}_{n_1, n_2}(1-\alpha/2)\right),$$

όπου, $T := \dfrac{\frac{1}{n_1}\sum_{i=1}^{n_1} X_i^2}{\frac{1}{n_2}\sum_{i=1}^{n_2} Y_i^2}$.

(4.79) Άσκηση. Υπό τις προϋποθέσεις του (4.75), δείξτε ότι τα ακόλουθα στοχαστικά διαστήματα, είναι διαστήματα εμπιστοσύνης επιπέδου α, βασισμένα σε ΕΠΠ, για τις παραμέτρους μ, σ^2, αντιστοίχως:

(α) $\left[\overline{X}_n - \dfrac{t_{n-1}(1-\alpha/2)\, S_n}{\sqrt{n}}, \ \overline{X}_n + \dfrac{t_{n-1}(1-\alpha/2)\, S_n}{\sqrt{n}}\right]$,

(β) $\left[\dfrac{(n-1) S_n^2}{\chi^2_{n-1}(1-\alpha/2)}, \ \dfrac{(n-1) S_n^2}{\chi^2_{n-1}(\alpha/2)}\right]$.

(4.80) Παράδειγμα: Έστω $\underline{X}$, $\underline{Y}$ τα δυο ανεξάρτητα κανονικά δείγματα του (4.77) με $\sigma_1^2 = \sigma_2^2$. Θα δείξουμε ότι το στοχαστικο διαστημα $I(\underline{X},\underline{Y}):=$

$$\left[\bar{X}-\bar{Y}-\sqrt{\frac{n_1+n_2}{n_1 n_2}}\,S_p\, t_{n_1+n_2-2}^{(1-\alpha/2)}\ ,\ \bar{X}-\bar{Y}+\sqrt{\frac{n_1+n_2}{n_1 n_2}}\,S_p\, t_{n_1+n_2-2}^{(1-\alpha/2)}\right],$$

ειναι ενα επιπεδου α, διαστημα εμπιστοσυνης για την διαφορα $\theta:=\mu_1-\mu_2$ των μεσων τιμων των δυο κανονικων κατανομων, βασιζεται δε στον μεγεθους α ΕΠΠ του $H:\ \theta=\theta_0$ vs $K:\ \theta\neq\theta_0$.

<u>Απαντ.</u> Παρατηρουμε οτι αν $Z_i:=X_i-\theta_0$, $i=1,\dots,n_1$, οι $Z_1,\dots,Z_{n_1}$ ειναι α.ι. $N(\mu_1^*,\sigma^2)$, οπου, $\mu_1^*:=\mu_1-\theta_0$. Αρα ο ελεγχος της H vs K ειναι ο ιδιος με τον της $\mu_1^*=\mu_2$ vs $\mu_1^*\neq\mu_2$, με τον οποιον ασχοληθηκαμε στο (4.77), η ελεγχοσυναρτηση του οποιου ειναι η:

$$d(\underline{x},\underline{y})=1\left(\frac{\sqrt{\frac{n_1 n_2}{n_1+n_2}}\,|\bar{x}-\theta_0-\bar{y}|}{S_p}>t_{n_1+n_2-2}^{(1-\alpha/2)}\right).$$

και αρα,

$$1-\alpha=P_{\theta_0}\left(\sqrt{\frac{n_1 n_2}{n_1+n_2}}\,\frac{|\bar{X}-\bar{Y}-\theta_0|}{S_p}\leq t_{n_1+n_2-2}^{(1-\alpha/2)}\right)$$

$$=P_{\theta_0}\left(\theta_0\in I(\underline{X},\underline{Y})\right),$$

και αυτο $\forall\,\theta_0\in\Theta=\mathbb{R}$, αρα το $I(\underline{X},\underline{Y})$ ειναι επιπεδου α διαστημα εμπιστοσυνης, προερχομενο απο ΕΠΠ, της παραμετρου $\theta\in\Theta=\mathbb{R}$.

