28
ΜΠΛΕ

ΠΑΝΕΠΙΣΤΗΜΙΟ ΙΩΑΝΝΙΝΩΝ

Η ΑΝΙΣΟΤΗΤΑ CRAMER-RAO
ΚΑΙ
ΟΙ ΕΦΑΡΜΟΓΕΣ ΤΗΣ

ΚΟΝΤΟΓΙΑΝΝΗ ΑΡΙΣΤΟΥΛΑ

ΜΑΘΗΜΑΤΙΚΟΣ

ΜΕΤΑΠΤΥΧΙΑΚΗ ΔΙΑΤΡΙΒΗ

ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ

ΤΜΗΜΑ ΜΑΘΗΜΑΤΙΚΩΝ

ΤΟΜΕΑΣ ΠΙΘΑΝΟΤΗΤΩΝ - ΣΤΑΤΙΣΤΙΚΗΣ

ΚΑΙ ΕΠΙΧΕΙΡΗΣΙΑΚΗΣ ΕΡΕΥΝΑΣ

ΙΩΑΝΝΙΝΑ 2006

Η παρούσα Μεταπτυχιακή Διατριβή εκπονήθηκε το Ακαδημαϊκό έτος 2005-2006, στα πλαίσια του Προγράμματος Μεταπτυχιακών σπουδών του Τμήματος Μαθηματικών της Σχολής Θετικών Επιστημών του Πανεπιστημίου Ιωαννίνων υπό την επίβλεψη του Καθηγητή Κοσμά Φερεντίνου.

ΤΡΙΜΕΛΗΣ ΕΠΙΤΡΟΠΗ ΚΡΙΣΗΣ ΔΙΑΤΡΙΒΗΣ

Φερεντίνος Κοσμάς, Καθηγητής, Τμήμα Μαθηματικών, Σχολή Θετικών Επιστημών, Πανεπιστήμιο Ιωαννίνων. (Επιβλέπων Καθηγητής)

Ζωγράφος Κωνσταντίνος, Αναπληρωτής Καθηγητής, Τμήμα Μαθηματικών, Σχολή Θετικών Επιστημών, Πανεπιστήμιο Ιωαννίνων.

Καρακώστας Κωνσταντίνος, Αναπληρωτής Καθηγητής, Τμήμα Μαθηματικών, Σχολή Θετικών Επιστημών, Πανεπιστήμιο Ιωαννίνων.

Ευχαριστίες

Η παρούσα διατριβή αποτελεί το τελευταίο μέρος για την απόκτηση του Μεταπτυχιακού Διπλώματος Ειδίκευσης στη Στατιστική και Επιχειρησιακή Έρευνα του τομέα Πιθανοτήτων – Στατιστικής και Επιχειρησιακής Έρευνας του τμήματος Μαθηματικών του Πανεπιστημίου Ιωαννίνων. Αρχικά θα ήθελα να ευχαριστήσω τον κ. Κοσμά Φερεντίνο καθηγητή του τμήματος Μαθηματικών, επιβλέποντα της μεταπτυχιακής μου διατριβής, που μου υπέδειξε το θέμα της και με βοήθησε στην συγγραφή της . Επίσης θα ήθελα να τον ευχαριστήσω για τις απαντήσεις στις απορίες μου καθώς και για την υπομονή και την ανοχή απέναντι μου.

Θα ήθελα να ευχαριστήσω τον κ. Κωνσταντίνο Καρακώστα, μέλος της τριμελούς επιτροπής, για τις εύστοχες παρατηρήσεις και διορθώσεις του. Ιδιαίτερα θα ήθελα να ευχαριστήσω τον κ. Κωνσταντίνο Ζωγράφο, επίσης μέλος της τριμελούς επιτροπής, για τις αναλυτικές διορθώσεις και σημαντικές παραινέσεις όσον αφορά το σωστό γράψιμο της διατριβής.

Ευχαριστώ, επίσης, το σύνολο των μελών του τομέα Πιθανοτήτων – Στατιστικής και Επιχειρησιακής Έρευνας του τμήματος Μαθηματικών του Πανεπιστημίου Ιωαννίνων για την συνεργασία που είχαμε κατά την διάρκεια των σπουδών μου για την απόκτηση του Μεταπτυχιακού Διπλώματος Ειδίκευσης . Κρίνω απαραίτητο να ευχαριστήσω το σύνολο των μελών του τμήματος Μαθηματικών του Πανεπιστημίου Ιωαννίνων για τις γνώσεις που αποκόμισα κατά τη διάρκεια των προπτυχιακών και μεταπτυχιακών μου σπουδών.

Ένα μεγάλο ευχαριστώ οφείλω στην οικογένεια μου για την υλική και ηθική υποστήριξη τους κατά την διάρκεια των σπουδών μου. Τέλος, ευχαριστώ όλους τους φίλους και τους ανθρώπους που πιστεύουν σε μένα και με στήριξαν ακόμα και τις στιγμές που πίστευα ότι όλα είχαν χαθεί. Χωρίς αυτούς η διατριβή αυτή δεν θα είχε ποτέ ολοκληρωθεί.

Κοντογιάννη Αριστούλα
Ιωάννινα 2006

Πρόλογος

Η μεταπτυχιακή αυτή διατριβή έχει ως θέμα την ανισότητα Cramer-Rao και τις εφαρμογές της. Στο εισαγωγικό κεφάλαιο δίνονται βασικές έννοιες της στατιστικής συμπερασματολογίας οι οποίες χρησιμοποιούνται στα κεφάλαια που ακολουθούν. Στο πρώτο κεφάλαιο αποδεικνύεται η ανισότητα Cramer-Rao για την μονοδιάστατη και για την πολυδιάστατη περίπτωση και δίνονται τα ανάλογα πορίσματα για το πότε ισχύει η ισότητα στην ανισότητας αυτή. Στο δεύτερο κεφάλαιο μελετάται η εφαρμογή της ανισότητας Cramer-Rao στην εκτιμητική για την εύρεση ενός Α.Ο.Ε.Δ (αμερόληπτος ομοιόμορφα ελαχίστης διασποράς) εκτιμητή. Στο τρίτο κεφάλαιο αναπτύσσονται δύο ακόμα εφαρμογές της ανισότητας Cramer-Rao που αφορούν την εύρεση ενός κάτω φράγματος για την διακύμανση μιας συνάρτησης της τυχαίας μεταβλητής X, η οποία ακολουθεί γνωστή κατανομή, και την ασυμπτωτική κατανομή του Ε.Μ.Π.(εκτιμητή μεγίστης πιθανοφάνειας). Τέλος, στο συμπέρασμα-επίλογο, γίνεται μια ανασκόπηση των όσων παρουσιάστηκαν καθώς και μια αναφορά σε κάποιες άλλες εφαρμογές της ανισότητας Cramer-Rao, που στα πλαίσια της παρούσας διατριβής δεν είναι δυνατόν να αναπτυχθούν.

ΠΕΡΙΕΧΟΜΕΝΑ

ΠΙΝΑΚΑΣ ΣΥΝΤΜΗΣΕΩΝ

τ.μ.	τυχαία μεταβλητή
τ.δ.	τυχαίο δείγμα
σ.π.	συνάρτηση πιθανότητας
σ.π.π.	συνάρτηση πυκνότητας πιθανότητας
α.σ.κ.	αθροιστική συνάρτηση κατανομής
σ.σ.	στατιστική συνάρτηση
Α.Ο.Ε.Δ.	αμερόληπτος ομοιόμορφα ελάχιστης διασποράς
Κ.Φ.C-R	κάτω φράγμα Cramer-Rao
Ε.Μ.Π.	εκτιμητής μέγιστης πιθανότητας
μ.τ.σ.	μέσο τετραγωνικό σφάλμα

Κεφάλαιο 0

Βασικές Έννοιες και Ορισμοί

0.1 Εισαγωγή

Το αντικείμενο της παρούσας διατριβής είναι η ανισότητα Cramer-Rao και οι εφαρμογές της. Στο εισαγωγικό αυτό κεφάλαιο δίνονται οι βασικές έννοιες και οι ορισμοί που χρησιμοποιούνται στα επόμενα κεφάλαια, για την καλύτερη κατανόηση των όσων ακολουθούν.

0.2 Εκτίμηση σε σημείο

Το πρόβλημα της εκτίμησης, όπως θα αντιμετωπιστεί στη συνέχεια, μπορεί να οριστεί ως εξής. Θεωρούμε ότι κάποια χαρακτηριστικά των στοιχείων ενός πληθυσμού μπορούν να αναπαρασταθούν μέσω του τ.δ $\mathbf{X}_1, \mathbf{X}_2, \ldots\ldots, \mathbf{X}_n$ του οποίου η σ.π.π. $f(\mathbf{x}, \boldsymbol{\theta})$ στην περίπτωση της συνεχούς τ.μ και αντίστοιχα η σ.π $p(\mathbf{x}, \boldsymbol{\theta})$ στην περίπτωση της διακριτής είναι γνωστή. Η μορφή της f θεωρείται γνωστή περιέχει όμως μια άγνωστη σταθερά $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots\ldots, \theta_n)'$,η οποία ονομάζεται παράμετρος ενώ το σύνολο των δυνατών τιμών της $\boldsymbol{\theta}$ συμβολίζεται με Θ και ονομάζεται παραμετρικός χώρος. Σκοπός μας λοιπόν, είναι να εκτιμήσουμε αυτή την άγνωστη παράμετρο $\boldsymbol{\theta}$ ή κάποια παραμετρική συνάρτηση αυτής, έστω $g(\boldsymbol{\theta})$, με μία ή περισσότερες στατιστικές συναρτήσεις, έστω $T(\mathbf{x})$. Αυτό μπορεί να γίνει με δυο τρόπους, είτε με την εκτίμηση σε σημείο, είτε με την εκτίμηση σε διάστημα. Στην παρούσα διατριβή θα αναφερθούμε μόνο σε μεθόδους που αφορούν την εκτίμηση σε σημείο.

Όσον αφορά την εκτίμηση σε σημείο, υπάρχουν δυο προβλήματα. Πρώτον να βρούμε ένα στατιστικό, δηλαδή μια συνάρτηση του δείγματος, το οποίο θα χρησιμοποιήσουμε ως εκτιμητή και δεύτερον, μέσω κάποιων αρχών, να επιλέξουμε τον «καλύτερο» μεταξύ πολλών πιθανών εκτιμητών.

Αρχικά, θα αναφερθούμε στις σημαντικότερες αρχές επιλογής εκτιμητών οι οποίες είναι οι:

(i) Αμεροληψία (Unbiasedness)

(ii) Συνέπεια (Consistency)

(iii) Επάρκεια (Efficiency)

(iv) Πληρότητα (Completeness)

Στη συνέχεια γίνεται αναφορά στο κάθε μία από αυτές τις αρχές ξεχωριστά και δίνονται και τα ανάλογα θεωρήματα.

Αμεροληψία

Ορισμός 0.1

Ο εκτιμητής $T(\mathbf{X}_1,\mathbf{X}_2,......,\mathbf{X}_n)$ της συνάρτησης $g(\boldsymbol{\theta})$ λέγεται αμερόληπτος εάν $E[T(\mathbf{X}_1,\mathbf{X}_2,......,\mathbf{X}_n)] = g(\boldsymbol{\theta}), \forall \boldsymbol{\theta} \in \Theta$. Εάν η $g(\boldsymbol{\theta})$ έχει έναν αμερόληπτο εκτιμητή, τότε η συνάρτηση $g(\boldsymbol{\theta})$ λέγεται εκτιμήσιμη συνάρτηση.

Συνέπεια

Ορισμός 0.2

Η ακολουθία των εκτιμητών $\{T_n\}$ (ή ο εκτιμητής T_n) μίας παραμέτρου θ είναι συνεπής εάν, για κάθε $\varepsilon > 0$ ισχύει

$$\lim_{n\to\infty} P(\theta-\varepsilon < T_n < \theta+\varepsilon) = 1$$

για κάθε $\theta \in \Theta$.

Με άλλα λόγια T_n είναι συνεπής, εάν η πιθανότητα, ότι ο T_n θα λάβει τιμή που διαφέρει αριθμητικά από το θ περισσότερο από μια αυθαίρετη σταθερή ποσότητα ε, τείνει στο μηδέν, όταν $n \to \infty$. Βλέπουμε ότι η συνέπεια είναι μια ασυμπτωτική ιδιότητα. Δεν είναι τίποτε άλλο παρά η σύγκλιση τυχαίων μεταβλητών κατά πιθανότητα που είναι γνωστή από την Θεωρία Πιθανοτήτων, δηλαδή $T_n \to \theta$. Το επόμενο θεώρημα δίνει επαρκείς συνθήκες, ώστε ο εκτιμητής T_n να είναι συνεπής.

Θεώρημα 0.1

Έστω T_n μία εκτιμήτρια σ.σ μίας παραμέτρου θ. Εάν

(i) η T_n είναι αμερόληπτη εκτιμήτρια της θ , και

(ii) $Var(T_n) \to 0$ όταν $n \to \infty$,

τότε η T_n είναι συνεπής εκτιμήτρια της θ. Το συμπέρασμα ισχύει και αν ακόμη η T_n είναι ασυμπτωτικά αμερόληπτος εκτιμητής, δηλαδή $E(T_n) \to \theta$ όταν $n \to \infty$.

Επάρκεια

Ο ορισμός που ακολουθεί είναι για τη διαπίστωση της επάρκειας χωρίς να είναι όμως αρκετά εύχρηστος καθώς εμπλέκει τη χρήση της δεσμευμένης πιθανότητας.

Ορισμός 0.3

Έστω η τ.μ $\mathbf{X} = (X_1, X_2, \ldots\ldots, X_n)$ με σ.π $p(\mathbf{x}, \boldsymbol{\theta})$ ή σ.π.π $f(\mathbf{x}, \boldsymbol{\theta})$, $\boldsymbol{\theta} \in \Theta$ και $\mathbf{T} = (T_1(\mathbf{X}), T_2(\mathbf{X}), \ldots\ldots, T_k(\mathbf{X}))$ μια σ.σ της τ.μ $\mathbf{X}$. Τότε λέμε ότι η $\mathbf{T}$ είναι στατιστικά επαρκής για το $\boldsymbol{\theta}$ ή την οικογένεια κατανομών που ορίζει το $\boldsymbol{\theta}$ εάν η δεσμευμένη κατανομή της $\mathbf{X}$, όταν δοθεί $\mathbf{T} = \mathbf{t}$, είναι ανεξάρτητη του $\boldsymbol{\theta}$, για όλες τις τιμές του $\mathbf{t}$ για τις οποίες ορίζεται η δεσμευμένη κατανομή.

Ο επόμενος ορισμός είναι χρήσιμος για να δείξουμε ότι μια σ.σ δεν είναι επαρκής.

Ορισμός 0.4

Έστω $X_1, X_2, \ldots\ldots, X_n$ τ.δ από έναν πληθυσμό με σ.π $p(\mathbf{x}, \boldsymbol{\theta})$ ή σ.π.π $f(\mathbf{x}, \boldsymbol{\theta})$. Μία στατιστική συνάρτηση $\mathbf{T} = \mathbf{T}(\mathbf{x})$ είναι επαρκής για το $\boldsymbol{\theta}$, εάν και μόνον εάν για κάθε άλλη σ.σ $\mathbf{U}$ η κατανομή $\mathbf{U}|\mathbf{T}$ είναι ανεξάρτητη του $\boldsymbol{\theta}$.

Επειδή οι παραπάνω ορισμοί δεν είναι εύχρηστοι και δεν μας δίνουν πληροφορίες για το ποιες σ.σ ενδέχεται να είναι επαρκείς χρησιμοποιούμε το παρακάτω κριτήριο.

Θεώρημα 0.2 (Παραγοντικό Θεώρημα των Neyman-Fisher)

Έστω $X_1, X_2, \ldots\ldots, X_n$ τ.δ από έναν πληθυσμό με σ.π $p(\mathbf{x},\boldsymbol{\theta})$ ή σ.π.π $f(\mathbf{x},\boldsymbol{\theta})$ $\boldsymbol{\theta} \in \Theta \subset \mathbb{R}^r$. Μία στατιστική συνάρτηση

$\mathbf{T} = \mathbf{T}(X_1, X_2, \ldots\ldots, X_n) = (T_1(\mathbf{X}), T_2(\mathbf{X}), \ldots\ldots, T_k(\mathbf{X}))'$ είναι επαρκής για την παράμετρο $\boldsymbol{\theta}$, εάν και μόνον εάν η σ.π ή σ.π.π του δείγματος μπορεί να γραφεί υπό τη μορφή

$$p(\mathbf{x},\boldsymbol{\theta}) = g(\mathbf{T}(\mathbf{x}),\boldsymbol{\theta})h(\mathbf{x})$$

ή

$$f(\mathbf{x},\boldsymbol{\theta}) = g(\mathbf{T}(\mathbf{x}),\boldsymbol{\theta})h(\mathbf{x}).$$

Αντίστοιχα για κάθε $\mathbf{x} \in \mathbb{R}^n$ και κάθε $\boldsymbol{\theta} \in \Theta$, όπου $g(\mathbf{T}(\mathbf{x}),\boldsymbol{\theta})$ είναι μια συνάρτηση που εξαρτάται από το $\mathbf{x}$ μόνο μέσω της $\mathbf{T}$ και $h(\mathbf{x})$ είναι μια άλλη συνάρτηση του $\mathbf{x}$.

Παρατήρηση 0.1

(α)Εάν $\boldsymbol{\theta}^* = \varphi(\boldsymbol{\theta})$ είναι μια αμφιμονοσήμαντη συνάρτηση της παραμέτρου $\boldsymbol{\theta}$ και επιπλέον αν $\mathbf{T}$ είναι μια επαρκής σ.σ για τη $\boldsymbol{\theta}$, τότε $\mathbf{T}$ είναι μια επαρκής σ.σ και για τη $\boldsymbol{\theta}^* = \varphi(\boldsymbol{\theta})$.

(β)Αν $\mathbf{T}^* = \mathbf{T}^*(\mathbf{T})$ είναι μια αμερόληπτη σ.σ του $\mathbf{T}$, και το $\mathbf{T}$ είναι επαρκές για το $\boldsymbol{\theta}$, τότε και το $\mathbf{T}^*$ είναι επαρκές για το $\boldsymbol{\theta}$.

Ένα θεώρημα που μας βοηθά στην επιλογή καλύτερου εκτιμητή με βάση την επάρκεια είναι το Θεώρημα του Blackwell-Rao.

Θεώρημα 0.3(Blackwell-Rao)

Έστω $X_1, X_2, \ldots\ldots, X_n$ τ.δ από έναν πληθυσμό με σ.π $p(x,\theta)$ ή σ.π.π $f(x,\theta), \theta \in \Theta \subset \mathbb{R}^r$ και $\mathbf{T} = (T_1(\mathbf{x}), T_2(\mathbf{x}), \ldots\ldots, T_k(\mathbf{x}))'$ μια επαρκής σ.σ για το θ. Έστω ακόμη $U = U(\mathbf{x})$ μια άλλη αμερόληπτη σ.σ του θ. Τότε η σ.σ $U^* = E(U|\mathbf{T})$ είναι αμερόληπτη σ.σ του θ με $Var(U^*) \leq Var(U)$ για κάθε $\theta \in \Theta$.

Πληρότητα

Ορισμός 0.5

Έστω $X_1, X_2, ..., X_n$ τ.δ από έναν πληθυσμό με σ.π $p(\mathbf{x}, \boldsymbol{\theta})$ ή σ.π.π $f(\mathbf{x}, \boldsymbol{\theta})$. Μία σ.σ $\mathbf{T} = \left(T_1(\mathbf{x}), T_2(\mathbf{x}), ..., T_k(\mathbf{x})\right)'$ ή ακριβέστερα η οικογένεια κατανομών της $\mathbf{T}$ είναι πλήρης, εάν και μόνο εάν για κάθε συνάρτηση $h(\mathbf{T})$ του $\mathbf{T}$ μόνο, η σχέση

$$E\left[h(\mathbf{T})\right] = 0 \text{ για κάθε } \boldsymbol{\theta} \in \Theta \text{ συνεπάγεται } h(\mathbf{t}) = 0 \text{ για κάθε } \mathbf{t},$$

εκτός πιθανόν από τα σημεία για τα οποία η πιθανότητα είναι μηδέν για κάθε $\boldsymbol{\theta} \in \Theta$.

Με άλλα λόγια μια σ.σ είναι πλήρης αν δεν υπάρχουν μη μηδενικοί αμερόληπτοι εκτιμητές του μηδενός ή η $\mathbf{T}$ είναι πλήρης εάν και μόνο εάν ο μόνος αμερόληπτος εκτιμητής του μηδενός, ο οποίος είναι συνάρτηση του $\mathbf{T}$ είναι ο εκτιμητής 0 με πιθανότητα 1.

Οι αρχές δεν είναι άμεσοι μέθοδοι εύρεσης εκτιμητών αλλά έμμεσοι. Για αυτό έχουμε τις μεθόδους ανάμεσα στις οποίες συγκαταλέγονται η μέθοδος της μέγιστης πιθανοφάνειας, η μέθοδος των ροπών, η μέθοδος των ελάχιστων τετραγώνων, η μέθοδος του Bayes και άλλες. Εκτενώς θα αναφερθούμε στην μέθοδο της μέγιστης πιθανοφάνειας.

Μέγιστη Πιθανοφανεία

Ορισμός 0.6

Έστω $\mathbf{x} = (X_1, X_2, ..., X_n)'$ τ.δ από έναν πληθυσμό με σ.π $p(\mathbf{x}, \boldsymbol{\theta})$ ή σ.π.π $f(\mathbf{x}, \boldsymbol{\theta})$. Τότε η συνάρτηση

$$L(\boldsymbol{\theta}|\mathbf{x}) = \prod_{i=1}^{n} p_{X_i}(x_i, \boldsymbol{\theta}) \qquad \text{ή} \qquad L(\boldsymbol{\theta}|\mathbf{x}) = \prod_{i=1}^{n} f_{X_i}(x_i, \boldsymbol{\theta})$$

που θεωρείται ως συνάρτηση της παραμέτρου $\boldsymbol{\theta}$, ονομάζεται συνάρτηση πιθανοφάνειας.

Ορισμός 0.7

Έστω $L(\boldsymbol{\theta}|\mathbf{x})$ η συνάρτηση πιθανοφάνειας ενός τ.δ. $X_1, X_2, ..., X_n$ Ο εκτιμητής $\hat{\boldsymbol{\theta}} = \hat{\boldsymbol{\theta}}(X_1, X_2, ..., X_n)$ λέγεται εκτιμητής μέγιστης πιθανοφάνειας (Ε.Μ.Π) του $\boldsymbol{\theta}$, εάν

$$L(\hat{\boldsymbol{\theta}}|\mathbf{x}) = \max_{\theta \in \Theta} L(\boldsymbol{\theta}|\mathbf{x})$$

όπου Θ ο παραμετρικός χώρος.

<u>Παρατήρηση 0.2</u>

Συνήθως για ευκολία αντί να μεγιστοποιούμε την $L(\boldsymbol{\theta}|\mathbf{x})$ μεγιστοποιούμε ισοδύναμα την $\ln L(\boldsymbol{\theta}|\mathbf{x})$, διότι αυτές οι συναρτήσεις παίρνουν μέγιστο στο ίδιο σημείο. Μπορεί να μην υπάρχει μέγιστο, να υπάρχει ένα ή πολλά.