<u>(4.81) Ασκηση.</u> Υπο τις προϋποθεσεις του (4.77 β) βρητε ενα επιπεδου α διαστημα εμπιστοσυνης για την παραμετρο $\theta:=\sigma_1^2/\sigma_2^2\in\Theta=(0,+\infty)$, βασισμενο σε Ε.Π.Π..

(4.82) Παρατηρηση. Στην περιπτωση του (4.75 a), εχουμε οτι

$$2\log\lambda(\underline{X}) = n\log\left\{1 + \frac{n(\bar{X}_n-\mu_0)^2}{\sum_{i=1}^n (X_i-\bar{X}_n)^2}\right\}$$

$$\approx [\sqrt{n}(\bar{X}_n-\mu_0)]^2/\hat{\sigma}_n^2 \quad ,\text{μια και } \log(1+z)\approx z .$$

Εχουμε δε οτι $\hat{\sigma}_n^2 \xrightarrow[n\to\infty]{P} \sigma^2$ — βλ. (3.123), και, υπο την H,

$$\sqrt{n}(\bar{X}_n-\mu_0) \xrightarrow[n\to\infty]{d} N(0,\sigma^2) \quad ,\text{ απο το ΚΟΘ}.$$

Ετσι, απο το Θεωρημα του Slutsky:

$$2\log\lambda(\underline{X}) \approx \frac{\sigma^2}{\hat{\sigma}_n^2}\cdot\left[\frac{\sqrt{n}(\bar{X}_n-\mu_0)}{\sigma}\right]^2$$

$$\xrightarrow[n\to]{d} [N(0,1)]^2 \stackrel{d}{=} \chi_1^2 \quad , \text{ \underline{υπο την } } H:\mu=\mu_0 .$$

Αυτο το αποτελεσμα ειναι μια ειδικη περιπτωση του κατωθι γενικοτερου αποτελεσματος, το οποιο δiδουμε χωρις αποδειξη:

Κατω απο συνθηκες ομαλοτητος του μοντελου $f(\cdot|\theta)$, $\theta\in\Theta$ (κυριως, παραγωγισιμοτητα ως προς θ), εχουμε οτι για την στατιστικη του πηλικου πιθανοφανειων $\lambda(\underline{X})$, του ΕΠΠ της $H:\theta\in\Theta_0$ vs $K:\theta\in\Theta\setminus\Theta_0$, με $d:=\dim\Theta-\dim\Theta_0 \geqslant 1$, ισχυει το εξης:

$$(4.83) \quad 2\log\lambda(\underline{X}) \xrightarrow[n\to\infty]{d} \chi_d^2 .$$

Αρα, ενας προσεγγιστικος (ασυμπτωτικος) ελεγχος της H vs K, μεγεθους α, ειναι ο εξης:

$$(4.84) \quad d_\alpha(\underline{X}) = 1\left(2\log\lambda(\underline{X}) > \chi_d^2(1-\alpha)\right).$$

Αυτος ο προσεγγιστικος ΕΠΠ εχει αποδειχθη πολυ καλος στην πραξη.

(4.85) Παραδειγμα. Προβαλεται προς ελεγχο η υποθεση οτι τα ποσοστα p_1, p_2, p_3, p_4 της εκπροσωπησης των τεσσαρων βασικων κατηγοριων αιματος A_1, A_2, A_3, A_4 σε ενα πληθυσμο ειναι:

$$H: p_1 = 0.16,\ p_2 = 0.48,\ p_3 = 0.20,\ p_4 = 0.16 \text{ vs } K: \text{αρνηση της } H.$$

Ειναι δεκτη αυτη η υποθεση σε επιπεδο $\alpha = 0.05$, αν ενα δειγμα μεγεθους $n = 770$, παρουσιασε $n_1 = 180$, $n_2 = 360$, $n_3 = 132$ και $n_3 = 98$;

Απαντ. Θα θεσουμε αυτο το συγκεκριμμενο προβλημα σε γενικοτερο πλαισιο:

Εστω $\underline{N} = (N_1, \dots, N_k) \sim$ Πολυωνυμικη$(n, p_1, \dots, p_k)$, δηλαδη,

$$P(N_i = n_i,\ i=1,\dots,k) = \frac{n!}{n_1! \cdots n_k!} p_1^{n_1} \cdots p_k^{n_k},$$

με $\Sigma_{i=1}^{k} n_i = n$ και $\Sigma_{i=1}^{k} p_i = 1$, $n_i \geq 0$, $i=1,\dots,k$

Θελουμε να ελεγξουμε την υποθεση:

$$H: p_1 = \pi_1, \dots, p_k = \pi_k \text{ vs } K: p_j \neq \pi_j \text{ για καποιο } j \in \{1,\dots,k\},$$

οπου, $\pi_1, \dots, \pi_{k-1} \in (0,1)$, δοσμενα, και $\pi_k = 1 - \sum_{i=1}^{k-1} \pi_i$.

$\theta = (p_1, \dots, p_{k-1}) \in \Theta = (0,1)^{k-1}$, δηλαδη, $\dim \Theta = k-1$ και $\Theta_0 = \{(\pi_1, \dots, \pi_{k-1})\}$, δηλαδη, $\dim \Theta_0 = 0$.

Για να κατασκευασουμε τον ΕΠΠ, αρκει να υπολογισουμε τις ε.μ.π. των $(p_1, \dots, p_{k-1}) \in \Theta$:

$$\ell(p_1, \dots, p_{k-1}) = \log \binom{n}{n_1, \dots, n_k} + \Sigma_{i=1}^{k-1} n_i \log p_i + (n - \Sigma_{i=1}^{k-1} n_i) \log(1 - \Sigma_{i=1}^{k-1} p_i)$$

$$\frac{\partial}{\partial p_j} \ell(p_1, \dots, p_{k-1}) = \frac{n_j}{p_j} + n_k \frac{-1}{p_k} = 0, \quad j=1,\dots,k-1$$

$$\Rightarrow \begin{cases} \dfrac{n_j}{\hat{p}_j} = \dfrac{n_k}{\hat{p}_k}, & j=1,\dots,k-1 \\ \hat{p}_k = 1 - \Sigma_{i=1}^{k-1} \hat{p}_j. \end{cases}$$

Αρα, $\begin{cases} \hat{p}_j = n_j \dfrac{\hat{p}_k}{n_k} & , j=1,\dots,k-1 \\ \sum_{i=1}^{k} \hat{p}_i = 1 & , \end{cases}$

και αρα, $1 = \sum_{j=1}^{k} \hat{p}_j = \left(\sum_{j=1}^{n} n_j\right) \dfrac{\hat{p}_k}{n_k} = n \dfrac{\hat{p}_k}{n_k}$

$$\Rightarrow \hat{p}_k = \frac{n_k}{n} \quad \Rightarrow \quad \hat{p}_j = n_j \frac{n_k/n}{n_k} = \frac{n_j}{n} \ , \ j=1,\dots,k-1.$$

Δηλαδη, $\hat{p}_j = \dfrac{n_j}{n}$, $j=1,\dots,k$,

και αρα,

$$L(\Theta \mid \underline{N}) = \binom{n}{N_1,\dots,N_k} \prod_{j=1}^{k} \left(\frac{N_j}{n}\right)^{N_j} ,$$

$$L(\Theta_0 \mid \underline{N}) = \binom{n}{N_1,\dots,N_k} \prod_{j=1}^{k} \pi_j^{N_j} .$$

Αρα,

$$\lambda(\underline{N}) = \prod_{j=1}^{k} \left(\frac{N_j}{n\pi_j}\right)^{N_j} \Rightarrow$$

$$(4.86) \quad d_\alpha(\underline{N}) = 1\left(2 \log \lambda(\underline{N}) > \chi^2_{k-1}(1-\alpha)\right)$$

$$= 1\left(2 \sum_{j=1}^{k} N_j \log\left(\frac{N_j}{n\pi_j}\right) > \chi^2_{k-1}(1-\alpha)\right)$$

Στην συγκεκριμενη μας περιπτωση,

$$2 \log \lambda(\underline{n}) = 2 \sum_{j=1}^{4} n_j \log\left(\frac{n_j}{n\pi_j}\right) \doteq 36 >$$

$$> 7.81473 = \chi^2_3(0.95) ,$$

και αρα απορριπτουμε (ανετα) την H.