Είτε εφαρμόζουμε για την εύρεση εκτιμητών αρχές, εμμέσως, είτε μεθόδους, αμέσως, το πρόβλημα στην εκτίμηση είναι το σφάλμα. Ως σφάλμα ενός εκτιμητή T της $g(\boldsymbol{\theta})$ μπορεί να θεωρηθεί η ποσότητα $|T-g(\boldsymbol{\theta})|$ ή, αντί αυτής το τετράγωνο $|T-g(\boldsymbol{\theta})|^2$ το οποίο ονομάζεται τετραγωνικό σφάλμα. Το $E[T-g(\boldsymbol{\theta})]^2$ λέγεται μέσο τετραγωνικό σφάλμα του εκτιμητή και συμβολίζεται με το $\mu\tau\sigma(T)$.

Το μέσο τετραγωνικό σφάλμα ενός εκτιμητή T γράφεται ως εξής:

$$\begin{aligned}\mu\tau\sigma(T)&=E[T-g(\boldsymbol{\theta})]^2=E\{[T-E(T)]-[g(\boldsymbol{\theta})-E(T)]\}^2\\&=E[T-E(T)]^2+E[g(\boldsymbol{\theta})-E(T)]^2-2E\{[T-E(T)][g(\boldsymbol{\theta})-E(T)]\}\\&=Var(T)+[g(\boldsymbol{\theta})-E(T)]^2.\end{aligned}$$

Τελικά

$$\mu\tau\sigma(T)=Var(T)+b^2(T),$$

όπου $b^2(T)$ είναι το ποσό μεροληψίας. Από την παραπάνω σχέση παρατηρούμε ότι αν το ποσό μεροληψίας είναι μηδέν, δηλαδή αν έχουμε αμερόληπτους εκτιμητές, τότε το μέσο τετραγωνικό σφάλμα ισούται με τη διακύμανση. Έτσι αν έχουμε αμερόληπτους εκτιμητές και θέσουμε τη συνθήκη η διακύμανση να είναι όσο το δυνατόν μικρότερη, τότε βάσει των παραπάνω ζητούμενο είναι η ελαχιστοποίηση του μέσου τετραγωνικού σφάλματος εφόσον στην περίπτωση αυτή θα έχουμε $E[T-g(\boldsymbol{\theta})]^2=Var(T)$.

Το κριτήριο της αμεροληψίας καθορίζει μια μεγάλη κλάση εκτιμητών, με την έννοια ότι υπάρχουν πολλοί αμερόληπτοι εκτιμητές της ίδιας ποσότητας. Το γεγονός αυτό υποδηλώνει την εισαγωγή του παρακάτω κριτηρίου, ελάχιστης διασποράς για τον περιορισμό της κλάσης των αμερόληπτων εκτιμητών.

Ορισμός 0.8 (Α.Ο.Ε.Δ εκτιμητής)

Ένας εκτιμητής $T(X_1, X_2, \ldots, X_n)$ της $g(\boldsymbol{\theta})$ λέγεται αμερόληπτος ομοιόμορφα ελάχιστης διασποράς (Α.Ο.Ε.Δ) εάν είναι αμερόληπτος και έχει ελάχιστη διασπορά (διακύμανση) μεταξύ όλων των αμερόληπτων εκτιμητών για κάθε $\boldsymbol{\theta} \in \Theta$. Δηλαδή αν T_1 είναι ένας άλλος αμερόληπτος εκτιμητής, τότε $Var(T) \leq Var(T_1)$, $\forall\ \boldsymbol{\theta} \in \Theta$.

Η εύρεση Α.Ο.Ε.Δ εκτιμητών μπορεί να γίνει με δύο τρόπους. Πρώτον, μέσω της ανισότητας Cramer-Rao, στην οποία όμως δεν θα αναφερθούμε εδώ εκτενώς καθώς αποτελεί το αντικείμενο της παρούσας διατριβής, και δεύτερον μέσω του θεωρήματος Lehmann-Scheffe το οποίο και ακολουθεί.

Θεώρημα 0.4 (Lehmann-Scheffe)

Έστω $X_1, X_2, \ldots, X_n$ τ.δ από έναν πληθυσμό με σ.π $p(\mathbf{x},\boldsymbol{\theta})$ ή σ.π.π $f(\mathbf{x},\boldsymbol{\theta})$. Έστω επίσης $\mathbf{T}$ μια πλήρης επαρκής σ.σ και $U = U(\mathbf{T})$ μια αμερόληπτη σ.σ του $g(\boldsymbol{\theta})$ η οποία είναι συνάρτηση του $\mathbf{T}$. Τότε U είναι Α.Ο.Ε.Δ εκτιμητής του $g(\boldsymbol{\theta})$.

Για περισσότερες λεπτομέρειες βλέπε Casella and Berger (2002), Rohatgi and Saleh (2001) και Παπαιωάννου και Φερεντίνος (2000).

Όλα τα παραπάνω αποτελούν βασικές έννοιες της στατιστικής συμπερασματολογίας στην εκτιμητική και βρίσκουν εφαρμογή στα κεφάλαια που ακολουθούν.

Κεφάλαιο 1^{o}

Ανισότητα Cramer-Rao

1.1 Εισαγωγή

Στο κεφάλαιο αυτό θα μελετηθεί η ανισότητα Cramer-Rao στη γενική της μορφή. Είναι πάρα πολύ χρήσιμη στη στατιστική συμπερασματολογία και αποδείχθηκε από τον C.R.Rao το 1944 και ανεξάρτητα από τον Cramer το 1946. (DeGroot, M.H (1987))

Συγκεκριμένα το έναυσμα για να οδηγηθεί ο Rao σε αυτό το αποτέλεσμα ήταν η ερώτηση ενός φοιτητή του κατά τη διάρκεια μίας διάλεξης. Ενώ αποδείκνυε την ασυμπτωτική ανισότητα του Fisher, ότι δηλαδή η ασυμπτωτική διακύμανση δεν είναι μικρότερη από το αντίστροφο του μέτρου του Fisher, κάποιος φοιτητής τον ρώτησε γιατί δεν το αποδείκνυε για πεπερασμένα δείγματα. Επέστρεψε σπίτι του κι αφού εργάστηκε όλη νύχτα το απέδειξε χρησιμοποιώντας την ανισότητα Cauchy-Schwarz και την ιδιότητα της αμεροληψίας.

Το κεφάλαιο αυτό χωρίζεται σε δύο μέρη. Στο πρώτο μέρος θα μελετηθεί η ανισότητα Cramer-Rao στην μονοδιάστατη περίπτωση(συνθήκες κανονικότητας, απόδειξη και πότε γίνεται ισότητα). Στο δεύτερο μέρος μελετάται η πολυδιάστατη περίπτωση ακολουθώντας τα ίδια βήματα.

1.2 Μονοδιάστατη Περίπτωση

Έστω $X_1,\dots,X_n$ τ.δ. από έναν πληθυσμό με σ.π.π. $f(x,\theta)$ με παράμετρο θ μονοδιάστατη, $\theta\in\Theta$ (στην περίπτωση της διακριτής τ.μ. τα συμπεράσματα προκύπτουν ανάλογα). Έστω ο εκτιμητής $T(\mathbf{x})=T(X_1,\dots,X_n)$ για τον οποίο ισχύει $E(T(\mathbf{x}))=g(\theta)$, όπου $g(\theta)$ μια διαφορίσιμη συνάρτηση της θ.

Οι παρακάτω συνθήκες, ονομάζονται συνθήκες κανονικότητας και είναι:

(i) Θ ένα διάστημα πεπερασμένο $\subseteq \mathbb{R}$.

(ii) Η $f(x,\theta)$ είναι θετική σε ένα σύνολο των x, το οποίο είναι ανεξάρτητο του θ.

(iii) $\dfrac{\partial}{\partial\theta}\log f(x,\theta)$ υπάρχει για όλα τα x και όλα τα θ.

(iv) $\dfrac{\partial}{\partial\theta}\int\dots\int\prod_{i=1}^{n} f(x_i,\theta)dx_i\dots\dots dx_n=\int\dots\int\dfrac{\partial}{\partial\theta}\prod_{i=1}^{n} f(x_i,\theta)dx_i\dots\dots dx_n$.

(v) $\dfrac{\partial}{\partial\theta}\int\dots\int \mathrm{T}(x_1,\dots,x_n)\prod_{i=1}^{n} f(x_i,\theta)dx_i\dots\dots dx_n$

$$= \int \ldots \int T(x_1,\ldots,x_n)\frac{\partial}{\partial\theta}\prod_{i=1}^{n} f(x_i,\theta)dx_i\ldots\ldots dx_n.$$

(vi) $\quad 0 < \mathrm{E}\left[\left[\frac{\partial}{\partial\theta}\log f(X,\theta)\right]^2\right] < \infty$ για όλα τα θ στο Θ.

Θεώρημα 1.1 (Ανισότητα Cramer-Rao)

Έστω η τ.μ X με σ.π.π $f(x,\theta)$ και η σ.σ T. Εάν ισχύουν οι παραπάνω συνθήκες κανονικότητας και $E(T(X)) = g(\theta)$, όπου g διαφορίσιμη συνάρτηση, τότε:

$$Var(T(X)) \geq \frac{[g'(\theta)]^2}{E\left[\left(\frac{\partial}{\partial\theta}\log f(X,\theta)\right)^2\right]}. \tag{1.1}$$

Απόδειξη

Η απόδειξη αυτού του θεωρήματος είναι εξαιρετικά απλή καθώς αποτελεί εφαρμογή της ανισότητας Cauchy-Schwarz.

Για δύο τυχαίες μεταβλητές X,Y η ανισότητα Cauchy-Schwarz οδηγεί :

$$[Cov(X,Y)]^2 \leq (Var(X))(Var(Y)). \tag{1.2}$$

Από την (1.2) προκύπτει ένα κάτω φράγμα για την διακύμανση της X, δηλαδή

$$Var(X) \geq \frac{[Cov(X,Y)]^2}{Var(Y)}. \tag{1.3}$$

Για τις τ.μ. $T = T(X)$ και $W = \frac{\partial}{\partial\theta}\log f(X,\theta)$ (score function) η σχέση (1.3) δίνει:

$$VarT(X) \geq \frac{[Cov(T,W)]^2}{VarW}. \tag{1.4}$$

Η συνδιακύμανση των T,W είναι :

$$Cov(T,W) = ETW - (ET)(EW). \tag{1.5}$$

Όμως :

$$E(W)=E\left(\frac{\partial}{\partial\theta}\log f(X,\theta)\right)=\int_A\left(\frac{\partial}{\partial\theta}\log f(x,\theta)\right)f(x,\theta)dx$$ (όπου A το πεδίο ορισμού της τ.μ. X)

$$=\int_A\left(\frac{\partial}{\partial\theta}f(x,\theta)\right)\frac{1}{f(x,\theta)}f(x,\theta)dx$$

$$=\int_A\left(\frac{\partial}{\partial\theta}f(x,\theta)\right)dx$$

$$=\frac{\partial}{\partial\theta}\int_A f(x,\theta)dx \qquad \text{(από τη συνθήκη vi)}$$

$$=\frac{\partial}{\partial\theta}(1)=0. \tag{1.6}$$

Επίσης:

$$E(TW)=\int_A TWdx$$

$$=\int_A T(x)\left(\frac{\partial}{\partial\theta}\log f(x,\theta)\right)f(x,\theta)dx$$

$$=\int_A\left(\frac{\partial}{\partial\theta}f(x,\theta)\right)T(x)dx$$

$$=\int_A\frac{\partial}{\partial\theta}\big(T(x)f(x,\theta)\big)dx$$

$$=\frac{\partial}{\partial\theta}\int_A T(x)f(x,\theta)dx \qquad \text{(από τη συνθήκη iv)}$$

$$=\frac{\partial}{\partial\theta}ET(X). \tag{1.7}$$

Βάσει των (1.6) και (1.7) η (1.5) δίνει:

$$Cov(T,W)=\frac{\partial}{\partial\theta}ET(X). \tag{1.8}$$

Από την (1.6)

$$VarW=EW^2-(EW)^2$$

$$=EW^2$$

$$=E\left(\frac{\partial}{\partial\theta}\log f\big(X,\theta\big)\right)^2. \tag{1.9}$$

Έτσι βάσει των (1.8) και (1.9), η (1.4) γίνεται:

$$VarT(X) \geq \frac{\left(\frac{\partial}{\partial\theta} ET(X)\right)^2}{E\left(\frac{\partial}{\partial\theta}\log f(X,\theta)\right)^2}$$

ή

$$VarT(X) \geq \frac{[g'(\theta)]^2}{\mathrm{I}_X^F(\theta)} \quad ,$$

όπου

$$ET(X) = g(\theta) \qquad \text{(από την υπόθεση)}$$

και

$$I_X^F(\theta) = E\left(\frac{\partial}{\partial\theta}\log f(X,\theta)\right)^2$$

είναι το μέτρο της πληροφορίας του Fisher.

Αν έχουμε τ.δ τότε η ανισότητα Cramer-Rao δίνεται στο παρακάτω πόρισμα.

Πόρισμα 1.1

Έστω $X_1,\ldots,X_n$ τ.δ. από έναν πληθυσμό με σ.π.π. $f(x,\theta)$ και $T(X)$ μια σ.σ. με $ET(X)$ διαφορίσιμη συνάρτηση, τότε:

$$VarT(X) \geq \frac{\left(\frac{d}{d\theta} ET(X)\right)^2}{nI_X^F(\theta)}. \qquad (1.10)$$

Απόδειξη

Η απόδειξη ακολουθεί τα ίδια βήματα με τα προηγούμενα, μόνο που εδώ έχουμε

$$W = \frac{\partial}{\partial\theta}\log f(X_1,\ldots,X_n,\theta)$$

$$= \frac{\partial}{\partial\theta}\log\prod_{i=1}^{n} f(X_i,\theta)$$

$$= \sum_{i=1}^{n}\left(\frac{\partial}{\partial\theta}\log f(X_i,\theta)\right).$$

Άρα, επειδή $X_1,\dots,X_n$ ανεξάρτητες και λόγω της (1.6), έχουμε:

$$VarW=\sum_{i=1}^{n}Var\left(\frac{\partial}{\partial\theta}\log f(X_i,\theta)\right)$$

$$=\sum_{i=1}^{n}\left[E\left(\frac{\partial}{\partial\theta}\log f(X_i,\theta)\right)^2-\left[E\left(\frac{\partial}{\partial\theta}\log f(X_i,\theta)\right)\right]^2\right]$$

$$=nI_X^F(\theta),$$

το οποίο βάσει και των προηγουμένων αποδεικνύει το ζητούμενο.

Η συνθήκη για την επίτευξη της ισότητας για το κάτω φράγμα της ανισότητας Cramer-Rao είναι αρκετά απλή. Η ισότητα προκύπτει έτσι ώστε η ανισότητα Cauchy-Schwarz να γίνεται ισότητα. Βάσει αυτού έχουμε το παρακάτω πόρισμα.

Πόρισμα 1.2

Έστω $X_1,\dots,X_n$ τ.δ. με σ.π.π. $f(x,\theta)$, όπου η $f(x,\theta)$ ικανοποιεί τις συνθήκες του Θεωρήματος 1.1. Έστω επίσης $L(\theta\mid X)=\prod_{i=1}^{n} f(X_i,\theta)$ η συνάρτηση πιθανοφάνειας. Αν $E\big(T(X)\big)=g(\theta)$, τότε το ίσον στην ανισότητα Cramer-Rao επιτυγχάνεται εάν και μόνον εάν

$$\frac{\partial}{\partial\theta}\log L(\theta\mid X)=K(\theta,n)\big[T(X)-g(\theta)\big] \qquad (1.11)$$

με

$$VarT(X)=\left|\frac{g'(\theta)}{K(\theta,n)}\right|$$

για κάποια συνάρτηση $K(\theta,n)$.

Απόδειξη

Στην ανισότητα Cauchy-Schwarz(βλέπε σχέση (1.2)), που στηρίζεται η απόδειξη της ανισότητας Cramer-Rao, το ίσον ισχύει εάν και μόνον εάν η μια είναι γραμμική συνάρτηση της άλλης. Έτσι στην περίπτωσή μας η $W=\frac{\partial}{\partial\theta}\log\prod_{i=1}^{n} f(X_i,\theta)$ αρκεί να είναι ανάλογη της $T(X)-g(\theta)$, δηλαδή να υπάρχει κάποια συνάρτηση $K(\theta,n)$ τέτοια ώστε

$$W=\frac{\partial}{\partial\theta}\log\prod_{i=1}^{n} f(X_i,\theta)X=K(\theta,n)\,[\,T(X)-g(\theta)\,]$$

ή

$$\frac{\partial}{\partial\theta}\log L(\theta\mid X)=K(\theta,n)[T(X)-g(\theta)].$$

Όπως αποδείχτηκε παραπάνω

$$VarW=nI_X^F(\theta). \tag{1.12}$$

Επίσης, εφόσον ισχύει το ίσον στην ανισότητα Cramer-Rao, προκύπτει:

$$VarT=\frac{[g'(\theta)]^2}{nI_X^F(\theta)} \tag{1.13}$$

Από την (1.11) προκύπτει:

$$W=K(\theta,n)\left[T(X)-g(\theta)\right]$$

ή

$$VarW=K^2(\theta,n)VarT(X). \tag{1.14}$$

Έτσι από τις (1.12) και (1.14) έχουμε

$$nI_X^F(\theta)=K^2(\theta,n)VarT(X). \tag{1.15}$$

Από την (1.13)

$$nI_X^F(\theta)=\frac{[g'(\theta)]^2}{VarT}, \tag{1.16}$$

οπότε οι (1.15) και (1.16) δίνουν:

$$K^2(\theta,n)VarT=\frac{[g'(\theta)]^2}{VarT}. \tag{1.17}$$

ή

$$Var^2T=\frac{[g'(\theta)]^2}{K^2(\theta,n)}. \tag{1.18}$$

Άρα

$$VarT=\left|\frac{g'(\theta)}{K(\theta,n)}\right|, \tag{1.19}$$

το οποίο είναι και το ζητούμενο.

Εκτός από την περίπτωση που περιγράφεται στο προηγούμενο πόρισμα, το ίσον επιτυγχάνεται (ισοδύναμα) στην ανισότητα Cramer-Rao και στην περίπτωση που η τ.μ. X ανήκει στην εκθετική οικογένεια κατανομών όπως αποδεικνύεται και στο πόρισμα που ακολουθεί.

Πόρισμα 1.3

Στην ανισότητα Cramer-Rao το ίσον επιτυγχάνεται αν και μόνον αν το X ανήκει στην εκθετική οικογένεια κατανομών, δηλαδή αν

$$f(x,\theta)=c(\theta)h(x)\exp[Q(\theta)T(X)]I_A(x)$$

με Α ανεξάρτητο του θ και όπου $I_A(x)$ η μοναδιαία συνάρτηση.

Απόδειξη

Βάσει της ανισότητας Cauchy-Schwarz για το ίσον στην C-R, όπως και στο Πόρισμα 1.2, έχουμε ότι

$$\frac{\partial}{\partial\theta}\log f(x,\theta)=K(\theta,n)\left[T(x)-g(\theta)\right] \tag{1.20}$$

Ολοκληρώνοντας ως προς θ τη σχέση (1.20) λαμβάνουμε:

$$\log f(x,\theta)+c_1=T(x)\int \mathrm{K}(\theta,n)d\theta-\int \mathrm{K}(\theta,n)g(\theta)d\theta \tag{1.21}$$

όπου c_1 σταθερά.

Έστω

$$\int \mathrm{K}(\theta,n)d\theta=Q(\theta)$$

$$\int \mathrm{K}(\theta,n)g(\theta)d\theta=c^*(\theta)$$

και

$$h^*(x)=-c_1.$$

Αντικαθιστώντας τα παραπάνω στη σχέση (1.21) έχουμε

$$\log f(x,\theta)=T(x)Q(\theta)-c^*(\theta)+h^*(x).$$

Επομένως

$$f(x,\theta)=e^{T(x)Q(\theta)}e^{-c^*(\theta)}e^{h^*(x)}$$

$$=e^{T(x)Q(\theta)}c(\theta)h(x),\, x\in \mathrm{A} \text{ ανεξάρτητο του } \theta \tag{1.22}$$

με

$$c(\theta)=e^{-c^*(\theta)},\, h(x)=e^{h^*(x)}.$$

Άρα η f ανήκει στην εκθετική οικογένεια κατανομών.

Αντίστροφα αν $f(x,\theta)$ ανήκει στην εκθετική οικογένεια κατανομών τότε

$$f(x,\theta) = c(\theta)h(x)\exp\{Q(\theta)T(X)\}$$

ή

$$\log f(x,\theta) = \log c(\theta) + \log h(x) + Q(\theta)T(X).$$

Έτσι

$$\frac{\partial}{\partial\theta}\log f(x,\theta) = \frac{c'(\theta)}{c(\theta)} + Q'(\theta)T(X)$$

$$= Q'(\theta)\left[T(x) - \left(-\frac{c'(\theta)}{c(\theta)Q'(\theta)}\right)\right]$$

$$= K(\theta,n)\left[T(x) - g(\theta)\right], \tag{1.23}$$

όπου $K(\theta,n) = Q'(\theta)$ και $g(\theta) = -\frac{c'(\theta)}{c(\theta)Q'(\theta)} = E\left[T(X)\right] = \mu(\theta)$.

Όσον αφορά το παραπάνω πόρισμα για περισσότερες λεπτομέρειες για μια αυστηρή απόδειξη, βλέπε Wijsman (1973).

1.3 Πολυδιάστατη Περίπτωση

Έστω $\mathbf{x}_1,\ldots,\mathbf{x}_n$ τ.δ. από πληθυσμό με σ.π.π. $f(\mathbf{x},\boldsymbol{\theta})$, όπου $\boldsymbol{\theta}$ είναι το διάνυσμα των παραμέτρων. Ο παραμετρικός χώρος Θ είναι είτε ο Ευκλείδειος r-χώρος R^r ή ένα ορθογώνιο στον R^r. Τα t_i είναι αμερόληπτοι εκτιμητές των g_i και έτσι ικανοποιούν τις σχέσεις :

$$\int t_i(\mathbf{x})f(\mathbf{x},\boldsymbol{\theta})dx = g_i(\theta) \quad , \quad i = 1,\ldots,n \tag{1.24}$$

για τα στατιστικά $t_i(\mathbf{x})$ και τις συναρτήσεις g_i , $i = 1,\ldots,n$ της παραμέτρου $\boldsymbol{\theta}$.

Έστω Σ ο πίνακας διακυμάνσεων-συνδιακυμάνσεων του $\mathbf{t}(\mathbf{x}) = (t_1(\mathbf{x}),...,t_n(\mathbf{x}))$ και $\mathbf{G} = \mathbf{G}(\boldsymbol{\theta})$ ο πίνακας των μερικών παραγώγων $\mathbf{G}(\theta) = \left\| \frac{\partial\, g_i(\boldsymbol{\theta})}{\partial \theta_j} \right\|_{n\,x\,r}$, $\boldsymbol{\theta} \in \Theta$ για $i = 1,...,n$ και $j = 1,...,r$, όπου $\| \; \|_{n\,x\,r}$ δηλώνει έναν nxr πίνακα.

Η συνάρτηση πιθανοφάνειας του δείγματος δίνεται από τη σχέση:

$$L(\boldsymbol{\theta}|\mathbf{x}) = \prod_{i=1}^{n} f(\mathbf{x}_i, \boldsymbol{\theta}) \tag{1.25}$$

και ο λογάριθμος της είναι:

$$l(\mathbf{x}, \boldsymbol{\theta}) = \log L(\boldsymbol{\theta}|\mathbf{x}) = \sum_{i=1}^{n} \log f(\mathbf{x}_i, \boldsymbol{\theta}). \tag{1.26}$$

Ορίζουμε τη συνάρτηση $\mathbf{s} = s(\mathbf{x}, \boldsymbol{\theta})$ (score function) ως εξής:

$$s(\mathbf{x}, \boldsymbol{\theta}) = \frac{\partial}{\partial \boldsymbol{\theta}} l(\mathbf{x}, \boldsymbol{\theta}) = \frac{1}{L(\boldsymbol{\theta}|\mathbf{x})} \frac{\partial}{\partial \boldsymbol{\theta}} L(\boldsymbol{\theta}|\mathbf{x}) = \frac{\partial}{\partial \boldsymbol{\theta}} \log L(\boldsymbol{\theta}|\mathbf{x}) \tag{1.27}$$

Η $s(\mathbf{x}, \boldsymbol{\theta})$ είναι ένα τυχαίο διάνυσμα το οποίο θα αναφέρεται απλά ως $\mathbf{s}$. Ο πίνακας διακυμάνσεων-συνδιακυμάνσεων του $\mathbf{s}$ καλείται πίνακας πληροφορίας του Fisher είναι συμμετρικός και θετικά ορισμένος και θα συμβολίζεται με $I_{\mathbf{X}}^F = I_{\mathbf{X}}^F(\boldsymbol{\theta})$. Δηλαδή ο πίνακας πληροφορίας του Fisher για την παραπάνω οικογένεια δίνεται ως εξής:

$$I_{\mathbf{X}}^F(\boldsymbol{\theta}) = \left\| \int f(\mathbf{x}, \boldsymbol{\theta}) \left[\frac{\partial}{\partial \theta_i} \log f(\mathbf{x}, \boldsymbol{\theta}) \frac{\partial}{\partial \theta_j} \log f(\mathbf{x}, \boldsymbol{\theta}) \right] d\mathbf{x} \right\|_{r\,x\,r}. \tag{1.28}$$

Όταν το $\boldsymbol{\theta}$ είναι r -διάστατο, τότε οι συνθήκες κανονικότητας του Θεωρήματος 1.1 προσαρμόζονται ως εξής :

i) Το σύνολο $A = \{\mathbf{x} : f(\mathbf{x}, \boldsymbol{\theta}) > 0\}$ δεν εξαρτάται από το $\boldsymbol{\theta}$.

ii)Οι μερικές παράγωγοι $\frac{\partial}{\partial \theta_i} f(\mathbf{x}, \boldsymbol{\theta})$ υπάρχουν για όλα τα $\mathbf{x} \in A$, $\boldsymbol{\theta} \in \Theta$ για $i = 1,...,r$ και είναι συνεχείς στο $\boldsymbol{\theta} \in \Theta$.

iii) Αν $\mathbf{t}$ είναι ένα οποιοδήποτε στατιστικό τότε:

$$\frac{\partial}{\partial \theta_i} \int \mathbf{t}(\mathbf{x}) f(\mathbf{x}, \boldsymbol{\theta}) d\mathbf{x} = \int \mathbf{t}(\mathbf{x}) \frac{\partial}{\partial \theta_i} f(\mathbf{x}, \boldsymbol{\theta}) d\mathbf{x}, \quad \theta \in \Theta \;, i = 1,...,r\,. \tag{1.29}$$

iv)Ο $I_{\mathbf{X}}^F(\boldsymbol{\theta})$ είναι θετικά ορισμένος.

Για να προχωρήσουμε στην διατύπωση και απόδειξη της ανισότητας Cramer-Rao για την πολυδιάστατη περίπτωση, κρίνεται σκόπιμο να διατυπωθούν προηγουμένως κάποια Λήμματα και Πορίσματα(βλέπε Mardia et al(1979)). Έτσι έχουμε τα παρακάτω.

Πόρισμα 1.4

Αν $\mathbf{s}=s(\mathbf{x},\boldsymbol{\theta})$ όπως ορίστηκε παραπάνω και $\mathbf{t}$ μια οποιαδήποτε συνάρτηση των $\mathbf{x}$ και $\boldsymbol{\theta}$, τότε, κάτω από τις συνθήκες κανονικότητας που αναφέρθηκαν παραπάνω, ισχύει:

$$E(\mathbf{s}\mathbf{t}')=\frac{\partial}{\partial\boldsymbol{\theta}}E(\mathbf{t}')-E\left(\frac{\partial\,\mathbf{t}'}{\partial\boldsymbol{\theta}}\right). \qquad (1.30)$$

Απόδειξη

Ισχύει ότι:

$$E(\mathbf{t}')=\int\mathbf{t}'L\,d\mathbf{x}. \qquad (1.31)$$

Παραγωγίζοντας και τα δύο μέλη της σχέσης (1.31) ως προς $\boldsymbol{\theta}$ και με την εναλλαγή της παραγώγου και του ολοκληρώματος (εξασφαλίζεται από τις συνθήκες κανονικότητας) προκύπτει:

$$\frac{\partial}{\partial\boldsymbol{\theta}}E(\mathbf{t}')=\int\mathbf{t}'\frac{\partial L}{\partial\boldsymbol{\theta}}d\mathbf{x}+\int L\frac{\partial\mathbf{t}'}{\partial\boldsymbol{\theta}}d\mathbf{x} \qquad (1.32)$$

$$=\int\frac{\partial\log L}{\partial\boldsymbol{\theta}}\mathbf{t}'L\,d\mathbf{x}+\int\frac{\partial\mathbf{t}'}{\partial\boldsymbol{\theta}}L\,d\mathbf{x} \qquad (1.33)$$

$$=\int\mathbf{s}\mathbf{t}'L\,d\mathbf{x}+\int\frac{\partial\mathbf{t}'}{\partial\boldsymbol{\theta}}L\,d\mathbf{x} \qquad (1.34)$$

$$=E(\mathbf{s}\mathbf{t}')+E\left(\frac{\partial\mathbf{t}'}{\partial\boldsymbol{\theta}}\right), \qquad (1.35)$$

από την οποία και το ζητούμενο.

Πόρισμα 1.5

Εάν $s(\mathbf{x},\boldsymbol{\theta})$ η συνάρτηση (score function), που αντιστοιχεί σε μια συνάρτηση πιθανοφάνειας για την οποία ισχύουν οι συνθήκες κανονικότητας, τότε $E(\mathbf{s})=0$.

Απόδειξη

$$E(\mathbf{s}) = \int \mathbf{s} L d\mathbf{x}$$

$$= \int \left(\frac{1}{L} \frac{\partial}{\partial \boldsymbol{\theta}} L \right) L d\mathbf{x} \qquad \text{(από σχέση (1.27))}$$

$$= \int \frac{\partial}{\partial \boldsymbol{\theta}} L d\mathbf{x}$$

$$= \frac{\partial}{\partial \boldsymbol{\theta}} \int L d\mathbf{x} \qquad \text{(συνθήκες κανονικότητας)}$$

$$= 0 .$$

Πόρισμα 1.6

Αν $\mathbf{s} = s(\mathbf{x}, \boldsymbol{\theta})$ η συνάρτηση (score function) που αντιστοιχεί σε μια συνάρτηση πιθανοφάνειας για την οποία ισχύουν οι συνθήκες κανονικότητας και $\mathbf{t}$ είναι ένας οποιοσδήποτε αμερόληπτος εκτιμητής του $\boldsymbol{\theta}$ τότε:

$$E(\mathbf{s}\mathbf{t}') = I, \qquad (1.36)$$

όπου I ο μοναδιαίος πίνακας.

Απόδειξη

Εφόσον το $\mathbf{t}$ σ.σ δεν εξαρτάται από το $\boldsymbol{\theta}$, τότε $\dfrac{\partial\, \mathbf{t}'}{\partial \boldsymbol{\theta}} = 0$.

Από το πόρισμα 1.4 ισχύει:

$$E(\mathbf{s}\mathbf{t}') = \frac{\partial}{\partial \boldsymbol{\theta}} E(\mathbf{t}') - E\left(\frac{\partial \mathbf{t}'}{\partial \boldsymbol{\theta}}\right) \qquad (1.37)$$

και λόγω αμεροληψίας $E(\mathbf{t}) = \boldsymbol{\theta}$.

Αντικαθιστώντας στη σχέση (1.37) προκύπτει

$$E(\mathbf{s}\mathbf{t}') = \frac{\partial}{\partial \boldsymbol{\theta}} E(\mathbf{t}')$$

$$= \left\| \frac{\partial}{\partial \theta_j} (\theta_1, \ldots\ldots, \theta_r) \right\|_{rxr}$$

$$
=\begin{pmatrix} 1 & \cdots\cdots & 0 \\ 0 & \ddots & 0 \\ \vdots & & \vdots \\ 0 & \cdots\cdots & 1 \end{pmatrix}
$$

$$
=I\,.
$$

Επίσης για να προχωρήσουμε στην γενίκευση του θεωρήματος Cramer-Rao για την πολυδιάστατη περίπτωση χρειάζεται ένα αποτέλεσμα από την θεωρία πινάκων το οποίο δίνεται στο παρακάτω Λήμμα. Για περισσότερες λεπτομέρειες βλέπε Mardia, Kent and Bibby(1979).

Λήμμα 1.1

Έστω $\boldsymbol{\alpha}$, $\mathbf{x}$ διανύσματα $n x 1$ και B συμμετρικός και θετικά ορισμένος πίνακας $n x n$ τότε:

$$
\max_{\mathbf{x}}\{(\boldsymbol{\alpha}'\mathbf{x})^2/(\mathbf{x}'B\mathbf{x})\}=\boldsymbol{\alpha}'B^{-1}\boldsymbol{\alpha}\quad. \tag{1.38}
$$

Θεώρημα 1.4 (Γενίκευση Θεωρήματος Cramer-Rao – Πολυδιάστατη Περίπτωση)

Αν $\mathbf{t}=t(\mathbf{x})$ είναι ένας αμερόληπτος εκτιμητής του $\boldsymbol{\theta}$ που βασίζεται σε μια συνάρτηση πιθανοφάνειας για την οποία ισχύουν οι συνθήκες κανονικότητας τότε :

$$
\Sigma \geq \left(\mathrm{I}_{\mathrm{X}}^{F}\right)^{-1}, \tag{1.39}
$$

όπου I_{X}^{F} είναι ο πίνακας πληροφορίας του Fisher και $\boldsymbol{\Sigma}$ είναι ο πίνακας διακυμάνσεων – συνδιακυμάνσεων του $\mathbf{t}$, όπως ορίστηκαν προηγουμένως.

Απόδειξη

Πριν προχωρήσουμε στην απόδειξη χρειάζεται να διατυπωθούν κάποιες σχέσεις που κρίνονται απαραίτητες για την κατανόηση της απόδειξης. Έστω $\mathbf{a},\mathbf{x},\mathbf{y}$ οποιαδήποτε διανύσματα και A, B οποιοιδήποτε πίνακες κατάλληλα ορισμένοι.

Ο πίνακας διακυμάνσεων – συνδιακυμάνσεων του τυχαίου διανύσματος $\mathbf{x}$ συμβολίζεται με $V(\mathbf{x})$ και ισχύει:

$$V(\mathbf{a}'\mathbf{x}) = \mathbf{a}'V(\mathbf{x})\mathbf{a}, \text{ όπου } \boldsymbol{\alpha} \text{ σταθερό διάνυσμα.} \tag{1.40}$$

Για τη συνδιακύμανση των τυχαίων διανυσμάτων $\mathbf{x},\mathbf{y}$ ισχύουν οι σχέσεις:

$$Cov[\mathbf{x},\mathbf{y}] = C[\mathbf{x},\mathbf{y}] = E[\mathbf{x}\mathbf{y}'] - E(\mathbf{x})E(\mathbf{y}) \tag{1.41}$$

$$Cov[A\mathbf{x},B\mathbf{y}] = C[A\mathbf{x},B\mathbf{y}] = AC[\mathbf{x},\mathbf{y}]B' \tag{1.42}$$

$$C(\mathbf{x},\mathbf{y}) = C(\mathbf{y},\mathbf{x})' \tag{1.43}$$

Η συνδιακύμανση θα συμβολίζεται με C ενώ η διακύμανση με V. Για τον συντελεστή συσχέτισης ισχύει

$$r^2[x,y] = C^2[x,y]/V(x)V(y) \tag{1.44}$$

και βάσει των προηγουμένων έχουμε

$$V(\mathbf{t}) = \Sigma, V(\mathbf{s}) = I_{\mathbf{x}}^F.$$

Επίσης στη συνέχεια για συμμετρικούς πίνακες A,B της ίδιας τάξης, όταν γράφουμε $A \geq B$ θα εννοούμε ότι ο πίνακας $A-B$ είναι θετικά ημιορισμένος.

Έστω $\beta = \boldsymbol{\alpha}'\mathbf{t}$, $\gamma = \mathbf{c}'\mathbf{s}$ όπου $\boldsymbol{\alpha}$, $\mathbf{c}$ κατάλληλα διανύσματα έτσι ώστε να ορίζονται τα β, γ και $\mathbf{t}$, $\mathbf{s}$ όπως ορίζονται παραπάνω.

Για την συνδιακύμανση των β, γ ισχύει:

$$C[\beta,\gamma] = C[\boldsymbol{\alpha}'\mathbf{t},\mathbf{c}'\mathbf{s}] = \boldsymbol{\alpha}'C(\mathbf{t},\mathbf{s})(\mathbf{c}')' = \boldsymbol{\alpha}'C(\mathbf{t},\mathbf{s})\mathbf{c} \tag{1.45}$$

Για την διακύμανση των β, γ ισχύει:

$$V(\beta) = \boldsymbol{\alpha}'V(\mathbf{t})\boldsymbol{\alpha} \tag{1.46}$$

$$V(\gamma) = V(\mathbf{c}'\mathbf{s}) = \mathbf{c}'V(\mathbf{s})\mathbf{c}. \tag{1.47}$$

Εφόσον $V(\mathbf{t}) = \Sigma$ και $V(\mathbf{s}) = \mathrm{I}_{\mathbf{x}}^F$ οι σχέσεις (1.46) και (1.47) δίνουν αντίστοιχα

$$V(\beta) = \boldsymbol{\alpha}'\Sigma\boldsymbol{\alpha}. \tag{1.48}$$

και

$$V(\gamma) = \mathbf{c}'\mathrm{I}_{\mathbf{x}}^F\mathbf{c}. \tag{1.49}$$

Έτσι έχουμε:

$$C(\mathbf{t},\mathbf{s}) = C(\mathbf{s},\mathbf{t})' \quad \text{(από σχέση (1.43))}$$

$$= \left[E(\mathbf{s}\mathbf{t}') - E(\mathbf{s})E(\mathbf{t})\right]' \text{(από Πόρισμα 1.5 } E(\mathbf{s}) = 0\text{)}$$

$$= \left[E(\mathbf{s}\mathbf{t}')\right]' \quad \text{(από πόρισμα 1.6 } E(\mathbf{s}\mathbf{t}') = I\text{)}$$

$$= I. \tag{1.50}$$

όπου I ο μοναδιαίος πίνακας.

Άρα η σχέση (1.45) γίνεται

$$C(\beta,\gamma) = \boldsymbol{\alpha}'\mathrm{I}\mathbf{c} = \boldsymbol{\alpha}'\mathbf{c}. \tag{1.51}$$

Χρησιμοποιώντας τα παραπάνω αποτελέσματα(σχέσεις (1.51), (1.48), (1.49)) για τον συντελεστή συσχέτισης των β, γ έχουμε:

$$r^2[\beta,\gamma] = \frac{C^2(\beta,\gamma)}{V(\beta)V(\gamma)}$$

$$= (\boldsymbol{\alpha}'\mathbf{c})^2 / \{\boldsymbol{\alpha}'\Sigma\boldsymbol{\alpha}\mathbf{c}'I^F\mathbf{c}\} \leq 1. \tag{1.52}$$

Το μέγιστο της (1.52) ως προς $\mathbf{c}$, χρησιμοποιώντας το Λήμμα 1.1 εφόσον τα $\boldsymbol{\alpha}$, $\mathbf{c}$ είναι κατάλληλα ορισμένα και ο πίνακας I_X^F είναι συμμετρικός και θετικά ορισμένος, έχει ως εξής :

$$\max_{\mathbf{c}}\left\{(\boldsymbol{\alpha}'\mathbf{c})^2 / \left\{\boldsymbol{\alpha}'\Sigma\boldsymbol{\alpha}\mathbf{c}'I_X^F\mathbf{c}\right\}\right\} = (1/\boldsymbol{\alpha}'\Sigma\boldsymbol{\alpha})\max_{\mathbf{c}}\left\{(\boldsymbol{\alpha}'\mathbf{c})^2 / \left\{\mathbf{c}'I_X^F\mathbf{c}\right\}\right\}$$

$$= \boldsymbol{\alpha}'\left(I_{\mathbf{x}}^F\right)^{-1}\boldsymbol{\alpha}/\boldsymbol{\alpha}'\Sigma\boldsymbol{\alpha} \leq 1 \quad \text{για κάθε } \boldsymbol{\alpha}. \tag{1.53}$$

Έτσι από την προηγούμενη σχέση

$$\boldsymbol{\alpha}'\Sigma\boldsymbol{\alpha} \geq \boldsymbol{\alpha}'\left(I_{\mathbf{x}}^F\right)^{-1}\boldsymbol{\alpha}$$

$$\boldsymbol{\alpha}'\{\Sigma - \left(I_{\mathbf{x}}^F\right)^{-1}\}\boldsymbol{\alpha} \geq 0 \text{ για κάθε } \boldsymbol{\alpha} \tag{1.54}$$

Η (1.54) δίνει το ζητούμενο.

Αν τώρα δε θέλουμε να εκτιμήσουμε το $\boldsymbol{\theta}$ αλλά μια συνάρτηση $\mathbf{g}(\boldsymbol{\theta})$ αυτού, τότε το παραπάνω θεώρημα διαμορφώνεται ως εξής.

Θεώρημα 1.5

Έστω $\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)'$ τυχαίο διάνυσμα από την κατανομή $f(\mathbf{x}, \boldsymbol{\theta})$ και ισχύουν οι συνθήκες κανονικότητας. Αν $\mathbf{t} = \mathbf{t}(\mathbf{x})$ είναι ένας αμερόληπτος εκτιμητής της $\mathbf{g}(\boldsymbol{\theta})$, τότε ισχύει:

$$\boldsymbol{\Sigma} \geq \mathbf{G}(I_{\mathbf{x}}^{F})^{-1}\mathbf{G}', \tag{1.55}$$

όπου $\mathbf{G} = \left\| \dfrac{\partial\, g_i(\boldsymbol{\theta})}{\partial \theta_j} \right\|_{n \times r}$ και g_i και θ_j είναι οι συνιστώσες των $\mathbf{g}$ και $\boldsymbol{\theta}$ αντίστοιχα.

Απόδειξη

Έστω $\beta = \boldsymbol{\alpha}'\mathbf{t}$, $\gamma = \mathbf{c}'\mathbf{s}$ και $\mathbf{s}, \mathbf{t}$ ομοίως με το προηγούμενο θεώρημα.

Ισχύουν τα παρακάτω, αντίστοιχα με την απόδειξη του προηγούμενου θεωρήματος:

$$C[\beta, \gamma] = \boldsymbol{\alpha}' C(\mathbf{t}, \mathbf{s})\mathbf{c} \tag{1.56}$$

$$V(\beta) = \boldsymbol{\alpha}'\boldsymbol{\Sigma}\boldsymbol{\alpha} \tag{1.57}$$

$$V(\gamma) = \mathbf{c}'\mathbf{I}_{\mathbf{x}}^{F}\mathbf{c} \tag{1.58}$$

Κατ' αναλογία με τα Πορίσματα (1.5) και (1.6) προκύπτει :

$$C(\mathbf{t}, \mathbf{s}) = C(\mathbf{s}, \mathbf{t})' \quad \text{(από σχέση (1.43)}$$

$$= [E(\mathbf{s}\mathbf{t}') - E(\mathbf{s})E(\mathbf{t})]' \ \text{(από Πόρισμα 1.5 } E(\mathbf{s}) = 0)$$

$$= [E(\mathbf{s}\mathbf{t}')]'$$

$$= \left[\frac{\partial}{\partial \boldsymbol{\theta}} E(\mathbf{t}') - E(\frac{\partial\, \mathbf{t}'}{\partial \boldsymbol{\theta}}) \right]' \ , \ (E(\frac{\partial\, \mathbf{t}'}{\partial \boldsymbol{\theta}}) = 0)$$

$$= \left[\left(\frac{\partial}{\partial \boldsymbol{\theta}} \mathbf{g}(\boldsymbol{\theta}) \right)' \right]'$$

$$= (\mathbf{G}')'$$

$$= \mathbf{G} \ . \tag{1.59}$$

Άρα η σχέση (1.56) έχει ως εξής:

$$C[\beta,\gamma] = \boldsymbol{\alpha}'\mathbf{G}\mathbf{c} \tag{1.60}$$

Χρησιμοποιώντας τώρα τα παραπάνω αποτελέσματα έχουμε:

$$r^2[\beta,\gamma] = \frac{C^2(\beta,\gamma)}{V(\beta)V(\gamma)}$$

$$= (\boldsymbol{\alpha}'\mathbf{G}\mathbf{c})^2 / \{\boldsymbol{\alpha}'\boldsymbol{\Sigma}\boldsymbol{\alpha}\mathbf{c}'I_{\mathbf{x}}^F\mathbf{c}\}. \tag{1.61}$$

Το μέγιστο της (1.61) ως προς $\mathbf{c}$,χρησιμοποιώντας το Λήμμα 1.1 όπως προηγουμένως και θεωρώντας ως διάνυσμα το $\boldsymbol{\alpha}'\mathbf{G}$ είναι:

$$\max_{\mathbf{c}}\left\{\left((\boldsymbol{\alpha}'\mathbf{G})\mathbf{c}\right)^2 / \left\{\boldsymbol{\alpha}'\boldsymbol{\Sigma}\boldsymbol{\alpha}\,\mathbf{c}'I_{\mathbf{x}}^F\mathbf{c}\right\}\right\} = \left(1/\boldsymbol{\alpha}'\boldsymbol{\Sigma}\boldsymbol{\alpha}\right)\max_{\mathbf{c}}\left\{\left((\boldsymbol{\alpha}'\mathbf{G})\mathbf{c}\right)^2 / \left\{\mathbf{c}'I_{\mathbf{x}}^F\mathbf{c}\right\}\right\}$$

$$= \boldsymbol{\alpha}'\mathbf{G}\left(I_{\mathbf{x}}^F\right)^{-1}\left(\boldsymbol{\alpha}'\mathbf{G}\right)' / \boldsymbol{\alpha}'\boldsymbol{\Sigma}\boldsymbol{\alpha}$$

$$= \boldsymbol{\alpha}'\mathbf{G}\left(I_{\mathbf{x}}^F\right)^{-1}\mathbf{G}'\mathbf{a}/\boldsymbol{\alpha}'\boldsymbol{\Sigma}\boldsymbol{\alpha} \le 1 \tag{1.62}$$

για κάθε $\boldsymbol{\alpha}$.