Υπαρχει μια εξαιρετικα καλη προσεγγιση του ελεγχου $d_\alpha(\underline{N})$, η οποια ειναι πολυ ευκολο να υπολογισθη, και η οποια βασιζεται στην προσεγγιση:

$$\log(1+z) = z - \frac{1}{2}z^2 + O(z^3) \ , \ z > -1 .$$

Εχουμε,

$$2 \log \lambda(\underline{n}) = 2\sum_{j=1}^{k} n_j \log\left(1 + \frac{n_j - n\pi_j}{n\pi_j}\right) \doteq$$

$$\doteq 2\sum_{j=1}^{k} n_j\left(\frac{n_j-n\pi_j}{n\pi_j}\right)-\sum_{j=1}^{k}\frac{n_j}{n\pi_j}\frac{(n_j-n\pi_j)^2}{n\pi_j}=$$

οπου, εχουμε παραλειψει τους ορους $\left(\frac{n_j-n\pi_j}{n\pi_j}\right)^m$, για $m\geqslant 3$, δεδομενου οτι $EN_j=n\pi_j$ και αρα οι αριθμητες ειναι κοντα στο μηδεν,

$$=2\sum_{j=1}^{k}\frac{(n_j-n\pi_j)^2}{n\pi_j}-\sum_{j=1}^{k}\frac{(n_j-n\pi_j)^3}{(n\pi_j)^2}-\sum_{j=1}^{k}\frac{(n_j-n\pi_j)^2}{n\pi_j}$$

$$=\sum_{j=1}^{k}\frac{(n_j-n\pi_j)^2}{n\pi_j}-\frac{1}{\sqrt{n}}\sum_{j=1}^{k}\frac{[\sqrt{n}(\hat{p}_j-\pi_j)]^3}{\pi_j^2}\doteq\sum_{j=1}^{k}\frac{(n_j-n\pi_j)^2}{n\pi_j}.$$

Αρα,

$$\chi^2:=\sum_{j=1}^{k}\frac{(N_j-n\pi_j)^2}{n\pi_j}\doteq 2\sum_{j=1}^{k}N_j\log\left(\frac{N_j}{n\pi_j}\right)\overset{\cdot}{\sim}\chi^2_{k-1},$$

και εχουμε λοιπον καταληξει στο γνωστο ελεγχο-χ^2:

$$(4.87)\quad d'_{\alpha}(\underline{N})=1\left(\chi^2>\chi^2_{k-1}(1-\alpha)\right),$$

ο οποιος ειναι γνωστος και σαν ελεγχος καλης εφαρμογης της υποθεσης H στα δεδομενα $\underline{N}=\underline{n}$.
Στο συγκεκριμμενο μας παραδειγμα,

$$\chi^2\approx 35>\chi^2_3(0.95)=7.81473,$$

και αρα και παλι απορριπτουμε (ανετα) την H.

(4.88) Ασκηση. Κατα τη θεωρια της κληρονομικοτητας του Mendel, ενα υβριδιο λουλουδιων θα δωσει ανθος χρωματος κοκκινου, κιτρινου, ασπρου με πιθανοτητες $\frac{9}{16}, \frac{3}{16}, \frac{4}{16}$ αντιστοιχως. Ενα πειραμα με 320 φυτα, παρουσιασε 150, 70 και 100 κοκκινα, κιτρινα και ασπρα ανθη αντιστοιχως. Ειναι η θεωρια δικαιολογημενη, σε επιπεδο $\alpha=0.05$. Χρησιμοποιειστε τον ελεγχο (4.86) και τον ελεγχο (4.87).