Έτσι από την προηγούμενη σχέση

$$\boldsymbol{\alpha}'\{\boldsymbol{\Sigma} - \mathbf{G}\left(I_{\mathbf{x}}^F\right)^{-1}\mathbf{G}'\}\boldsymbol{\alpha} \ge 0 \text{ για όλα τα } \boldsymbol{\alpha}, \tag{1.63}$$

το οποίο δίνει το ζητούμενο όπως και προηγουμένως.

Το ίσον στην ανισότητα Cramer – Rao επιτυγχάνεται στην πολυδιάστατη περίπτωση ανάλογα με την μονοδιάστατη περίπτωση, όπως φαίνεται και στα Πορίσματα που ακολουθούν.

Πόρισμα 1.6

Η ανισότητα Cramer-Rao γίνεται ισότητα ένα και μόνο εάν

$$\frac{\partial}{\partial\theta_i}\log f\left(\mathbf{x},\theta_1,\ldots,\theta_r\right) = \sum_{j=1}^{n} u_{ij}\left[\mathrm{t}_j(\mathbf{x}) - g_j(\boldsymbol{\theta})\right] \tag{1.64}$$

όπου οι u_{ij} εξαρτώνται από το $\boldsymbol{\theta}$ αλλά όχι από το $\mathbf{x}$ και $ET_j(\mathbf{x}) = g_j(\boldsymbol{\theta})$, $i,j = 1,\ldots,n$.

Απόδειξη

Από το Θεώρημα 1.5 έχουμε ότι η ανισότητα Cramer-Rao δίνεται ως εξής

$$\boldsymbol{\Sigma} \ge \mathbf{G}(I_{\mathbf{x}}^F)^{-1}\mathbf{G}'. \tag{1.65}$$

Προκύπτει έτσι λοιπόν ένα κάτω φράγμα για τον πίνακα διακυμάνσεων- συνδιακυμάνσεων Σ το οποίο επιτυγχάνεται αν

$$\det\left\{\boldsymbol{\Sigma}-\mathbf{G}(I_{\mathbf{x}}^{F})^{-1}\mathbf{G}'\right\}=0. \tag{1.66}$$

(βλέπε Zacks (1971))

Αυτό όμως ισχύει αν ο συντελεστής συσχέτισης μεταξύ των $\mathbf{s}(\mathbf{x},\boldsymbol{\theta})$ και $\mathbf{t}(\mathbf{x})$ ισούται με τη μονάδα.

Επομένως $\det\left\{\boldsymbol{\Sigma}-\mathbf{G}(I_{\mathbf{x}}^{F})^{-1}\mathbf{G}'\right\}=0$ αν και μόνον αν $Cov^2(\mathbf{s},\mathbf{t})=Var(\mathbf{s})Var(\mathbf{t})$ γεγονός που ισχύει αν και μόνον αν υπάρχει ένας πίνακας με στοιχεία σταθερές, ο

$$U(\boldsymbol{\theta})=\left\|u_{ij}(\boldsymbol{\theta}),i,j=1,....,n\right\| \tag{1.67}$$

ανεξάρτητος του $\mathbf{X}$ και τέτοιος ώστε

$$\mathbf{s}(\mathbf{x},\boldsymbol{\theta})=U(\boldsymbol{\theta})(\mathbf{t}(\mathbf{x})-\mathbf{g}(\boldsymbol{\theta})). \tag{1.68}$$

Δηλαδή

$$\begin{pmatrix}\frac{\partial}{\partial\theta_1}\log f(\mathbf{x},\theta_1,...,\theta_n)\\ \vdots\\ \vdots\\ \frac{\partial}{\partial\theta_n}\log f(\mathbf{x},\theta_1,...,\theta_n)\end{pmatrix}_{nx1}=\begin{pmatrix}u_{11} & \cdots\cdots & u_{1n}\\ \vdots & & \vdots\\ \vdots & \ddots & \vdots\\ \vdots & & \vdots\\ u_{n1} & \cdots\cdots & u_{n1}\end{pmatrix}_{nxn}\begin{pmatrix}t_1(\mathbf{x})-g_1(\boldsymbol{\theta})\\ \vdots\\ \vdots\\ \vdots\\ t_n(\mathbf{x})-g_n(\boldsymbol{\theta})\end{pmatrix}_{nx1},\tag{1.69}$$

από όπου προκύπτει η ζητούμενη σχέση

$$\frac{\partial}{\partial\theta_i}\log f(\mathbf{x},\theta_1,...,\theta_r)=\sum_{j=1}^{n}u_{ij}\left[\mathrm{t}_j(\mathbf{x})-g_j(\boldsymbol{\theta})\right]. \tag{1.70}$$

<u>Πόρισμα 1.7</u>

Αν ισχύουν οι συνθήκες κανονικότητας, το ίσον επιτυγχάνεται στην ανισότητα Cramer-Rao αν και μόνον αν

$$f(\mathbf{x},\boldsymbol{\theta})=h(\mathbf{x})c(\boldsymbol{\theta})e^{\mathbf{Q}'(\boldsymbol{\theta})t(\mathbf{x})}I_A(\mathbf{x}), \tag{1.71}$$

με A ανεξάρτητο του $\boldsymbol{\theta}$, $c(\boldsymbol{\theta})>0$, $h(\mathbf{x})>0$ και κάθε στοιχείο του διανύσματος $\mathbf{Q}(\boldsymbol{\theta})$ έχει συνεχείς πρώτες μερικές παραγώγους.

Απόδειξη

Από την ανισότητα Cramer-Rao έχουμε

$$\boldsymbol{\Sigma} \geq \mathbf{G}\left(I_{\mathbf{x}}^{F}\right)^{-1}\mathbf{G}' \ , \ \boldsymbol{\theta} \in \Theta \tag{1.72}$$

από την οποία

$$I_{\mathbf{x}}^{F} \geq \mathbf{G}'\boldsymbol{\Sigma}^{-1}\mathbf{G} \, , \ \boldsymbol{\theta} \in \Theta \tag{1.73}$$

(βλέπε Mardia et al (1979))

Το ίσον στις σχέσεις (1.72) ή (1.73) επιτυγχάνεται αν και μόνον αν το διάνυσμα της συνάρτησης $\mathbf{s}$ (score function)

$$\mathbf{s}' = \left(\frac{\partial}{\partial \theta_1} \log f, \ldots, \frac{\partial}{\partial \theta_r} \log f \right) \tag{1.74}$$

είναι γραμμική συνάρτηση του $\mathbf{G}'\boldsymbol{\Sigma}^{-1}\mathbf{t}$ (Zacks (1971)). Έτσι:

$$\mathbf{s} = \kappa(\boldsymbol{\theta})\mathbf{G}'\boldsymbol{\Sigma}^{-1}\mathbf{t} + W(\boldsymbol{\theta}) \tag{1.75}$$

ή

$$\frac{\partial}{\partial \theta_s} \log f = \kappa(\boldsymbol{\theta}) \sum_{i,j=1}^{r} \frac{\partial g_i(\boldsymbol{\theta})}{\partial \theta_s} \sigma^{ij} t_j(\mathbf{x}) + W_s(\boldsymbol{\theta}) \ , \ s = 1,\ldots,r, \tag{1.76}$$

για όλα τα $\mathbf{x}$, όπου $\kappa(\boldsymbol{\theta})$ και $W_s(\boldsymbol{\theta})$ είναι πραγματικές συναρτήσεις του $\boldsymbol{\theta}$ και σ^{ij} είναι το (i,j) στοιχείο του Σ^{-1}. Ολοκληρώνοντας τα δυο μέλη της σχέσης (1.77) ως προς θ_s, από ένα γνωστό θ_s^0 έως το θ_s , για όλα τα $s = 1,\ldots,r$ και προσθέτοντας τις εξισώσεις που προκύπτουν για όλα τα $\theta, \theta_{10}\theta_{20},\ldots,\theta_{r0}$ στο Θ, με

$$\theta_{s0}' = (\theta_1^0,\ldots,\theta_s^0\theta_{s+1},\ldots,\theta_r) \, , \ \theta_c^0 \text{ γνωστό } c = 1,\ldots,s \text{ και } s = 1,\ldots,r,$$

προκύπτει ότι:

$$\log f(\mathrm{x},\theta) - \log f(\mathrm{x},\theta_{r0}) = \sum_{i,j=1}^{n} \left[q_{ij}(\boldsymbol{\theta}) - q_{ij}(\boldsymbol{\theta}_{r0}) \right] t_j(\mathrm{x}) + b(\theta) - b(\theta_{r0}), \tag{1.77}$$

για κάθε $\boldsymbol{\theta}$ και θ_{r0} γνωστά στο Θ . Όπως και στη μονοδιάστατη περίπτωση η προηγούμενη σχέση οδηγεί στο ότι η $f(\mathbf{x},\boldsymbol{\theta})$ ανήκει στην εκθετική οικογένεια κατανομών με $q_{ij}(\boldsymbol{\theta})$ και $b(\boldsymbol{\theta})$ πραγματικές συναρτήσεις του θ τέτοιες ώστε:

$$\frac{\partial}{\partial \theta_s} q_{ij}(\boldsymbol{\theta}) = \kappa(\boldsymbol{\theta}) \frac{\partial g_i(\boldsymbol{\theta})}{\partial \theta_s} \sigma^{ij}$$

και

$$\frac{\partial}{\partial \theta_s} b(\boldsymbol{\theta}) = -\sum_{i,j=1}^{r} \frac{\partial q_{ij}(\boldsymbol{\theta})}{\partial \theta_s} g_j(\boldsymbol{\theta})$$

για $\boldsymbol{\theta} \in \Theta$, $s = 1,\ldots,r$ και $i, j = 1,\ldots,n$.

Αντιστρόφως, αν η σ.π.π. f δίνεται από την σχέση (1.71) τότε ικανοποιείται η σχέση (1.76) και επομένως η s (score function) είναι γραμμική συνάρτηση του $\mathbf{G}'\boldsymbol{\Sigma}^{-1}\mathbf{t}$ το οποίο συνεπάγεται την ισότητα στη σχέση (1.72) ή (1.73)(βλέπε Zografos and Ferentinos (1994)).

Κεφάλαιο 2^{ο}

Εφαρμογές της Ανισότητας Cramer-Rao στην Εκτιμητική

2.1 Εισαγωγή

Όπως έχει ήδη αναφερθεί η ανισότητα Cramer-Rao έχει πάρα πολλές εφαρμογές στη στατιστική συμπερασματολογία. Στο κεφάλαιο αυτό θα μελετηθούν αυτές που αφορούν την εκτιμητική σε σημείο. Θα δοθούν παραδείγματα που αφορούν την εκτίμηση σε σημείο, μέσω της ανισότητας Cramer-Rao, παραμετρικών συναρτήσεων, τόσο στη μονοδιάστατη όσο και στην πολυδιάστατη περίπτωση.

2.2 Εκτίμηση σε σημείο (Μονοδιάστατη περίπτωση)

Έστω $X_1,\ldots,X_n$ τ.δ. από έναν πληθυσμό με σ.π.π. $f(x,\theta)$ (ή σ.π $p(x,\theta)$). Η εύρεση ενός αμερόληπτα ομοιόμορφα ελάχιστης διασποράς (Α.Ο.Ε.Δ) εκτιμητή μιας συνάρτησης του θ, έστω $g(\theta)$, είναι ένα κλασικό πρόβλημα της στατιστικής. Η ανισότητα Cramer-Rao και η πληρότητα σε συνδυασμό με την επάρκεια (βλέπε θεώρημα Lehmann-Scheffe) δίνουν τις δυο κύριες μεθόδους για την εύρεση ενός Α.Ο.Ε.Δ εκτιμητή.

Η ανισότητα Cramer-Rao δεν εγγυάται την ύπαρξη ενός Α.Ο.Ε.Δ εκτιμητή αλλά δίνει έναν τρόπο για την αναζήτηση του. Η αναζήτηση αυτή γίνεται ως εξής: πρώτα βρίσκεται το κάτω φράγμα για τις διακυμάνσεις όλων των αμερόληπτων εκτιμητών για τους οποίους ισχύουν οι συνθήκες κανονικότητας. Στη συνέχεια το ζητούμενο είναι ένας αμερόληπτος εκτιμητής, του οποίου η διακύμανση να συμπίπτει με το κάτω φράγμα που προσδιορίστηκε προηγουμένως. Εάν αυτό επιτευχθεί τότε έχει βρεθεί ο Α.Ο.Ε.Δ εκτιμητής.Το προηγούμενο (η εύρεση αμερόληπτου εκτιμητή) δεν είναι εύκολο. Τα παρακάτω όμως διευκολύνουν περισσότερο την άμεση εύρεση Α.Ο.Ε.Δ εκτιμητή.

Στο 1^{ο} κεφάλαιο αποδείχθηκε ότι το ίσον στην ανισότητα Cramer-Rao επιτυγχάνεται σε δύο ισοδύναμες περιπτώσεις.

A) Εάν και μόνον εάν ισχύει η σχέση

$$\frac{\partial}{\partial\theta}\log L(\theta\,|\,X) = K(\theta,n)[U(X)-g(\theta)], \tag{2.1}$$

για κάποια συνάρτηση $K(\theta,n)$. Επομένως εάν υπάρχει ένας εκτιμητής U με $EU = g(\theta)$ έτσι ώστε να ισχύει αυτή η σχέση τότε ο U είναι Α.Ο.Ε.Δ της $g(\theta)$ με

$$Var\left[U(X)\right] = \left|\frac{g'(\theta)}{K(\theta,n)}\right|. \tag{2.2}$$

B) Εάν και μόνον εάν το X ανήκει στην εκθετική οικογένεια κατανομών, δηλαδή εάν:

$$f(x,\theta) = c(\theta)h(x)\exp[Q(\theta)T(X)]I_A(x) \tag{2.3}$$

με A ανεξάρτητο του θ.

Στην περίπτωση αυτή είναι εύκολο να αποδειχθεί ότι ο

$$T^* = \frac{1}{n}\sum_{i=1}^{n} T(X_i) \tag{2.4}$$

είναι ο Α.Ο.Ε.Δ της

$$\mu(\theta) = \mathrm{E}\left[\mathrm{T}(\mathrm{X})\right] = \frac{-c'(\theta)}{c(\theta)Q'(\theta)} \tag{2.5}$$

με

$$VarT^* = \left|\frac{\mu'(\theta)}{nQ'(\theta)}\right| = \frac{\left[\mu'(\theta)\right]^2}{n\mathrm{I}_x^F(\theta)}, \tag{2.6}$$

όπου $c'(\theta) = \frac{\partial}{\partial\theta}c(\theta)$, $Q'(\theta) = \frac{\partial}{\partial\theta}Q(\theta)$, $\mu'(\theta) = \frac{\partial}{\partial\theta}\mu(\theta)$.

(Για περισσότερες λεπτομέρειες βλέπε Ferentinos (1984))

Από την πρώτη περίπτωση παρατηρούμε ότι η αναζήτηση Α.Ο.Ε.Δ εκτιμητών μέσω της ανισότητας Cramer-Rao περιορίζεται μόνο στις συναρτήσεις $g(\theta) = \mu(\theta) = E\left[T(X)\right]$ ή γραμμικές συναρτήσεις αυτών, όπου $T(X)$ η σ.σ που υπεισέρχεται στη σχέση (2.3).

Η δεύτερη περίπτωση μας δείχνει ότι η αναζήτηση Α.Ο.Ε.Δ εκτιμητών μέσω της ανισότητας Cramer-Rao περιορίζεται μόνο σε κατανομές που ανήκουν στην εκθετική οικογένεια. Όπως και προηγουμένως η μόνη παραμετρική συνάρτηση για την οποία μπορεί να βρεθεί Α.Ο.Ε.Δ εκτιμητής μέσω της Cramer-Rao είναι της μορφής $\mu(\theta)$ ή $\lambda_1 + \lambda_2\mu(\theta)$ και ο Α.Ο.Ε.Δ της είναι της μορφής $\lambda_1 + \lambda_2 T^*$, όπου λ_1, λ_2 είναι σταθερές ανεξάρτητες της παραμέτρου θ και των παρατηρήσεων $X_1,\ldots,X_n$.

Επομένως αν η οικογένεια δεν είναι εκθετική τότε δεν μπορεί να βρεθεί Α.Ο.Ε.Δ εκτιμητής μέσω της ανισότητας Cramer-Rao. Αν είναι εκθετική, τότε υπολογίζεται η $g(\theta)=\mu(\theta)$ και το T^* δίνεται από τη σχέση (2.4) και είναι ο Α.Ο.Ε.Δ εκτιμητής της $g(\theta)$. Αν θέλουμε να εκτιμήσουμε γραμμική συνάρτηση της $\mu(\theta)$ δηλαδή $\lambda_1+\lambda_2\mu(\theta)$, τότε ο Α.Ο.Ε.Δ εκτιμητής της είναι ο $\lambda_1+\lambda_2 T^*$.

Στην ουσία οι δύο περιπτώσεις είναι ισοδύναμες, η διαφορετική τους όμως έκφραση μας βοηθάει από τεχνική άποψη στην εύρεση της συναρτησιακής έκφρασης του Α.Ο.Ε.Δ εκτιμητή. Δηλαδή με την περίπτωση (Α) βρίσκουμε άμεσα την συναρτησιακή έκφραση του Α.Ο.Ε.Δ, ενώ από τη (Β) έχουμε ότι για εύρεση Α.Ο.Ε.Δ μέσω της ανισότητας C-R μόνο για εκθετική οικογένεια έχει έννοια να ψάχνουμε και μάλιστα εκτιμάμε τις παραμετρικές συναρτήσεις $\mu(\theta)=E\left[T(X)\right]$ ή γραμμικές συναρτήσεις αυτών.

Τα παραδείγματα που ακολουθούν μας δείχνουν την άμεση εφαρμογή των προηγουμένων που διευκολύνουν την εύρεση Α.Ο.Ε.Δ εκτιμητών μέσω της ανισότητας Cramer-Rao.

<u>Παράδειγμα 2.1 (Bernoulli)</u>

Έστω $X_1,\ldots,X_n$ τ.δ. από έναν πληθυσμό με κατανομή την $\mathrm{B}(1,p)$. Ζητούμενο είναι ένας Α.Ο.Ε.Δ εκτιμητής για την παράμετρο p. Ισχύουν οι συνθήκες κανονικότητας και η $\mathrm{B}(1,p)$ ανήκει στην εκθετική οικογένεια κατανομών με $T(X)=X$ και $ET(X)=p$. Έτσι ο Α.Ο.Ε.Δ της p βρίσκεται ως εξής:

Έχουμε ότι

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

και

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

Έτσι

$$\frac{\partial}{\partial p}\log L=\frac{1}{p}\sum_{i=1}^{n}x_i-\frac{\left(n-\sum_{i=1}^{n}x_i\right)}{(1-p)}$$

$$=\frac{n}{p(1-p)}\left(\frac{1}{n}\sum_{i=1}^{n}x_i-p\right).$$

Άρα ο $T^*=\overline{X}$ είναι ο Α.Ο.Ε.Δ της p με $VarT^*=\frac{p(1-p)}{n}$.

Παράδειγμα 2.2 (Γεωμετρική κατανομή)

Έστω $X_1,\ldots,X_n$ τ.δ. από έναν πληθυσμό με κατανομή :

$$p(x,\theta)=\theta(1-\theta)^x I_{(0,1,2,\ldots)}(x)\quad,\qquad 0<\theta<1$$

Ζητούμενο είναι ο Α.Ο.Ε.Δ εκτιμητής της $g(\theta)=\frac{(1-\theta)}{\theta}$.

Ισχύουν οι συνθήκες κανονικότητας και η $p(x,\theta)$ ανήκει στην εκθετική οικογένεια με $c(\theta)=\theta$, $T(x)=x$, $Q(\theta)=\log(1-\theta)$.

Έτσι, βάσει των προηγουμένων ο Α.Ο.Ε.Δ της

$$\begin{aligned}ET(X)&=E(X)\\&=\frac{-c'(\theta)}{c(\theta)Q'(\theta)}\\&=\frac{1-\theta}{\theta}\\&=g(\theta).\end{aligned}\tag{2.7}$$

είναι ο

$$\begin{aligned}T^*&=\frac{1}{n}\sum_{i=1}^{n}T(X_i)\\&=\frac{1}{n}\sum_{i=1}^{n}X_i\\&=\overline{X}.\end{aligned}\tag{2.8}$$

Παράδειγμα 2.3 (Εκθετική Κατανομή)

Έστω $X_1,\dots,X_n$ τ.δ. από έναν πληθυσμό με κατανομή την Εκθ(θ). Ζητούμενο είναι ο Α.Ο.Ε.Δ. εκτιμητής της $g(\theta)=\frac{1}{\theta}$. Ισχύουν οι συνθήκες κανονικότητας η X ανήκει στην εκθετική οικογένεια με $T(X)=X$ και $ET(X)=\frac{1}{\theta}$.

Η σ.π.π της X είναι η

$$f(x,\theta)=\frac{1}{\theta}e^{-\frac{1}{\theta}x},\ 0\le x<\infty,\ \theta>0 \tag{2.9}$$

Έτσι βάσει των προηγουμένων έχουμε:

$$\begin{aligned} L(\theta/X)&=\prod_{i=1}^{n} f(x_i,\theta)\\ &=\frac{1}{\theta^n}e^{-\frac{1}{\theta}\sum_{i=1}^{n}X_i}. \end{aligned} \tag{2.10}$$

$$\log L(\theta/X)=-n\log\theta-\frac{1}{\theta}\sum_{i=1}^{n}X_i \tag{2.11}$$

και

$$\begin{aligned} \frac{\partial}{\partial\theta}\log L(\theta/X)&=-\frac{n}{\theta}+\frac{1}{\theta^2}\sum_{i=1}^{n}X_i\\ &=\frac{n}{\theta^2}\left(\frac{1}{n}\sum_{i=1}^{n}X_i-\frac{1}{\theta}\right). \end{aligned} \tag{2.12}$$

Άρα ο $\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$ είναι ο Α.Ο.Ε.Δ της $g(\theta)=\frac{1}{\theta}$.

Παράδειγμα 2.4 (Κανονική Κατανομή)

Έστω $X_1,\dots,X_n$ τ.δ. από έναν πληθυσμό με κατανομή $N(\mu,\sigma^2)$, όπου $\mu=\mu_0$ γνωστό. Ζητούμενο είναι ο Α.Ο.Ε.Δ εκτιμητής της σ^2.
Έχουμε:

$$f(x,\sigma^2)=\frac{1}{(2\pi\sigma^2)^{1/2}}e^{-\frac{1}{2\sigma^2}(X_i-\mu)^2}$$

$$L(\mu,\sigma^2\mid x)=\frac{1}{(2\pi\sigma^2)^{n/2}}e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(X_i-\mu)^2}$$

$$\log L(\mu,\sigma^2 \mid x) = -\frac{n}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(X_i-\mu)^2$$

$$\frac{\partial \log L(\mu,\sigma^2 \mid x)}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{n}{2\sigma^4}\sum_{i=1}^{n}(X_i-\mu)^2$$

$$\frac{\partial \log L(\mu,\sigma^2 \mid x)}{\partial \sigma^2} = \frac{n}{2\sigma^4}\left(\sum_{i=1}^{n}\frac{(X_i-\mu)^2}{n} - \sigma^2\right).$$

Χρησιμοποιώντας τη σχέση (2.1) και θεωρώντας

$$\mathrm{K}(\sigma^2,n) = \frac{n}{2\sigma^4},$$

$$U = \sum_{i=1}^{n}\frac{(X_i-\mu)^2}{n} \quad \text{και} \quad g(\sigma^2)=\sigma^2 \tag{2.13}$$

προκύπτει ότι ο Α.Ο.Ε.Δ για την σ^2 είναι ο $\sum_{i=1}^{n}\frac{(X_i-\mu)^2}{n}$, ο οποίος υπολογίζεται μόνο αν το μ είναι γνωστό. Αν το μ δεν είναι γνωστό, βλέπε παράδειγμα 2.9.

Παράδειγμα 2.5 (Pareto Κατανομή)

Έστω $X_1,\ldots,X_n$ τ.δ. από έναν πληθυσμό με σ.π.π :

$$f(x,\theta) = \theta\beta^{\theta}x^{-(\theta+1)} \quad , x \ge \beta \ , \theta > 0$$

όπου β γνωστό.

Η σ.π.π. $f(x,\theta)$ μπορεί να γραφεί ως εξής:

$$f(x,\theta) = \theta\beta^{\theta}e^{-\theta\log x}e^{-\log x}$$

Επομένως ανήκει στην εκθετική οικογένεια κατανομών με

$$c(\theta)=\theta\beta^{\theta},\ Q(\theta)=-(\theta+1),\ T(X)=\log X,\ h(x)=e^{-\log x}$$

Χρησιμοποιώντας τις σχέσεις (2.4) , (2.5) και (2.6) προκύπτει ότι ο $T^* = \frac{1}{n}\sum_{i=1}^{n}\log X_i$ είναι ο Α.Ο.Ε.Δ. εκτιμητής της

$$\mu(\theta) = \frac{-c'(\theta)}{c(\theta)Q'(\theta)} = \frac{-\beta(1+\theta\log\beta)}{\theta\beta^{\theta}(-1)} = \frac{1}{\theta} + \log\beta, \tag{2.14}$$

με διακύμανση

$$VarT^* = \left|\frac{\mu'(\theta)}{nQ'(\theta)}\right| = \left|\frac{-\frac{1}{\theta^2}}{n}\right| = \frac{1}{n\theta^2}. \tag{2.15}$$

Θεωρώντας ως $g(\theta) = \lambda_1 + \lambda_2\mu(\theta)$ με $\lambda_1 = -\log\beta$ και $\lambda_2 = 1$ προκύπτει:

$$g(\theta) = -\log(\beta) + \frac{1}{\theta} + \log\beta = \frac{1}{\theta}. \tag{2.16}$$

Χρησιμοποιώντας τα παραπάνω συμπεράσματα ο Α.Ο.Ε.Δ. της $g(\theta)$ είναι ο

$$\begin{aligned}\lambda_1 + \lambda_2 T^* &= -\log\beta + \frac{1}{n}\sum_{i=1}^{n}\log X_i \\ &= \frac{1}{n}\sum_{i=1}^{n}\left[\log X_i - \log\beta\right]. \\ &= \frac{1}{n}\sum_{i=1}^{n}\log\frac{X_i}{\beta} \\ &= \frac{1}{n}\left[\log\left(\prod_{i=1}^{n}\frac{X_i}{\beta}\right)\right]\end{aligned}$$

με διακύμανση $\frac{1}{n\theta^2}$.

<u>Παράδειγμα 2.6 (Poisson Κατανομή)</u>

Έστω $X_1,\ldots,X_n$ τ.δ. από έναν πληθυσμό με κατανομή την Poisson(θ).

Εδώ $c(\theta) = e^{-\theta}, Q(\theta) = \log\theta, T(X) = X$. Χρησιμοποιώντας όπως και προηγουμένως τις σχέσεις (2.4), (2.5) και (2.6) καταλήγουμε στο ότι ο Α.Ο.Ε.Δ εκτιμητής της

$$\mu(\theta) = \frac{-c'(\theta)}{c(\theta)Q'(\theta)} = \theta,$$

είναι ο $T^* = \frac{1}{n}\sum_{i=1}^{n}X_i = \bar{X}$ με $VarT^* = \left|\frac{\mu'(\theta)}{nQ'(\theta)}\right| = \frac{\theta}{n}$

Αν $g(\theta) = \theta^2$ τότε δεν μπορεί να βρεθεί Α.Ο.Ε.Δ εκτιμητής της $g(\theta) = \theta^2$ μέσω της ανισότητας Cramer-Rao γιατί η $g(\theta)$ δεν είναι γραμμική συνάρτηση της $\mu(\theta) = \theta$. Ωστόσο με τη θεωρία της επάρκειας και της πληρότητας (βλέπε θεώρημα Lehmann-Scheffe) προκύπτει ότι ο

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

είναι ο Α.Ο.Ε.Δ εκτιμητής της θ^2.

Παρατήρηση 2.1

Γενικά, το κάτω φράγμα Cramer-Rao δεν είναι πάντα δυνατό να επιτευχθεί καθώς συχνά υπάρχει ένα κάτω φράγμα για την διακύμανση το οποίο είναι μεγαλύτερο από το κάτω φράγμα Cramer-Rao. Ο λόγος είναι ότι ενώ οι συνθήκες κανονικότητας ισχύουν οι παραμετρικές συναρτήσεις που θέλουμε να εκτιμήσουμε δεν είναι οι $\mu(\theta) = ET(X)$ ή γραμμικές συναρτήσεις αυτής.

Για παράδειγμα έστω ότι έχουμε μια τ.μ X από έναν πληθυσμό με κατανομή την Poisson(θ) και θέλουμε να βρούμε έναν Α.Ο.Ε.Δ εκτιμητή για την $g(\theta) = e^{-\theta}$. Εδώ $ET(X) = E(X) = \theta$. Άρα $g(\theta)$ δεν είναι γραμμική συνάρτηση της $\mu(\theta)$ οπότε το κάτω φράγμα Cramer-Rao δεν είναι δυνατόν να επιτευχθεί και έτσι δεν μπορεί να βρεθεί Α.Ο.Ε.Δ εκτιμητής της $e^{-\theta}$ μέσω της ανισότητας Cramer-Rao. Πράγματι έχουμε

$$p(x,\theta) = \frac{e^{-\theta}\theta^x}{x!}, x = 0,1,2,\ldots$$

$$\log p(x,\theta) = -\theta + x\log\theta - \log x!$$

$$\frac{\partial}{\partial\theta}\log p(x,\theta) = -1 + \frac{x}{\theta}$$

$$E\left[\left(\frac{\partial}{\partial\theta}\log p(x,\theta)\right)^2\right] = E\left[\left(\frac{X}{\theta} - 1\right)^2\right] = \frac{1}{\theta^2}E\left[(X-\theta)^2\right] = \frac{1}{\theta^2}Var(X) = \frac{1}{\theta}$$

Άρα

$$I_x^F(\theta) = \frac{1}{\theta}.$$

Επίσης

$$g(\theta) = e^{-\theta}, g'(\theta) = -e^{-\theta}$$

και επομένως

$$\text{Κ.Φ.}C-R=\frac{\left[g'(\theta)\right]^2}{I_x^F(\theta)}=\theta e^{-\theta}. \tag{2.17}$$

Έστω
$$T=\begin{cases}1, X=1\\ 0, \alpha\lambda\lambda o\acute{\upsilon}\end{cases}$$

τότε

$$E(T)=1P[X=0]+0P[X\neq 0]=e^{-\theta}.$$

Άρα T είναι αμερόληπτος εκτιμητής της $e^{-\theta}$.

Η διακύμανση του T έχει ως εξής:

$$\begin{aligned}VarT &= E(T^2)-[E(T)]^2\\ &= e^{-\theta}-\left(e^{-\theta}\right)^2\\ &= e^{-\theta}\left(1-e^{-\theta}\right).\end{aligned}$$

Άρα $VarT = e^{-\theta}\left(1-e^{-\theta}\right)$. (2.18)

Συγκρίνοντας τη διακύμανση του T με το Κ.Φ.C-R από τις σχέσεις (2.17) και (2.18) έχουμε ότι

$$VarT \geq \text{Κ.Φ.}C-R,$$

καθώς

$$e^{-\theta}(1-e^{-\theta})\geq \theta e^{-2\theta}.$$

Αυτό ήταν αναμενόμενο διότι $g(\theta)=e^{-\theta}\neq \mu(\theta)=\theta$ και ούτε είναι γραμμική συνάρτηση της $\mu(\theta)$. Εντούτοις με άλλη μεθοδολογία (βλέπε θεώρημα Lehmann-Scheffe) υπάρχει ο Α.Ο.Ε.Δ της $e^{-\theta}$ και είναι ο $\left(\frac{n-1}{n}\right)^{\sum X_i}$ για τ.δ $X_1,\ldots,X_n$.

Επομένως ένας Α.Ο.Ε.Δ εκτιμητής για κάποια παραμετρική συνάρτηση μπορεί να υπάρχει χωρίς να συμπίπτει η διακυμανσή του με το Κ.Φ.C-R και όταν ακόμα ισχύουν οι συνθήκες κανονικότητας.

Παράδειγμα 2.7

Έστω $X_1,\ldots,X_n$ ανεξάρτητες τυχαίες μεταβλητές με Poisson κατανομή και μέσες τιμές $\exp\{\theta z_1\},\ldots,\exp\{\theta z_n\}$,αντίστοιχα, όπου $z_1,\ldots,z_n$ είναι γνωστοί πραγματικοί αριθμοί. Μέσω της διαδικασίας που περιγράφεται παραπάνω θα βρεθεί ένας Α.Ο.Ε.Δ εκτιμητής για μια συνάρτηση $g(\theta)$. Εδώ η παράμετρος της Poisson είναι $e^{\theta z}$.

Έχουμε

$$X_i \sim P\left(\exp\{\theta z_i\}\right),\ i=1,\ldots,n$$

$$p(x_i,\theta)=\frac{e^{\theta z_i x_i}e^{-\theta z_i}}{x_i!}$$

$$L(\theta)=\prod_{i=1}^{n}p(x_i,\theta)=\frac{e^{\theta\sum_{i=1}^{n}z_i x_i}\,e^{-\sum_{i=1}^{n}e^{\theta z_i}}}{\prod_{i=1}^{n}x_i!}$$

$$\log L(\theta)=\theta\sum_{i=1}^{n}z_i x_i-\sum_{i=1}^{n}e^{\theta z_i}-\log\left(\prod_{i=1}^{n}x_i!\right)$$

$$\frac{\partial}{\partial\theta}\log L(\theta)=\left(\sum_{i=1}^{n}z_i x_i-\sum_{i=1}^{n}z_i e^{\theta z_i}\right).$$

Άρα

$$\frac{\partial}{\partial\theta}\log L(\theta)=K(\theta,n)[U(X_i)-g(\theta)]$$

με $$K(\theta,n)=1,U(X_i)=\sum_{i=1}^{n}z_i x_i,g(\theta)=\sum_{i=1}^{n}z_i e^{\theta z_i}.$$

Επομένως βάσει των παραπάνω ο Α.Ο.Ε.Δ εκτιμητής για την

$$g(\theta)=\sum_{i=1}^{n}z_i e^{\theta z_i}$$

είναι ο $U(X_i)=\sum_{i=1}^{n}z_i x_i$ με διακύμανση $VarU=\left|\frac{g'(\theta)}{K(\theta,n)}\right|=\sum_{i=1}^{n}z_i^2 e^{\theta z_i}$.

(βλέπε Ferguson(1996)).

2.3 Εκτίμηση σε σημείο (Πολυδιάστατη Περίπτωση)

Ενδιαφέρον παρουσιάζει η εύρεση ενός Α.Ο.Ε.Δ εκτιμητή στην περίπτωση που η παράμετρος $\boldsymbol{\theta}$ είναι πολυδιάστατη. Όπως και στη μονοδιάστατη περίπτωση θα δοθούν παραδείγματα στα οποία σκιαγραφείται η εφαρμογή της ανισότητας Cramer-Rao για την εύρεση Α.Ο.Ε.Δ εκτιμητών όταν η παράμετρος $\boldsymbol{\theta}$ είναι πολυδιάστατη.

Αν $\boldsymbol{\theta}$ είναι r -διάστατη στο 1ο Κεφάλαιο αποδείχθηκε ότι η ανισότητα Cramer-Rao γίνεται ισότητα σε δύο ισοδύναμες περιπτώσεις:

Α) Εάν και μόνο εάν

$$\frac{\partial}{\partial\theta_i}\log f\left(\mathbf{x},\theta_1,\ldots,\theta_r\right)=\sum_{j=1}^{r}u_{ij}\left[T_j\left(\mathbf{x}\right)-g_j\left(\boldsymbol{\theta}\right)\right] \qquad (2.19)$$

όπου οι u_{ij} εξαρτώνται από το $\boldsymbol{\theta}$ αλλά όχι από το $\mathbf{x}$ και $ET_j\left(\mathbf{x}\right)=g_j\left(\boldsymbol{\theta}\right)$ $i,j=1,\ldots,r$.

Β) Εάν και μόνο εάν η $f\left(\mathbf{x},\boldsymbol{\theta}\right)$ είναι η r -παραμετρική εκθετική οικογένεια, δηλαδή

$$f\left(\mathbf{x},\boldsymbol{\theta}\right)=c\left(\boldsymbol{\theta}\right)e^{\sum_{i=1}^{r}Q_j(\boldsymbol{\theta})T_j(x)}h\left(x\right)I_A\left(x\right) \qquad (2.20)$$

με $\mathbf{x}\in A$ ανεξάρτητο του $\boldsymbol{\theta}$ και $\boldsymbol{\theta}'=\left(\theta_1,\ldots,\theta_r\right)$.

(βλέπε Zacks(1971), Zografos and Ferentinos (1994)).

Συνδυάζοντας τα δύο παραπάνω αποτελέσματα (όπως και στη μονοδιάστατη περίπτωση) μπορεί να βρεθεί η κλάση των παραμετρικών συναρτήσεων και οι αντίστοιχοι Α.Ο.Ε.Δ εκτιμητές μέσω της ανισότητας Cramer-Rao. Θα διευκρινιστούν τα παραπάνω για $\boldsymbol{\theta}$ διδιάστατο.

Έστω $r=2$ και X τ.μ με σ.π.π η οποία δίνεται από τη σχέση (2.20) και $X_1,\ldots,X_n$ τ.δ από αυτόν τον πληθυσμό. Ένας εκτιμητής $T'=\left(T_1,T_2\right)$ του $\boldsymbol{\theta}'=\left(\theta_1,\theta_2\right)$ είναι Α.Ο.Ε.Δ μέσω της ανισότητας Cramer-Rao αν

$$ET_1=\theta_1, ET_2=\theta_2, Var\left(T_1\right)=I_{11}^{-1}\left(\boldsymbol{\theta}\right)=\frac{I_{22}\left(\boldsymbol{\theta}\right)}{D\left(\boldsymbol{\theta}\right)},\; Var\left(T_2\right)=I_{22}^{-1}\left(\boldsymbol{\theta}\right)=\frac{I_{11}\left(\boldsymbol{\theta}\right)}{D\left(\boldsymbol{\theta}\right)},$$

όπου $I_{ii}\left(\boldsymbol{\theta}\right)=E\left(\frac{\partial}{\partial\theta_i}\log f\left(x,\boldsymbol{\theta}\right)\right)^2, D\left(\boldsymbol{\theta}\right)=Det\left(I_x^F\left(\boldsymbol{\theta}\right)\right)$ και $I_{ii}^{-1}\left(\boldsymbol{\theta}\right)$ είναι το (i,i) στοιχείο του πίνακα $\left(I_x^F\left(\boldsymbol{\theta}\right)\right)^{-1}$ για $i=1,\ldots,r$.

Οι σχέσεις (2.19) βάσει της σχέσεως (2.20) γίνονται:

$$\frac{\partial}{\partial\theta_1}\log\prod_{i=1}^{n} f(x_i,\boldsymbol{\theta}) = nQ'_{1\theta_1}\left[\frac{1}{n}\sum_i T_1(x_i) - g_i(\boldsymbol{\theta})\right] + nQ'_{2\theta_1}\left[\frac{1}{n}\sum_i T_2(x_i) - g_2(\boldsymbol{\theta})\right] \quad (2.21)$$

$$\frac{\partial}{\partial\theta_2}\log\prod_{i=1}^{n} f(x,\boldsymbol{\theta}) = nQ'_{1\theta_2}\left[\frac{1}{n}\sum_i T_1(x_i) - g_1(\boldsymbol{\theta})\right] + nQ'_{2\theta_2}\left[\frac{1}{n}\sum_i T_2(x_i) - g_2(\boldsymbol{\theta})\right] \quad (2.22)$$

όπου

$$(ET_1, ET_2)' = (g_1(\boldsymbol{\theta}), g_2(\boldsymbol{\theta}))' = [Q(\boldsymbol{\theta})]^{-1} c^*(\boldsymbol{\theta})$$

με

$$Q(\boldsymbol{\theta}) = \left\|Q'_{i\theta_j}\right\|_{2x2}, \ (c^*(\boldsymbol{\theta}))' = (c_1', c_2')$$

$$Q'_{i\theta_j} = \frac{\partial}{\partial\theta_j}Q_i(\boldsymbol{\theta}), \ c_i' = \frac{\partial}{\partial\theta_i}\log c(\boldsymbol{\theta}), \quad i,j = 1,2.$$

Έτσι ο εκτιμητής $T' = (T_1^*, T_2^*)$ με $T_1^* = \frac{1}{n}\sum_i T_1(x_i)$, $T_2^* = \frac{1}{n}\sum_i T_2(x_i)$, είναι Α.Ο.Ε.Δ για την $g'(\boldsymbol{\theta}) = (g_1(\boldsymbol{\theta}), g_2(\boldsymbol{\theta}))$. Οι μόνοι άλλοι Α.Ο.Ε.Δ εκτιμητές μέσω της ανισότητας Cramer-Rao για άλλες παραμετρικές συναρτήσεις είναι γραμμικοί μετασχηματισμοί της $g(\theta)$. Με άλλα λόγια η μόνη παραμετρική συνάρτησή που είναι εκτίμησιμη μέσω της ανισότητας Cramer-Rao είναι της μορφής

$$(g'(\boldsymbol{\theta}))^* = [k_1 + k_2 g_1(\boldsymbol{\theta}), k_3 + k_4 g_2(\boldsymbol{\theta})] \quad (2.23)$$

και ο Α.Ο.Ε.Δ εκτιμητής της είναι της μορφής

$$(T')^* = [k_1 + k_2 T_1^*(\boldsymbol{\theta}), k_3 + k_4 T_2^*(\boldsymbol{\theta})] \quad (2.24)$$

όπου k_i, $i = 1,2,3,4$ είναι σταθερές ανεξάρτητες της παραμέτρου $\boldsymbol{\theta}$ και των παρατηρήσεων $X_1, \ldots, X_n$.

Επομένως αν η οικογένεια δεν είναι εκθετική δεν μπορεί να βρεθεί Α.Ο.Ε.Δ εκτιμητής μέσω της ανισότητας Cramer-Rao. Αν είναι εκθετική τότε υπολογίζονται τα $g'(\boldsymbol{\theta}) = (g_1(\boldsymbol{\theta}), g_2(\boldsymbol{\theta}))$, $g^*(\boldsymbol{\theta})$ και $T^*(\boldsymbol{\theta})$ όπως δίνονται παραπάνω.

Αν $r > 2$ τότε εκφράσεις για τα $E(T_j)$ δίνονται από τον Jonson et al (1979). Τα παραπάνω αποτελέσματα διευκρινίζονται-εφαρμόζονται στα παρακάτω παραδείγματα.

Παράδειγμα 2.8 (Τριωνυμική Κατανομή)

Έστω το τ.δ (X_1, X_2, X_3) με τριωνυμική κατανομή με παραμέτρους $\theta_1, \theta_2, \theta_3$. Η από κοινού σ.π.π δίνεται από τη σχέση

$$P(x_1, x_2, x_3; \theta_1, \theta_2, \theta_3) = \frac{n}{x_1! x_2! x_3!} \theta_1^{x_1} \theta_2^{x_2} (1-\theta_1-\theta_2)^{x_3}$$

με $\theta_1 + \theta_2 + \theta_3 = 1$ και $x_1 + x_2 + x_3 = n$.

Εδώ

$$c(\boldsymbol{\theta}) = (1-\theta_1-\theta_2)^n,\ Q_1(\boldsymbol{\theta}) = \frac{1}{1-\theta_1-\theta_2},\ Q_2(\boldsymbol{\theta}) = \frac{\theta_2}{1-\theta_1-\theta_2}$$

$$T_1(X) = X_1,\ T_2(X) = X_2,\ h(x) = \frac{n!}{x_1! x_2! x_3!} I_A(x)$$

Επομένως από τις σχέσεις (2.21) και (2.22) προκύπτει

$$ET_1 = n\theta_1,\ ET_2 = n\theta_2,$$

$$\frac{\partial}{\partial\theta_1} \log P = \frac{n(1-\theta_2)}{\theta_1(1-\theta_1-\theta_2)}\left(\frac{X_1}{n} - \theta_1\right) + \frac{n}{(1-\theta_1-\theta_2)}\left(\frac{X_2}{n} - \theta_2\right),$$

$$\frac{\partial}{\partial\theta_2} \log P = \frac{n}{(1-\theta_1-\theta_2)}\left(\frac{X_1}{n} - \theta_1\right) + \frac{n(1-\theta_1)}{\theta_2(1-\theta_1-\theta_2)}\left(\frac{X_2}{n} - \theta_2\right),$$

$$\left[I_x^F(\boldsymbol{\theta})\right]^{-1} = \frac{1}{n}\begin{pmatrix} \theta_1(1-\theta_1) & -\theta_1\theta_2 \\ -\theta_1\theta_2 & \theta_2(1-\theta_2) \end{pmatrix}.$$

Άρα $T' = (T_1^*, T_2^*) = \left(\frac{X_1}{n}, \frac{X_2}{n}\right)$ είναι Α.Ο.Ε.Δ εκτιμητής για την $\theta' = (\theta_1, \theta_2)$ με

$$VarT_1^* = I_{11}^{-1}(\boldsymbol{\theta}) = \frac{\theta_1(1-\theta_1)}{n}, VarT_2^* = I_{22}^{-1}(\boldsymbol{\theta}) = \frac{\theta_2(1-\theta_2)}{n}.$$

(βλέπε Zografos and Ferentinos (1994))

<u>Παράδειγμα 2.9 (Κανονική Κατανομή, $\theta_1=\mu$, $\theta_2=\sigma^2$)</u>

Εδώ $f\left(X_i,\theta_1,\theta_2\right)=\left(\dfrac{1}{2\pi\theta_2}\right)^{\frac{1}{2}} e^{-\frac{1}{2\theta_2}\left(X_i-\theta_1\right)^2}$

$$\prod_{i=1}^{n} f\left(X_i,\theta_1,\theta_2\right)=\left(\frac{1}{2\pi\theta_2}\right)^{\frac{n}{2}} e^{-\frac{1}{2\theta_2}\sum_{i=1}^{n}\left(X_i-\theta_1\right)^2}$$

$$=\left(\frac{1}{2\pi\theta_2}\right)^{\frac{n}{2}} e^{-\frac{1}{2\theta_2}\sum_{i=1}^{n}\left(X_i^2-2X_i\theta_1+\theta_1^2\right)}$$

$$=\left(\frac{1}{2\pi\theta_2}\right)^{\frac{n}{2}} e^{-\frac{n\theta_1^2}{2\theta_2}} e^{-\frac{1}{2\theta_2}\sum_{i=1}^{n}X_i^2} e^{\frac{\theta_1}{\theta_2}X_i}$$

με $\quad c(\boldsymbol{\theta})=\left(\dfrac{1}{2\pi\theta_2}\right)^{\frac{n}{2}} e^{-\frac{n\theta_1^2}{2\theta_2}}$, $Q_1(\boldsymbol{\theta})=\dfrac{\theta_1}{\theta_2}$, $Q_2(\boldsymbol{\theta})=-\dfrac{1}{2\theta_2^2}$, $T_1(X)=X$, $T_2(X)=X^2$, $h(x)=1$

και

$ET_1=\theta_1=\mu$, $ET_2=\theta_1^2+\theta_2=\mu^2+\sigma^2$.

Επομένως από τις σχέσεις (2.21) και (2.22) προκύπτει:

$$\frac{\partial}{\partial\theta_1}\log\prod_{i=1}^{n} f\left(X_i,\theta_1,\theta_2\right)=\frac{n}{\theta_2}\left[\frac{1}{n}\sum_i X_i-\theta_1\right]$$

$$\frac{\partial}{\partial\theta_2}\log\prod_{i=1}^{n} f\left(X_i,\theta_1,\theta_2\right)=-\frac{n\theta_1}{\theta_2^2}\left[\frac{1}{n}\sum_i X_i-\theta_1\right]+\frac{n}{2\theta_2^2}\left[\frac{1}{n}\sum_i X_i^2-\left(\theta_1^2+\theta_2\right)\right].$$

Ο πίνακας πληροφορίας του Fisher έχει ως εξής:

$$I_x^F(\boldsymbol{\theta})=\begin{bmatrix}\dfrac{n}{\theta_2} & 0\\ 0 & \dfrac{n}{2\theta_2^2}\end{bmatrix}, \left[I_x^F(\boldsymbol{\theta})\right]^{-1}=\begin{bmatrix}\dfrac{\theta_2}{n} & 0\\ 0 & \dfrac{2\theta_2^2}{n}\end{bmatrix}.$$

Άρα

$$T' = \left(T_1^*, T_2^*\right) = \left(\frac{1}{n}\sum_i X_i, \frac{1}{n}\sum_i X_i^2\right)$$

είναι Α.Ο.Ε.Δ εκτιμητής της

$$g'(\boldsymbol{\theta}) = \left(g_1(\boldsymbol{\theta}), g_2(\boldsymbol{\theta})\right) = \left(\mu, \mu^2 + \sigma^2\right) = \left(\theta_1, \theta_1^2 + \theta_2\right)$$

με

$$VarT_1^* = I_{11}^{-1}(\boldsymbol{\theta}) = \frac{\sigma^2}{n}, \; VarT_2^* = I_{22}^{-1}(\boldsymbol{\theta}) = \frac{2\sigma^4}{n}.$$

Έτσι δεν μπορεί να εκτιμηθεί δια μέσου της ανισότητας Cramer-Rao η παράμετρος $\boldsymbol{\theta}' = \left(\mu, \sigma^2\right)$,γιατί η μόνη παραμετρική εκτιμίσιμη συνάρτηση δια μέσου της ανισότητας Cramer-Rao είναι της μορφής

$$g'(\boldsymbol{\theta}) = \left(ET_1, ET_2\right) = \left(\mu, \mu^2 + \sigma^2\right)$$

ή γραμμικοί συνδυασμοί της $g(\boldsymbol{\theta})$.

(βλέπε Zografos and Ferentinos (1994)

Το παράδειγμα που ακολουθεί αποτελεί εφαρμογή της εκτίμησης σε σημείο για την πολυδιάστατη περίπτωση με $r = 5$.

Παράδειγμα 2.10 (Διδιάστατη Κανονική κατανομή)

Έστω

$$\mathbf{X}_i = \begin{pmatrix} X_{11} \\ X_{12} \end{pmatrix} \sim N_2\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}\right)$$

για $i = 1, 2, \ldots, n$ να είναι ανεξάρτητα τυχαία διανύσματα από την κανονική διδιάστατη κατανομή με $\boldsymbol{\theta} = (\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)' = (\theta_1, \theta_2, \theta_3, \theta_4, \theta_5)' \in \Theta$ όπου :

$$\Theta = \left\{\left(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho\right)' : -\infty < \alpha_{j1} < \mu_j < \alpha_{j2} < \infty, 0 < b_{j1} < \sigma_j^2 < b_{j2} < \infty, -1 < c_1 < \rho < c_2 < 1, j = 1, 2\right\}$$

Η σ.π.π δίνεται ως εξής

$$f(x,\theta)=\prod_{i=1}^{n}\frac{1}{\sigma_1\sigma_2\sqrt{1-\rho^2 2\pi}}\exp\left\{\frac{-1}{2\left(1-\rho^2\right)}\sum_{i=1}^{n}\left[\left(\frac{x_{i1}-\mu_1}{\sigma_1}\right)^2-2\left(\frac{x_{i1}-\mu_1}{\sigma_1}\right)\left(\frac{x_{i2}-\mu_2}{\sigma_2}\right)+\left(\frac{x_{i2}-\mu_2}{\sigma_2}\right)^2\right]\right\}$$

Ανήκει στην εκθετική οικογένεια και από τη σχέση (2.20) έχουμε ότι

$$\mathbf{T}(\mathbf{x})=\left(T_1(\mathbf{x}),T_2(\mathbf{x}),T_3(\mathbf{x}),T_4(\mathbf{x}),T_5(\mathbf{x})\right)'$$

$$=\left(\sum_{i=1}^{n}x_{i1},\sum_{i=1}^{n}x_{i2},\sum_{i=1}^{n}x_{i1}{}^2,\sum_{i=1}^{n}x_{i2}{}^2,\sum_{i=1}^{n}x_{i1}x_{i2}\right)',$$

$$c(\boldsymbol{\theta})=\prod_{i=1}^{n}\frac{1}{\sigma_1\sigma_2\sqrt{1-\rho^2 2\pi}},$$

$$\mathbf{Q}(\boldsymbol{\theta})=\left(Q_1(\boldsymbol{\theta}),Q_2(\boldsymbol{\theta}),Q_3(\boldsymbol{\theta}),Q_4(\boldsymbol{\theta}),Q_5(\boldsymbol{\theta})\right)'$$

$$=\left(\frac{\mu_1-\mu_2\rho\sigma_1/\sigma_2}{\left(1-\rho^2\right)\sigma_1{}^2},\frac{\mu_2-\mu_1\rho\sigma_2/\sigma_1}{\left(1-\rho^2\right)\sigma_2{}^2},\frac{-1}{2\left(1-\rho^2\right)\sigma_1{}^2},\frac{-1}{2\left(1-\rho^2\right)\sigma_2{}^2},\frac{\rho}{\left(1-\rho^2\right)\sigma_1\sigma_2}\right)',$$

$$h(\mathbf{x})=1,$$

και

$$E\left[\mathbf{T}(\mathbf{x})\right]=\left(n\mu_1,n\mu_2,n\left(\sigma_1{}^2+\mu_1{}^2\right),n\left(\sigma_2{}^2+\mu_2{}^2\right),n\left(\rho\sigma_1\sigma_2+\mu_1\mu_2\right)\right)'.$$

Ο πίνακας πληροφορίας του Fisher και ο αντίστροφος του έχουν ως εξής

$$I_{\mathbf{x}}^F(\boldsymbol{\theta})=\frac{1}{\left(1-p^2\right)}\begin{pmatrix}1/\sigma_1^2 & -\rho/(\sigma_1\sigma_2) & 0 & 0 & 0\\ -\rho/(\sigma_1\sigma_2) & 1/\sigma_2^2 & 0 & 0 & 0\\ 0 & 0 & (2-\rho^2)/4(\sigma_1^4) & -\rho^2/4(\sigma_1^2\sigma_2^2) & -\rho/2(\sigma_1^2)\\ 0 & 0 & -\rho^2/4(\sigma_1^2\sigma_2^2) & (2-\rho^2)/4(\sigma_2^4) & -\rho/2(\sigma_2^2)\\ 0 & 0 & -\rho/2(\sigma_1^2) & -\rho/2(\sigma_2^2) & (1+\rho^2)/(1-\rho^2)\end{pmatrix}$$

$$
\left(I_{\mathbf{x}}^{F}(\boldsymbol{\theta})\right)^{-1}=\begin{pmatrix}
\sigma_1^2 & \rho\sigma_1\sigma_2 & 0 & 0 & 0\\
\rho\sigma_1\sigma_2 & \sigma_2^2 & 0 & 0 & 0\\
0 & 0 & 2\sigma_1^4 & 2\rho^2\sigma^2{}_1\sigma^2{}_2 & \rho(1-\rho^2)\sigma_1^2\\
0 & 0 & 2\rho^2\sigma^2{}_1\sigma^2{}_2 & 2\sigma_2^4 & \rho(1-\rho^2)\sigma_2^2\\
0 & 0 & \rho(1-\rho^2)\sigma_1^2 & \rho(1-\rho^2)\sigma_2^2 & (1-\rho^2)^2
\end{pmatrix}.
$$

Άρα ο

$$
\mathbf{T}^*(\mathbf{x})=\left(T_1^*(\mathbf{x}),T_2^*(\mathbf{x}),T_3^*(\mathbf{x}),T_4^*(\mathbf{x}),T_5^*(\mathbf{x})\right)'
$$

$$
=\left(\frac{1}{n}\sum_{i=1}^{n}x_{i1},\frac{1}{n}\sum_{i=1}^{n}x_{i2},\frac{1}{n}\sum_{i=1}^{n}x_{i1}{}^2,\frac{1}{n}\sum_{i=1}^{n}x_{i2}{}^2,\frac{1}{n}\sum_{i=1}^{n}x_{i1}x_{i2}\right)'
$$

είναι ο Α.Ο.Ε.Δ εκτιμητής της

$$
\mathbf{g}(\boldsymbol{\theta})=\left(g_1(\boldsymbol{\theta}),g_2(\boldsymbol{\theta}),g_3(\boldsymbol{\theta}),g_4(\boldsymbol{\theta}),g_5(\boldsymbol{\theta})\right)'
$$

$$
=\left(\mu_1,\mu_2,\left(\sigma_1{}^2+\mu_1{}^2\right),\left(\sigma_2{}^2+\mu_2{}^2\right),\left(\rho\sigma_1\sigma_2+\mu_1\mu_2\right)\right)'
$$

$$
=\left(\theta_1,\theta_2,\left(\theta_3+\theta_1{}^2\right),\left(\theta_4+\theta_2{}^2\right),\left(\theta_5\theta_3\theta_4+\theta_1\theta_2\right)\right)'
$$

με

$$
VarT_1^*=\frac{\sigma_1{}^2}{n}=\frac{\theta_3}{n}
$$

$$
VarT_2^*=\frac{\sigma_2{}^2}{n}=\frac{\theta_4}{n}
$$

$$
VarT_3^*=\frac{2\sigma_1{}^4}{n}=\frac{2\theta_3{}^2}{n} \qquad (\text{αφού } VarT_i^*=\frac{I_{ii}{}^{-1}(\theta)}{n})
$$

$$
VarT_4^*=\frac{2\sigma_2{}^4}{n}=\frac{2\theta_4{}^2}{n}
$$

$$
VarT_5^*=\frac{\left(1-\rho^2\right)^2}{n}=\frac{\left(1-\theta_5{}^2\right)^2}{n}.
$$

Έτσι δεν μπορεί να εκτιμηθεί δια μέσου της ανισότητας Cramer-Rao η παράμετρος $\boldsymbol{\theta}=(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)'=(\theta_1,\theta_2,\theta_3,\theta_4,\theta_5)'\in\Theta$, γιατί η μόνη παραμετρική εκτιμίσιμη συνάρτηση δια μέσου της ανισότητας Cramer-Rao είναι της μορφής

$$
\mathbf{v}(\boldsymbol{\theta})=\left(ET_1,ET_2,ET_3,ET_4,ET_5\right)'
$$

$$=\left(n\mu_1, n\mu_2, n\left(\sigma_1^2+\mu_1^2\right), n\left(\sigma_2^2+\mu_2^2\right), n\left(\rho\sigma_1\sigma_2+\mu_1\mu_2\right)\right)'$$

$$=\left(n\theta_1, n\theta_2, n\left(\theta_3+\theta_1^2\right), n\left(\theta_4+\theta_2^2\right), n\left(\theta_5\theta_3\theta_4+\theta_1\theta_2\right)\right)'.$$

ή γραμμικοί συνδυασμοί της $\mathbf{v}(\boldsymbol{\theta})$.

Έστω, ως εφαρμογή των παραπάνω, ότι θέλουμε να βρούμε έναν Α.Ο.Ε.Δ εκτιμητή για την $\mathbf{g}(\boldsymbol{\theta})=\mu_1-\mu_2$. Ορίζουμε $\mathbf{H}(\mathbf{x})=\frac{1}{n}\left(\sum_{i=1}^{n}x_{i1}-\sum_{i=1}^{n}x_{i2}\right)$ που είναι γραμμικός μετασχηματισμός των συνιστωσών του $\mathbf{T}$ από σχέση (2.24). Εφόσον η $f(\mathbf{x},\boldsymbol{\theta})$ ανήκει στην εκθετική οικογένεια η ανισότητα Cramer-Rao γίνεται ισότητα.

Άρα ο $\mathbf{H}(\mathbf{x})$ είναι Α.Ο.Ε.Δ εκτιμητής της

$$\mathbf{g}(\boldsymbol{\theta})=\mu_1-\mu_2$$

με

$$Var\mathbf{H}(\mathbf{x})=\mathbf{G}\left(I_X^F(\boldsymbol{\theta})\right)^{-1}\mathbf{G}' \qquad \text{(από σχέση (1.55))}$$

και $\mathbf{G}=(1,-1,0,0,0)$.

Τελικά μετά από πράξεις

$$Var\mathbf{H}(\mathbf{x})=\sigma_1^2+\sigma_2^2-2\rho\sigma_1\sigma_2$$

(βλέπε Ferguson (1996)).

Κεφάλαιο 3^{o}

Άλλες εφαρμογές της ανισότητας Cramer-Rao

3.1 Εισαγωγή

Στο κεφάλαιο αυτό θα μελετηθούν δυο επιπλέον εφαρμογές της ανισότητας Cramer-Rao στην στατιστική συμπερασματολογία. Η πρώτη αναφέρεται στη χρήση της ανισότητας στην εύρεση ενός κάτω φράγματος για την διακύμανση μιας συνάρτησης, έστω $g(x)$, μιας τυχαίας μεταβλητής X η οποία ακολουθεί κάποια γνωστή κατανομή (βλέπε Cacoullos (1982)). Αρχικά θα διατυπωθεί η θεωρία και στη συνέχεια θα ακολουθήσουν εφαρμογές της με τη X να ακολουθεί διαφορετικές κατανομές. Στη δεύτερη μελετάται μια βέλτιστη ιδιότητα των εκτιμητών μέγιστης πιθανοφάνειας (ασυμπτωτική κατανομή) που εμπεριέχει χρήση του κάτω φράγματος της ανισότητας Cramer-Rao. Τέλος, θα δοθούν εφαρμογές αυτής της ιδιότητας τόσο για την μονοδιάστατη όσο και για την πολυδιάστατη περίπτωση (βλέπε Klotz (2004), Ferguson (1996))

3.2 Κάτω φράγμα για διακύμανση συνάρτησης μιας τυχαίας μεταβλητής

Στη στατιστική είναι γνωστό ότι δεν μπορεί να βρεθεί, πάντα, η διακύμανση μιας συνάρτησης, έστω $g(x)$, μιας τυχαίας μεταβλητής X.Ένας τρόπος αντιμετώπισης αυτού του προβλήματος είναι όταν, χρησιμοποιώντας τις σειρές Taylor, προκύπτει μια προσεγγιστική έκφραση της διακύμανσης όπως φαίνεται παρακάτω.

Έστω $X_1,\ldots,X_n$ τ.μ. με αναμενόμενες τιμές $\theta_1,\ldots,\theta_n$, αντίστοιχα, και $\mathbf{X}=(X_1,\ldots,X_n)'$ και $\boldsymbol{\theta}=(\theta_1,\ldots,\theta_n)'$. Έστω, επίσης, μια διαφορίσιμη συνάρτηση $g(\mathbf{X})$ (εκτιμητής κάποιας παραμέτρου) για την οποία ζητούμενο είναι μια προσεγγιστική έκφραση της διακύμανσής της.

Έστω:

$$g_i'(\boldsymbol{\theta})=\frac{\partial}{\partial\theta_i}g(\boldsymbol{\theta}),i=1,\ldots,n. \tag{3.1}$$

Το ανάπτυγμα της g σε σειρά Taylor με κέντρο το $\boldsymbol{\theta}$ είναι:

$$g(\mathbf{x})=g(\boldsymbol{\theta})+\sum_{i=1}^{n}g_i'(\boldsymbol{\theta})(x_i-\theta_i)+R(\boldsymbol{\theta}), \tag{3.2}$$

όπου με $R(\boldsymbol{\theta})$ συμβολίζεται το υπόλοιπο. (βλέπε Ντούγιας (1998), Apostol(1974))

Για τη στατιστική προσέγγιση παραλείπεται το $R(\boldsymbol{\theta})$ και έχουμε :

$$g(\mathbf{x}) \approx g(\boldsymbol{\theta}) + \sum_{i=1}^{n} g_i'(\boldsymbol{\theta})(x_i - \theta_i) \ . \tag{3.3}$$

Παίρνοντας αναμενόμενες τιμές και στα δυο μέλη της σχέσης (3.3) προκύπτει:

$$\begin{aligned} E[g(\mathbf{X})] &\approx g(\boldsymbol{\theta}) + \sum_{i=1}^{n} g_i'(\boldsymbol{\theta}) E(X_i - \theta_i) \\ &= g(\boldsymbol{\theta}). \quad (\ E(X_i) = \theta_i \ ) \end{aligned} \tag{3.4}$$

Η προσεγγιστική διακύμανση της $g(\mathbf{X})$ είναι:

$$\begin{aligned} Var[g(\mathbf{X})] &= E\left([g(\mathbf{X}) - g(\boldsymbol{\theta})]^2 \right) && \text{(από 3.4)} \\ &= E\left(\left(\sum_{i=1}^{n} g_i'(\boldsymbol{\theta})(x_i - \theta_i) \right)^2 \right) && (3.5) \\ &= Var\left(\sum_{i=1}^{n} g_i'(\boldsymbol{\theta})(x_i - \theta_i) \right) + \left[E\left(\sum_{i=1}^{n} g_i'(\boldsymbol{\theta})(x_i - \theta_i) \right) \right]^2 \\ &= Var\left(\sum_{i=1}^{n} g_i'(\boldsymbol{\theta})(x_i - \theta_i) \right) \quad (\text{αφού } E(X_i) = \theta_i \ ) \\ &= \sum_{i=1}^{n} [g_i'(\boldsymbol{\theta})]^2 Var(X_i) + 2 \sum_{i<j}^{n} g_i'(\boldsymbol{\theta}) g_j'(\boldsymbol{\theta}) Cov(X_i, X_j) \ . \end{aligned}$$

Η τελική σχέση προέκυψε χρησιμοποιώντας γνωστό αποτέλεσμα σύμφωνα με το οποίο ισχύει:

$$Var\left(\sum_{i=1}^{n} X_i \right) = \sum_{i=1}^{n} Var(X_i) + 2 \sum_{1 \le i < j \le n} Cov(X_i, X_j) \ .$$

Η παραπάνω προσεγγιστική έκφραση της διακύμανσης είναι πολύ χρήσιμη γιατί χρησιμοποιεί, μόνο, απλές διακυμάνσεις και συνδιακυμάνσεις. (βλέπε Casella, Berger (2002))

Πέρα όμως από την προσεγγιστική έκφραση της διακύμανσης μιας συνάρτησης, ένας άλλος τρόπος αντιμετώπισης του προβλήματος είναι αν γνωρίζουμε ένα διάστημα, μέσα στο οποίο η διακύμανση της συνάρτησης παίρνει τιμές. Προς αυτή την κατεύθυνση έχουν βρεθεί εκφράσεις για το κάτω και άνω φράγμα της διακύμανσης μιας συνάρτησης (βλέπε Cacoullos (1982)).

Χρησιμοποιώντας την ανισότητα Cramer-Rao προκύπτει ένα κάτω φράγμα για την διακύμανση μιας συνάρτησης μιας τ.μ. X, με την σ.π.π. της X να πληρεί γνωστές συνθήκες κανονικότητας. Θα ακολουθήσουν εφαρμογές της παραπάνω θεωρίας για διάφορες τ.μ. που ακολουθούν γνωστές κατανομές. Επιπλέον, χρησιμοποιώντας στοιχεία από τη θεωρία πινάκων, θα δοθεί ένα κάτω φράγμα, πάλι με εφαρμογή της ανισότητας Cramer-Rao, για την πολυδιάστατη περίπτωση. Το κοινό χαρακτηριστικό των κάτω φραγμάτων για την διακύμανση είναι ότι εμπλέκουν, τις δύο πρώτες παραγώγους της g, στην συνεχή περίπτωση και τις δύο πρώτες διαφορές της g στην διακριτή.

Αρχικά θα δοθούν δυο λήμματα για τον τρόπο που προκύπτει γενικά ένα κάτω φράγμα για την διακύμανση μιας συνάρτησης και θα ακολουθήσουν εφαρμογές αυτών με την τ.μ X να ακολουθεί διάφορες κατανομές.

Λήμμα 3.1

Έστω $X_1,\ldots,X_n$ ανεξάρτητες τ.μ με κοινή σ.π.π. $f(x,\theta)$, $\theta\in\Theta$, όπου Θ ένα ανοιχτό διάστημα στο $\mathbb{R}$. Έστω ότι η f ικανοποιεί τις εξής συνθήκες κανονικότητας:

(i) Η $f(x,\theta)$ είναι θετική σε ένα διάστημα D ανεξάρτητο του θ.

(ii) Για κάθε $\theta\in\Theta$

$$E\left[\frac{\partial}{\partial\theta}\log f(X_i,\theta)\right]=0,\ i=1,\ldots,n.$$

Έστω, τώρα, $g:\mathbb{R}^n\to\mathbb{R}$ μια συνάρτηση τέτοια ώστε να υπάρχει η

$$E\left[g(\mathbf{X})\frac{\partial}{\partial\theta}\log\prod_{i=1}^{n}f(\mathbf{X}_i,\theta)\right]\equiv a(\theta) \qquad (3.6)$$

και $\mathbf{X}=(X_1,\ldots,X_n)$.

Τότε
$$Var\left[g(\mathbf{X})\right]\geq\frac{\left[a(\theta)\right]^2}{nI_x^F(\theta)}, \qquad (3.7)$$

όπου $I_x^F(\theta)$ είναι το μέτρο πληροφορίας του Fisher με

$$I_x^F(\theta)=E\left[\frac{\partial}{\partial\theta}\log f(X_i,\theta)\right]^2. \qquad (3.8)$$

Απόδειξη

Η απόδειξη ακολουθεί τα ίδια βήματα με την απόδειξη της ανισότητας Cramer-Rao για την μονοδιάστατη περίπτωση (βλέπε Θεώρημα 1.1- Θεώρημα 1.2)

Για τις τ.μ $g(\mathbf{X})$ και $W = \frac{\partial}{\partial\theta}\log\prod_{i=1}^{n} f(X_i,\theta)$ έχουμε

$$Var\left[g(\mathbf{X})\right] \geq \frac{\left[Cov\left(g(\mathbf{X}),W\right)\right]^2}{VarW}. \tag{3.9}$$

Ισχύει

$$\begin{aligned} W &= \frac{\partial}{\partial\theta}\log\prod_{i=1}^{n} f(X_i,\theta) \\ &= \sum_{i=1}^{n}\left(\frac{\partial}{\partial\theta}\log f(X_i,\theta)\right). \end{aligned}$$

Επειδή $X_1,\ldots,X_n$ ανεξάρτητες και λόγω του ότι $E\left[\frac{\partial}{\partial\theta}\log f(X_i,\theta)\right]=0$ έχουμε:

$$\begin{aligned} VarW &= \sum_{i=1}^{n} Var\left(\frac{\partial}{\partial\theta}\log f(X_i,\theta)\right) \\ &= \sum_{i=1}^{n}\left[E\left(\frac{\partial}{\partial\theta}\log f(X_i,\theta)\right)^2 - \left[E\left(\frac{\partial}{\partial\theta}\log f(X_i,\theta)\right)\right]^2\right] \\ &= nI_X^F(\theta). \end{aligned} \tag{3.10}$$

Ισχύει

$$\begin{aligned} Cov\left[g(\mathbf{X}),W\right] &= E\left[g(\mathbf{X})W\right] - E\left[g(\mathbf{X})\right]E(W) \\ &= E\left[g(\mathbf{X})\frac{\partial}{\partial\theta}\log\prod_{i=1}^{n} f(\mathbf{X}_i,\theta)\right] \equiv a(\theta).\ (\text{αφού } E(W)=0) \end{aligned} \tag{3.11}$$

Αντικαθιστώντας στη σχέση (3.9) από τις (3.10), (3.11) προκύπτει το ζητούμενο.

Όσον αφορά την πολυδιάστατη περίπτωση, το ανάλογο είναι το επόμενο .

Λήμμα 3.2

Έστω X τ.μ. με σ.π.π. $f(x,\boldsymbol{\theta})$ όπου $\boldsymbol{\theta}=(\theta_1,\ldots,\theta_n)$ και $\mathbf{g}(x)=(g_1(x),\ldots,g_r(x))$ τέτοια ώστε $Var[g_i(X)]<\infty$, $i=1,\ldots,r$. Τότε, ο πίνακας διακυμάνσεων-συνδιακυμάνσεων $\mathbf{D}\left[\mathbf{g}(X)\right]=(\sigma_{ij})$, με $\sigma_{ij}=Cov(g_i(X),g_j(X))$, ικανοποιεί τη σχέση

$$\mathbf{D}\left[\mathbf{g}(X)\right]\geq \mathbf{\Lambda}\left(I_x^F(\boldsymbol{\theta})\right)^{-1}\mathbf{\Lambda}', \tag{3.12}$$

όπου $\left(I_x^F(\boldsymbol{\theta})\right)^{-1}$ είναι ο αντίστροφος του πίνακα πληροφορίας του Fisher, δηλαδή:

$$I_x^F(\boldsymbol{\theta})=\left(I_{ij}\right) \text{ με } I_{ij}=\left\{E\frac{\partial^2 \log f(x,\boldsymbol{\theta})}{\partial\theta_i\partial\theta_j}\right\}, \qquad i,j=1,\ldots,n,$$

και

$$\mathbf{\Lambda}=\left(\lambda_{ij}\right)$$

με

$$\lambda_{ij}=\int_{-\infty}^{+\infty} g_i(x)\frac{\partial f(x,\boldsymbol{\theta})}{\partial\theta_j}dx \quad , \; i=1,\ldots,r \;,\; j=1,\ldots,n .$$

Η ανισότητα (3.9) γίνεται κατανοητή χρησιμοποιώντας το γνωστό αποτέλεσμα από την άλγεβρα ότι αν Α, Β μη-αρνητικά ορισμένοι πίνακες τότε $A\geq B$ αν και μόνον αν $A-B$ είναι μη αρνητικά ορισμένος.

(Για τις εφαρμογές που ακολουθούν βλέπε Cacoullos (1982))

Εφαρμογή 3.1 Η κανονική κατανομή

Θα εφαρμόσουμε το Λήμμα 3.1 για να λάβουμε ένα κάτω φράγμα για τη διακύμανση της συνάρτησης μιας τυχαίας μεταβλητής X που ακολουθεί την τυπική κανονική κατανομή $N(0,1)$.

Πρόταση 3.1

Έστω η τ.μ X ακολουθεί την τυπική κανονική κατανομή $N(0,1)$ και $g(x)$ μια συνάρτηση που και αυτή και η πρώτη παράγωγος της, $g'(x)$, παίρνουν πραγματικές τιμές στο $\mathbb{R}$ με $Var[g(X)]<\infty$ και $E\left\{\left|g'(X)\right|\right\}<\infty$. Τότε

$$Var[g(X)]\geq E^2[g'(X)]. \tag{3.13}$$

Η ανισότητα (3.13) γίνεται ισότητα αν και μόνον αν η $g(x)$ είναι γραμμική συνάρτηση.

Απόδειξη

Εφαρμόζουμε το Λήμμα 3.1 με $n=1$ και $\theta=\mu$ για μια τ.μ. X με $N(\mu,1)$.

Έχουμε:

$$\alpha(\theta)\equiv E\left[g(X)\frac{\partial}{\partial\theta}\log f(X,\theta)\right]$$

$$\overset{\theta=\mu}{\equiv}\int_{-\infty}^{+\infty}g(x)\frac{\partial}{\partial\mu}\left(\frac{1}{\sqrt{2\pi}}\exp\left[{-(x-\mu)^2}\Big/{2}\right]\right)dx$$

$$=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}g(x)(x-\mu)\exp\left[{-(x-\mu)^2}\Big/{2}\right]dx$$

$$=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}g(x)d\left\{\exp\left[{-(x-\mu)^2}\Big/{2}\right]\right\}$$

$$=\frac{1}{\sqrt{2\pi}}g(x)\exp\left[{-(x-\mu)^2}\Big/{2}\right]\Bigg|_{-\infty}^{+\infty}+\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}g'(x)\exp\left[{-(x-\mu)^2}\Big/{2}\right]dx$$

$$=\int_{-\infty}^{+\infty}g'(x)\frac{1}{\sqrt{2\pi}}\exp\left[{-(x-\mu)^2}\Big/{2}\right]dx$$

$$=\int_{-\infty}^{+\infty}g'(x)f(x,\theta)dx=E\left[g'(X)\right]$$

$$=E\left[g'(X)\right]$$

Παραπάνω χρησιμοποιήθηκε το γεγονός ότι

$$\lim_{x\to\pm\infty}g(x)\exp\left[{-(x-\mu)^2}\Big/{2}\right]=0\ , \qquad (E\{|g'(X)|\}<\infty).$$

Επίσης

$$I_x^F(\theta)=E\left[\frac{\partial}{\partial\theta}\log f(X,\theta)\right]^2$$

$$=E\left[\frac{\partial}{\partial\mu}\log f(X,\theta)\right]^2$$

$$=E\left[\frac{\partial}{\partial\mu}\left(-\log\sqrt{2\pi}-\frac{(X-\mu)^2}{2}\right)\right]^2$$

$$=E\left[(X-\mu)\right]^2$$

$$= Var(X)$$

$$= 1$$

για κάθε $\mu \in \mathbb{R}$.

Αντικαθιστώντας στη σχέση (3.7) του Λήμματος 3.1 όπου $\alpha(\theta) = \mathrm{E}\left[g'(X)\right]$ και $I_x^F(\theta) = 1$ προκύπτει το ζητούμενο, δηλαδή

$$Var[g(X)] \geq E^2[g'(X)].$$

Η ανισότητα γίνεται ισότητα αν $g(x) = ax + b$ για κάποιες σταθερές a,b.

Πράγματι

$$Var[g(X)] = Var[aX + b]$$

$$= a^2 Var(X)$$

$$= a^2.$$

$$g'(x) = a, E^2\left[g'(X)\right] = a^2.$$

Άρα

$$Var[g(X)] = E^2\left[g'(X)\right], \text{όταν}\, g(x) = ax + b.$$

Αντίστροφα αν $Var[g(X)] = E^2\left[g'(X)\right]$ τότε $g(x) = ax + b$,με την απόδειξη να είναι ανάλογη όπως στην ανισότητα Cramer-Rao , χρησιμοποιώντας την ανισότητα Cauchy-Scwarz.

<u>Παρατήρηση 3.1</u>

Εάν η $X \sim N(\mu, \sigma^2)$ τότε από το γνωστό μας μετασχηματισμό $X = \sigma Z + \mu$ $Z \sim N(0,1)$. Επομένως θέτοντας $g(x) = g_0(z)$ έχουμε:

$$g_o(z) = g(\sigma z + \mu)$$
$$g_o'(z) = \sigma g'(\sigma z + \mu)$$

Άρα $g_0'(z) = \sigma g'(x)$. Εφαρμόζοντας λοιπόν την Πρόταση 3.1 για την συνάρτηση $g_o(z)$ προκύπτουν τα ακόλουθα.

Πρόταση 3.2

Έστω $X \sim N(\mu,\sigma^2)$ και $g(x)$ με τις προϋποθέσεις της Πρότασης 3.1 (δηλαδή $\theta=\mu$ και σ^2 γνωστό). Τότε

$$Var[g(X)] \geq \sigma^2 E^2\left[g'(X)\right] \qquad (3.14)$$

με ισότητα αν και μόνον αν $g(x)=ax+b$ για κάποιες σταθερές α, b.

Απόδειξη

Εφόσον $Z \sim N(0,1)$ από την Πρόταση 3.1 έχουμε $Var[g_o(Z)] \geq E^2\left[g_o'(Z)\right]$ και αντικαθιστώντας όπου $g(x)=g_0(z)$ και $g_0'(z)=\sigma g'(x)$ προκύπτει το ζητούμενο.

Στη συνέχεια εφαρμόζουμε το Λήμμα 3.1 με την τ.μ. X να ακολουθεί την κατανομή $N(0,\sigma^2)$ και παίρνοντας ως $\theta=\sigma^2$ οδηγούμαστε στα ακόλουθα.

Πρόταση 3.3

Έστω $X \sim N(0,\sigma^2)$ και $g(x), g'(x)$ απολύτως συνεχείς και τέτοιες ώστε $E\{|g''(x)|\}<\infty$. Τότε

$$Var[g(X)] \geq \frac{\sigma^4}{2} E^2\left[g''(X)\right], \qquad (3.15)$$

όπου η ανισότητα γίνεται ισότητα αν και μόνον αν $g(x)=ax^2+b$ για κάποιες σταθερές α και b.

Απόδειξη

Έχουμε

$$a(\theta) = E\left(g(x)\frac{\partial}{\partial\theta}\log f(x,\theta)\right)$$

$$= \int g(x)\frac{\partial}{\partial\theta} f(x,\theta)dx$$

$$= \int_{-\infty}^{+\infty} g(x)\frac{\partial}{\partial\theta}\left[\frac{1}{\sqrt{2\pi\theta}}e^{-x^2/2\theta}\right]dx$$

$$= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} g(x)\left[-\frac{1}{2\theta}e^{-x^2/2\theta}+\frac{x^2}{2\theta^{5/2}}e^{-x^2/2\theta}\right]dx$$

$$=\frac{1}{2\sqrt{2\pi\theta}}\int_{-\infty}^{+\infty}\frac{1}{\theta}g(x)e^{-x^2/2\theta}dx+\frac{1}{2\sqrt{2\pi\theta}}\int_{-\infty}^{+\infty}g(x)\frac{x^2}{\theta^2}e^{-x^2/2\theta}dx$$

$$=\frac{1}{2\theta\sqrt{2\pi\theta}}\int_{-\infty}^{+\infty}g(x)e^{-x^2/2\theta}dx+\frac{1}{2\theta\sqrt{2\pi\theta}}\int_{-\infty}^{+\infty}g(x)x\frac{x}{\theta}e^{-x^2/2\theta}dx$$

$$=\frac{1}{2\theta\sqrt{2\pi\theta}}\int_{-\infty}^{+\infty}g(x)e^{-x^2/2\theta}dx-\frac{1}{2\theta\sqrt{2\pi\theta}}\left[g(x)x\left(e^{-x^2/2\theta}\right)'\right]_{-\infty}^{+\infty}$$

$$+\frac{1}{2\theta\sqrt{2\pi\theta}}\int_{-\infty}^{+\infty}[g'(x)x+g(x)]e^{-x^2/2\theta}dx$$

$$=\frac{1}{2\sqrt{2\pi\theta}}\int_{-\infty}^{+\infty}g'(x)\frac{x}{\theta}e^{-x^2/2\theta}dx$$

$$=-\frac{1}{2\sqrt{2\pi\theta}}\int_{-\infty}^{+\infty}g'(x)e^{-x^2/2\theta}dx$$

$$=\frac{1}{2\sqrt{2\pi\theta}}\left[g'(x)e^{-x^2/2\theta}\right]_{-\infty}^{+\infty}+\frac{1}{2\sqrt{2\pi\theta}}\int_{-\infty}^{+\infty}g''(x)e^{-x^2/2\theta}dx$$

$$=\frac{1}{2}\int_{-\infty}^{+\infty}g''(x)f(x,\theta)dx=\frac{1}{2}E[g''(X)]$$

$$=\frac{1}{2}E[g''(X)].$$

Άρα

$$a(\theta)\equiv\frac{1}{2}E[g''(X)].$$

$$I_x^F(\theta)=E\left(\frac{\partial}{\partial\theta}\log f(x,\theta)\right)^2=-E\left(\frac{\partial^2}{\partial\theta^2}\log f(x,\theta)\right)$$

$$=-E\left(\frac{1}{2\theta^2}-\frac{x^2}{\theta^3}\right)$$

$$=\frac{1}{\theta^2}-\frac{1}{2\theta^2}=\frac{1}{2\theta^2}=\frac{1}{2\sigma^4}.$$

Αντικαθιστώντας στην σχέση (3.7) προκύπτει ότι:

$$Var[g(X)] \geq \frac{\sigma^4}{2} E^2[g''(X)].$$

Η ισότητα ισχύει αν και μόνον αν (ανισότητα Cauchy-Swharz)

$$\frac{\partial}{\partial\theta}\log f(x,\theta) = -\frac{1}{2\theta} + \frac{x^2}{2\theta^2} = cg(x).$$

Δηλαδή αν και μόνον αν $g(x) = ax^2 + b$ για κάποιες σταθερές α και b.

<u>Παράδειγμα 3.1</u>

Θεωρούμε ως $g(X) = X^2$ και $X \sim N(0,\sigma^2)$. Από την παραπάνω πρόταση έχουμε, άμεσα, ότι:

$$Var[g(X)] \geq \frac{\sigma^4}{2} E^2[g''(X)].$$

$$g(x) = x^2, g'(x) = 2x, g''(x) = 2.$$

Άρα $$Var[X^2] \geq 4\frac{\sigma^4}{2}.$$

$$Var[X^2] \geq 2\sigma^4.$$

Και εφόσον $g(x)$ της μορφής $ax^2 + b$ με $a = 1$ και $b = 0$ προκύπτει $Var[X^2] = 2\sigma^4$. Στη συνέχεια πηγαίνουμε ένα βήμα παρακάτω παίρνοντας τ.μ. με $X \sim N(\mu,\sigma^2)$.

<u>Πρόταση 3.4</u>

Έστω $X \sim N(\mu,\sigma^2)$ (δηλαδή εδώ και το μ και το σ^2 είναι άγνωστα) και $g(x)$ όπως δίνεται στην Πρόταση 3.3 τέτοια ώστε $E\{|g'(X)|\} < \infty$, $E\{|g''(X)|\} < \infty$. Τότε

$$Var[g(X)] \geq \sigma^2 E^2[g'(X)] + \frac{1}{2}\sigma^4 E^2[g''(X)]. \tag{3.16}$$

Η ανισότητα γίνεται ισότητα αν και μόνον αν $g(x) = ax^2 + bx + c$ για κάποιες σταθερές a, b και c.

__Απόδειξη__

Αυτή η πρόταση αποτελεί εφαρμογή του Λήμματος 3.2 με $\boldsymbol{\theta}=(\theta_1,\theta_2)=(\mu,\sigma^2)$ και $X\sim N(\mu,\sigma^2)$. Η $g(x)$ ομοίως με πριν, δηλαδή $i=1$. Εφόσον $\boldsymbol{\theta}=(\theta_1,\theta_2)$ το $n=2$ και $j=1,2$.

Ο πίνακας πληροφορίας του Fisher, ως γνωστόν, δίνεται από τη σχέση:

$$I_x^F(\boldsymbol{\theta})=\begin{pmatrix}\frac{1}{\theta_2} & 0\\ 0 & \frac{1}{2\theta_2}\end{pmatrix},$$

και

$$\left[I_x^F(\boldsymbol{\theta})\right]^{-1}=\begin{pmatrix}\theta_2 & 0\\ 0 & 2\theta_2\end{pmatrix}. \tag{3.17}$$

Ο πίνακας $\boldsymbol{\Lambda}=(\lambda_{11},\lambda_{12})$ με λ_{ij} από τη σχέση

$$\lambda_{ij}=\int_{-\infty}^{+\infty} g_i(x)\frac{\partial f(x,\boldsymbol{\theta})}{\partial\theta_j}dx\ ,\ \begin{matrix} i=1\\ j=1,2\end{matrix}.$$

Ισχύει:

$$\lambda_{11}=\int_{-\infty}^{+\infty} g(x)\frac{\partial f(x,\theta_1,\theta_2)}{\partial\theta_1}dx$$

$$=\int_{-\infty}^{+\infty} g(x)\frac{\partial}{\partial\theta_1}\left(\frac{1}{\sqrt{2\pi\theta_2}}e^{\frac{-(x-\theta_1)^2}{2\theta_2}}\right)dx$$

$$=\int_{-\infty}^{+\infty} g(x)\frac{1}{\sqrt{2\pi\theta_2}}\frac{(x-\theta_1)}{\theta_2}\left(e^{\frac{-(x-\theta_1)^2}{2\theta_2}}\right)dx$$

$$=\int_{-\infty}^{+\infty} -g(x)\frac{1}{\sqrt{2\pi\theta_2}}\left(e^{\frac{-(x-\theta_1)^2}{2\theta_2}}\right)' dx$$

$$= \left[-\frac{1}{\sqrt{2\pi\theta_2}} g(x) e^{\frac{-(x-\theta_1)^2}{2\theta_2}} \right]_{-\infty}^{+\infty} + \int_{-\infty}^{+\infty} g'(x) \frac{1}{\sqrt{2\pi\theta_2}} e^{\frac{-(x-\theta_1)^2}{2\theta_2}} dx$$

$$= \int_{-\infty}^{+\infty} g'(x) f(x, \theta_1, \theta_2) dx = E\left[g'(X)\right]$$

$$= E\left[g'(X)\right] \tag{3.18}$$

Ομοίως αποδεικνύεται ότι

$$\lambda_{12} = \frac{1}{2} E[g''(X)]. \tag{3.19}$$

Άρα από (3.17) και (3.18):

$$\Lambda = (E[g'(X)], \frac{1}{2} E[g''(X)]). \tag{3.20}$$

Επειδή $i = 1$ έχουμε ότι $\Sigma \equiv Var[g(X)]$. Επομένως από τη σχέση (3.9) και από τις (3.16) και (3.19) προκύπτει:

$$Var[g(X)] \geq (E[g'(X)], \frac{1}{2} E[g''(X)]) \begin{pmatrix} \theta_2 & 0 \\ 0 & 2\theta_2 \end{pmatrix} \begin{pmatrix} E[g'(X)] \\ \frac{1}{2} E[g''(X)] \end{pmatrix}$$

ή

$$Var[g(X)] \geq \theta_2 E^2 \left[g'(X)\right] + \frac{1}{2} \theta_2^{\,2} E^2 \left[g''(X)\right].$$

Αντικαθιστώντας $\theta_2 = \sigma^2$ έχουμε:

$$Var[g(X)] \geq \sigma^2 E^2 \left[g'(X)\right] + \frac{\sigma^4}{2} E^2 \left[g''(X)\right]$$

το οποίο είναι και το ζητούμενο.

<u>Παρατήρηση 3.2</u>

Από τις παραπάνω προτάσεις και από τις σχέσεις (3.13), (3.14), (3.15) και (3.16) παρατηρείται ότι δεν μπορούμε, από την τελική σχέση (3.16) και αντικαθιστώντας κατάλληλα είτε το μ είτε το σ^2, να καταλήξουμε στις υπόλοιπες.

Παρατηρείται ότι στην (3.16) το κάτω φράγμα βελτιώνεται, με την έννοια ότι μεγαλώνει, συγκρινόμενο με τις σχέσεις (3.14) και (3.15). Αυτό συμβαίνει γιατί το κάτω φράγμα της (3.14) εφαρμόζεται είτε το σ^2 είναι γνωστό είτε άγνωστο (ο $\bar{X}$ είναι αμερόληπτος και επαρκής εκτιμητής της μ), ενώ το κάτω φράγμα της (3.15) εφαρμόζεται μόνο αν το μ είναι γνωστό. Γενικά, αν στο πρόβλημα της εκτίμησης εμπλέκονται αρκετές παράμετροι $\theta_1,\ldots,\theta_n$ τότε το κάτω φράγμα Cramer-Rao για ένα εκτιμητή της παραμέτρου, έστω, θ_i μπορεί να μη λαμβάνεται εκτός και αν ο εκτιμητής τυγχάνει να είναι ανεξάρτητος των άλλων παραμέτρων. Αλλιώς, όπως στην συγκεκριμένη περίπτωση, το κάτω φράγμα αυξάνει.

Από την άλλη πλευρά το κάτω φράγμα για την διακύμανση της g(x), αν το δούμε σαν έναν εκτιμητή της αναμενόμενης τιμής $E[g(X)]=\gamma(\mu,\sigma^2)$, μπορεί να εκφραστεί με όρους του πίνακα πληροφορίας του Fisher $I_x^F(\boldsymbol{\theta})$ και των μερικών παραγώγων της γ. Η έκφραση τελικά που προκύπτει είναι ανεξάρτητη από κάθε άλλο επιλεγμένο αμερόληπτο εκτιμητή της $\gamma(\mu,\sigma^2)$, όπως είναι η $g(X)$. Όντως, από την ανισότητα Cramer-Rao για την πολυδιάστατη περίπτωση προκύπτει:

$$Var\left[g(X)\right]\geq\sigma^2\left(\frac{\partial\gamma}{\partial\mu}\right)^2+2\sigma^4\left(\frac{\partial\gamma}{\partial\sigma^2}\right)^2 \tag{3.21}$$

(βλέπε Cacoullos (1982)).

Παρόλα αυτά, αυτό το κάτω φράγμα εμπλέκει τις δυο πρώτες παραγώγους της g, όπως φαίνεται και στην σχέση (3.16), κάτι που δείχνει την εξάρτηση από την συνάρτηση g.

Στην συνέχεια θα δούμε πώς προκύπτουν κάτω όρια για τη διακύμανση μιας συνάρτησης $g(x_1,\ldots,x_n)$, με τα x_i να ακολουθούν την κανονική κατανομή. Χρησιμοποιώντας το Λήμμα 3.1 και δουλεύοντας ανάλογα με την Πρόταση 3.1 προκύπτουν τα ακόλουθα.

Πρόταση 3.5

Έστω $\mathbf{X}=(X_1,\ldots,X_n)$ όπου τα X_i ανεξάρτητες τ.μ. που ακολουθούν κανονική κατανομή $N(\mu,\sigma^2)$. Υποθέτουμε ότι η συνάρτηση $g(x_1,\ldots,x_n)$ έχει μερικές παραγώγους $g_i(x_1,\ldots,x_n)=\frac{\partial}{\partial x_i}g(x_1,\ldots,x_n)$ και $E\{|g_i(X)|\}<\infty$, $i=1,\ldots,n$. Τότε

$$Var[g(\mathbf{X})]\geq \frac{\sigma^2}{n}\left\{\sum_{i=1}^{n}E[g_i(\mathbf{X})]\right\}^2. \tag{3.22}$$

Όπου η ανισότητα γίνεται ισότητα αν και μόνο αν $g(\mathbf{x})=a\sum_{i=1}^{n}x_i+b$ για κάποιες σταθερές α και b.

Εφαρμόζοντας πάλι το Λήμμα 3.1 με $\theta=\sigma^2$ και προχωρώντας όπως στην Πρόταση 3.3 προκύπτουν τα ακόλουθα.

Πρόταση 3.6

Έστω $\mathbf{X}=(X_1,\ldots,X_n)$ όπου τα X_i ανεξάρτητες τ.μ. που ακολουθούν $N(\mu,\sigma^2)$. Έστω $g(x)$ απολύτως συνεχής με $g_i(x)=\frac{\partial}{\partial x_i}g(x)$ απολύτως συνεχής και $E\{|g_{ii}(X)|\}<\infty$, όπου $g_{ii}(x)=\frac{\partial^2}{\partial x_i^2}g(x)$, $i=1,\ldots,n$.

Τότε

$$Var[g(\mathbf{X})]\geq \frac{\sigma^4}{2n}E^2[\sum_{i=1}^{n}g_{ii}(\mathbf{X})]. \tag{3.23}$$

Όπου η ανισότητα γίνεται ισότητα αν και μόνο αν $g(\mathbf{x})=a\sum_{i=1}^{n}x_i^2+b$ για κάποιες σταθερές α και b.

Σαν μια εφαρμογή του Λήμματος 3.2, θεωρούμε την περίπτωση της πολυδιάστατης κανονικής $N(\boldsymbol{\mu},\boldsymbol{\Sigma})$ με $\boldsymbol{\theta}=\boldsymbol{\mu}=(\mu_1,\ldots,\mu_n)'$ και $r=1$.

Πρόταση 3.7

Έστω $\mathbf{X} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ και $g(\mathbf{x})$ μια παραγωγίσιμη συνάρτηση που παίρνει πραγματικές τιμές και είναι ορισμένη στο $\mathbb{R}^n$. Ορίζουμε ως $\mathbf{G} = (g_1, \ldots, g_n)'$ το ανάστροφο διάνυσμα των μερικών παραγώγων με $g_i(\mathbf{x}) = \frac{\partial}{\partial x_i} g(\mathbf{x})$. Υποθέτουμε ότι $E\{|g_i(\mathbf{X})|\} < \infty$. Τότε

$$Var[g_i(\mathbf{X})] \geq E[\mathbf{G}'] \boldsymbol{\Sigma} E[\mathbf{G}] \tag{3.24}$$

Η ανισότητα γίνεται ισότητα αν και μόνο αν $g(\mathbf{x}) = a_1 x_1 + \ldots + a_n x_n + b$ για κάποιες σταθερές τιμές $a_1, \ldots, a_n$ και b.

Απόδειξη

Θεωρώντας $\boldsymbol{\theta} = \boldsymbol{\mu}$ ο πίνακας πληροφορίας του Fischer έχει ως εξής:

$$I_x^F(\boldsymbol{\theta}) = \left(-E\left(\frac{\partial^2 \log f(\mathbf{x}, \boldsymbol{\mu})}{\partial \mu_i \partial \mu_j} \right) \right) = \boldsymbol{\Sigma}^{-1} \tag{3.25}$$

και ισούται με τον αντίστροφο πίνακα διακυμάνσεων-συνδιακυμάνσεων. Ενώ το διάνυσμα $\boldsymbol{\Lambda} = (\lambda_1, \ldots, \lambda_n)'$ δίνεται ως εξής:

$$\begin{aligned}
\lambda_j &= \int_{\mathbb{R}^n} g(\mathbf{x}) \frac{\partial f(\mathbf{x}, \boldsymbol{\mu})}{\partial \mu_j} d\mathbf{x} \\
&= -\int_{\mathbb{R}^n} g(\mathbf{x}) f_j(\mathbf{x}, \boldsymbol{\mu}) d\mathbf{x} \\
&= \left[g(\mathbf{x}) f(\mathbf{x}, \boldsymbol{\mu}) \right] \Big\| + \int_{\mathbb{R}^n} \left[g_j(\mathbf{x}) \right] f(\mathbf{x}, \boldsymbol{\mu}) d\mathbf{x} \\
&= \int_{\mathbb{R}^n} \left[g_j(\mathbf{x}) \right] f(\mathbf{x}, \boldsymbol{\mu}) d\mathbf{x} = E[g_j(\mathbf{X})].
\end{aligned}$$

Όπου

$$g_j(\mathbf{x}) = \frac{\partial g(\mathbf{x})}{\partial x_j}$$

και

$$f_j(\mathbf{x}, \boldsymbol{\mu}) = \frac{\partial f(\mathbf{x}, \boldsymbol{\mu})}{\partial x_j}.$$

Άρα $\quad \boldsymbol{\Lambda} = E[\mathbf{G}']$ και $\boldsymbol{\Lambda}' = E[\mathbf{G}]$

Αντικαθιστώντας από την σχέση (3.9) προκύπτει το ζητούμενο

$$Var[g(\mathbf{x})] \geq E[\mathbf{G}']\Sigma^{-1} E[\mathbf{G}].$$

Θα ολοκληρώσουμε την εφαρμογή της κανονικής κατανομής με μια σχέση ανάλογη της (3.23) που περιλαμβάνει πίνακες.

Πρόταση 3.8

Έστω $\mathbf{X}$ ακολουθεί κανονική κατανομή $N(0,I)$ και έστω $\mathbf{g}(\mathbf{X})=(g_1(\mathbf{X}),...,g_r(\mathbf{X}))$ όπου κάθε g_i είναι διαφορίσιμη στο $\mathbb{R}^n$. Θέτουμε

$g_{ij}=\dfrac{\partial g_i(\mathbf{x})}{\partial x_j}$, $\mathbf{\Gamma}=(g_{ij})$, $i=1,...,r,\ j=1,...,n.$ Τότε

$$\mathbf{D}[\mathbf{g}(\mathbf{X})] \geq (E\mathbf{\Gamma})(E\mathbf{\Gamma})'. \qquad (3.26)$$

(Όπου I ο μοναδιαίος πίνακας)

Η ανισότητα γίνεται ισότητα αν και μόνο αν κάθε g_i είναι γραμμική συνάρτηση, δηλαδή αν, $\mathbf{g}(\mathbf{X})=\mathrm{A}\mathbf{X}+\mathbf{b}$, όπου Α ένας πίνακας με στοιχεία σταθερές και $\mathbf{b}$ ένα σταθερό διάνυσμα.

Απόδειξη

Αυτό το κάτω φράγμα θα προκύψει χρησιμοποιώντας το Λήμμα 3.2 για την κανονική κατανομή $N(\boldsymbol{\mu},I)$ και τη συνάρτηση-διάνυσμα $\mathbf{g}$. Έχουμε

$$\begin{aligned}
\lambda_{ij} &= \int_{\mathbb{R}^n} g_i(\mathbf{x})\frac{\partial f(\mathbf{x},\boldsymbol{\mu})}{\partial \mu_j}d\mathbf{x} = -\int_{\mathbb{R}^n} g_i(\mathbf{x})D_j f(\mathbf{x},\boldsymbol{\mu})d\mathbf{x} \\
&= \int_{\mathbb{R}^n} \frac{\partial g_i(\mathbf{x})}{\partial x_j} f(\mathbf{x},\boldsymbol{\mu})d\mathbf{x} \\
&= \int_{\mathbb{R}^n} g_{ij}(\mathbf{x}) f(\mathbf{x},\boldsymbol{\mu})d\mathbf{x} \\
&= E\left[g_{ij}(\mathbf{X})\right], i=1,..r, j=1,...n.
\end{aligned}$$

Ομοίως με πριν

$$I_x^F(\boldsymbol{\theta}) = \left(-E\left(\frac{\partial^2 \log f(\mathbf{x},\boldsymbol{\mu})}{\partial\mu_i\partial\mu_j}\right)\right) = I^{-1} = I.$$

Για τον πίνακα Λ έχουμε ότι $\Lambda = E\Gamma$.

Επομένως αντικαθιστώντας τα παραπάνω στη σχέση (3.9) προκύπτει το ζητούμενο

$$\mathrm{D}\left[\mathrm{g}(\mathrm{X})\right] \geq (E\Gamma)(E\Gamma)'.$$

Εφαρμογή 3.2 Η Γάμμα κατανομή

Πρόταση 3.9

Έστω $g(x)$ μια διαφορίσιμη συνάρτηση τέτοια ώστε η $g(X)$ να έχει πεπερασμένη διακύμανση και η τ.μ. X να ακολουθεί κατανομή με σ.π.π.

$$f(x,\lambda) = \frac{\lambda^{\kappa}}{\Gamma(\kappa)} x^{\kappa-1} e^{-\lambda x},\ (x>0,\kappa>0), \tag{3.27}$$

τότε

$$Var[g(X)] \geq \frac{1}{\kappa} E^2[Xg'(X)] \tag{3.28}$$

με ισότητα αν και μόνον αν η $g(x)$ είναι γραμμική συνάρτηση του x.

Απόδειξη

Εφαρμόζοντας το Λήμμα 3.1 για $n=1$ και $\theta=\lambda$ για την σ.π.π. που δίνεται από τη σχέση (3.26) προκύπτει:

$$\begin{aligned}
a(\theta) &\equiv \int_{-\infty}^{+\infty} g(x) \frac{\partial}{\partial\theta} f(x,\theta) dx \\
&= \int_0^{+\infty} g(x) \frac{x^{\kappa-1}}{\Gamma(\kappa)} [\kappa\lambda^{\kappa-1} e^{-\lambda x} - \lambda^{\kappa} x e^{-\lambda x}] dx \\
&= \int_0^{+\infty} g(x) \frac{x^{\kappa-1}}{\Gamma(\kappa)} \kappa\lambda^{\kappa-1} e^{-\lambda x} dx + \int_0^{+\infty} g(x) \frac{\lambda^{\kappa-1}}{\Gamma(\kappa)} x^{\kappa} \left(e^{-\lambda x}\right)' dx \\
&= \int_0^{+\infty} g(x) \frac{x^{\kappa-1}}{\Gamma(\kappa)} \kappa\lambda^{\kappa-1} e^{-\lambda x} dx + \left[g(x) \frac{\lambda^{\kappa-1}}{\Gamma(\kappa)} x^{\kappa} e^{-\lambda x} \right]_0^{+\infty} \\
&\quad - \int_0^{+\infty} \frac{\lambda^{\kappa-1}}{\Gamma(\kappa)} e^{-\lambda x} \left(g'(x) x^{\kappa} + \kappa g(x) x^{\kappa-1} \right) dx \\
&= -\frac{1}{\lambda} \int_0^{+\infty} g'(x) x \frac{\lambda^{\kappa}}{\Gamma(\kappa)} x^{\kappa-1} e^{-\lambda x} dx \\
&= -\frac{1}{\lambda} \int_0^{+\infty} g'(x) x f(x,\lambda) dx
\end{aligned}$$

$$= -\frac{1}{\lambda} E\left(Xg'(X)\right).$$

Ισχύει $I_x^F(\theta) = \frac{\kappa}{\lambda^2}$.

Αντικαθιστώντας στη σχέση (3.7) προκύπτει ότι

$$Var[g(X)] \geq \frac{1}{\kappa} E^2[Xg'(X)],$$

το οποίο είναι και το ζητούμενο.

Είναι γνωστό ότι η κατανομή Γάμμα αποτελεί γενίκευση της εκθετικής κατανομής. Επομένως μπορούμε να πάρουμε αντίστοιχο συμπέρασμα με την Πρόταση 3.9 εφαρμόζοντας το Λήμμα 3.1 με $n=1$, $\kappa=1$ και θ να ισούται με την παράμετρο θέσης λ με την τ.μ. X να ακολουθεί κατανομή Εκθ(θ).

Πρόταση 3.10 (Η εκθετική κατανομή)

Έστω X τ.μ. που ακολουθεί την εκθετική κατανομή με παράμετρο θ και έστω $g(x)$ μια διαφορίσιμη συνάρτηση τέτοια ώστε η $g(x)$ να έχει πεπερασμένη διακύμανση. Τότε

$$Var[g(X)] \geq E^2[Xg'(X)], \qquad (3.29)$$

με ισότητα αν και μόνον αν η $g(x)$ είναι γραμμική συνάρτηση του x.

Παράδειγμα 3.2

Έστω $g(X) = X^2 + X$ με X τ.μ. που έχει σ.π.π. όπως δίνεται από τη σχέση (3.26). Από την Πρόταση 3.9 έχουμε ότι:

$$Var[g(X)] \geq \frac{1}{\kappa} E^2[Xg'(X)].$$

$$g(x) = x^2 + x, g'(x) = 2x + 1.$$

$$E[Xg'(X)] = E[2X^2 + X] = 2E(X^2) + E(X)$$

$$= \left\{2Var(X) + [E(X)]^2\right\} + E(X)$$

$$= 2\left\{\frac{\kappa}{\lambda^2} + \frac{\kappa}{\lambda^2}\right\} + \frac{\kappa}{\lambda}$$

$$= \frac{2(\kappa + \kappa^2)}{\lambda^2} + \frac{\kappa}{\lambda}.$$

Άρα

$$Var[g(X)] \geq \kappa\left(\frac{2+2\kappa+\lambda}{\lambda^2}\right)^2$$

Εάν όμως η g(x) είναι γραμμική π.χ. $g(X)=3X+2$ τότε, άμεσα, $Var[g(X)]=\frac{3}{\lambda}$.

Στην περίπτωση που $\mathbf{X}=(X_1,\ldots,X_n)$ με X_i ανεξάρτητες τ.μ. που ακολουθούν την εκθετική κατανομή με την ίδια παράμετρο θ, προκύπτει αντίστοιχη σχέση με την (3.28) δηλαδή:

$$Var[g(\mathbf{X})] \geq \frac{1}{n}E^2[\mathbf{X}G]. \tag{3.30}$$

Ομοίως, θεωρώντας διαφορετικές παραμέτρους για τα X_i καταλήγουμε στην εξής ανισότητα:

$$Var[g(\mathbf{X})] \geq \sum_{i=1}^{n} E^2[X_i g_i(\mathbf{X})], \tag{3.31}$$

όπου η ανισότητα γίνεται ισότητα αν η g είναι γραμμική συνάρτηση των $X_1,\ldots,X_n$.

Ας δούμε τώρα μερικές εφαρμογές με τ.μ διακριτές Τα κάτω φράγματα προκύπτουν πάλι από το Λήμμα 3.1 ανάλογα με τη συνεχή περίπτωση αντικαθιστώντας όμως τα ολοκληρώματα με αθροίσματα .

Εφαρμογή 3.3 Η Poisson κατανομή

Πρόταση 3.11

Έστω X τ.μ. που ακολουθεί την Poisson κατανομή με παράμετρο λ και $g(x)$ μια συνάρτηση που παίρνει πραγματικές τιμές, ορισμένη στο σύνολο $U_0=\{0,1,2,\ldots\}$ έτσι ώστε η $Var[g(X)]$ να είναι πεπερασμένη. Τότε

$$Var[g(X)] \geq \lambda E^2[\Delta g(X)] \tag{3.32}$$

όπου $\Delta g(x)=g(x+1)-g(x)$ η διαφορά πρώτης τάξεως προς το εμπρός στο σημείο x.Η ανισότητα γίνεται ισότητα αν η $g(x)$ είναι σταθερή.

Απόδειξη

Εφαρμόζουμε το Λήμμα 3.1 για την σ.π.π. της Poisson κατανομής $f(x,\theta)$ με $\theta=\lambda$. Έτσι προκύπτει

$$
\begin{aligned}
a(\theta) &\equiv \mathrm{E}\left[g(X)\frac{\partial}{\partial\theta}\log f(X,\theta)\right] \\
&= \sum_{x=0}^{+\infty}\left[g(x)\frac{\partial}{\partial\theta}f(x,\theta)\right] \\
&= \sum_{x=0}^{+\infty}\left\{g(x)\frac{x\theta^{x-1}e^{-\theta}-\theta^{x}e^{-\theta}}{x!}\right\} \\
&= e^{-\theta}\left\{\sum_{x=0}^{+\infty}g(x)\frac{x\theta^{x-1}}{x!}-\sum_{x=0}^{+\infty}g(x)\frac{\theta^{x}}{x!}\right\} \\
&= e^{-\theta}\left\{\sum_{x=0}^{+\infty}g(x)\frac{\theta^{x-1}}{(x-1)!}-\sum_{x=0}^{+\infty}g(x)\frac{\theta^{x}}{x!}\right\} \\
&= e^{-\theta}\left\{\sum_{x=1}^{+\infty}g(x)\frac{\theta^{x-1}}{(x-1)!}-\sum_{x=0}^{+\infty}g(x)\frac{\theta^{x}}{x!}\right\} \\
&= e^{-\theta}\left\{\sum_{x-1=0}^{+\infty}g(x)\frac{\theta^{x-1}}{(x-1)!}-\sum_{x=0}^{+\infty}g(x)\frac{\theta^{x}}{x!}\right\} \\
&= e^{-\theta}\left\{\sum_{x=0}^{+\infty}g(x+1)\frac{\theta^{x}}{x!}-\sum_{x=0}^{+\infty}g(x)\frac{\theta^{x}}{x!}\right\} \\
&= e^{-\theta}\sum_{x=0}^{+\infty}\frac{\theta^{x}}{x!}[g(x+1)-g(x)] \\
&= \sum_{x=0}^{+\infty}\frac{\theta^{x}e^{-\theta}}{x!}[g(x+1)-g(x)] \\
&= E[\Delta g(X)], \qquad (3.33)
\end{aligned}
$$

όπου $\Delta g(x)=g(x+1)-g(x)$ δηλώνει την διαφορά πρώτης τάξεως προς το εμπρός στο σημείο x και

$$I_x^F(\theta)=\frac{1}{\theta}=\frac{1}{\lambda} \qquad (3.34)$$

Από εφαρμογή του Λήμματος 3.1 και τις σχέσεις (3.35) και (3.36) προκύπτει:

$$Var[g(X)]\geq\lambda E^2[\Delta g(X)]. \qquad (3.35)$$

Παράδειγμα 3.3

Έστω $X \sim P(\lambda)$ και $g(X) = X^2$.

$\Delta g(x) = g(x+1) - g(x) = (x+1)^2 - x^2 = 2x+1$

$E[\Delta g(X)] = E(2X+1) = 2E(X)+1 = 2\lambda+1$.

Από την άμεση εφαρμογή της σχέσης (3.37) προκύπτει:

$$Var[g(X)] \geq \lambda(2\lambda+1)^2.$$

Θεωρώντας, τώρα, τυχαίο διάνυσμα $\mathbf{X} = (X_1, \ldots, X_n)$ με τις ανεξάρτητες τ.μ. $X_i \sim P(\theta)$ το κάτω φράγμα για την διακύμανση μιας συνάρτησης $g(\mathbf{X})$ έχει ως εξής:

$$Var[g(\mathbf{X})] \geq \frac{\lambda}{n} E^2 \sum_{i=1}^{n} \Delta_\kappa g(\mathbf{X}), \qquad (3.36)$$

όπου το Δ_κ είναι η διαφορά πρώτης τάξης προς τα εμπρός στο σημείο x_κ, δηλαδή

$$\Delta_\kappa g(\mathbf{X}) = g(x_1, \ldots, x_\kappa + 1, x_{\kappa+1}, \ldots, x_n) - g(x_1, \ldots, x_\kappa, x_{\kappa+1}, \ldots, x_n).$$

Αν όμως τα X_i είναι ανεξάρτητες τ.μ. αλλά ακολουθούν κατανομή Poisson με διαφορετική παράμετρο λ_i το καθένα τότε το κάτω φράγμα έχει ως εξής:

$$Var[g(\mathbf{X})] \geq \sum_{i=1}^{n} \lambda_i E^2 \left[\Delta_\kappa g(\mathbf{X})\right], \qquad (3.37)$$

όπου η ανισότητα γίνεται ισότητα αν και μόνον αν η $g(\mathbf{X})$ είναι γραμμική.

3.3 Ασυμπτωτική Κατανομή Ε.Μ.Π.

Είναι γνωστό το πόσο σημαντικοί είναι οι εκτιμητές μέγιστης πιθανοφάνειας (ΕΜΠ) στη στατιστική συμπερασματολογία. Οι ΕΜΠ έχουν κάποιες σημαντικές ιδιότητες, μια εκ των οποίων είναι η ασυμπτωτική κατανομή τους. Το ιδιαίτερο χαρακτηριστικό αυτής της ιδιότητας είναι ότι η ασυμπτωτική κατανομή δε δίνεται σε όρους της κατανομής του ΕΜΠ αλλά σε όρους της σ.π.π. της κατανομής του δείγματος. Συγκεκριμένα, η διακύμανση αυτής της ασυμπτωτικής κατανομής είναι το κάτω φράγμα της ανισότητας Cramer-Rao. Σε αυτό το τελευταίο χαρακτηριστικό θα δοθεί έμφαση καθώς εμπλέκεται η ανισότητα Cramer – Rao, που αποτελεί αντικείμενο της παρούσας διατριβής.

Μονοδιάστατη περίπτωση

Έστω τ.μ. X με σ.π.π. $f(x,\theta)$, $\theta \in \Theta$ όπου Θ ένα ανοιχτό διάστημα στο $\mathbb{R}$ και έστω τυχαίο δείγμα $X_1,\ldots,X_n$ από την σ.π.π. $f(x,\theta)$.

Ορίζουμε τις παρακάτω συνθήκες ομαλότητας :

i) Για όλα τα $x \in \mathbb{R}$ υπάρχουν οι παράγωγοι:

$$\frac{\partial}{\partial\theta}\log f(x,\theta)\,,\ \frac{\partial^2}{\partial\theta^2}\log f(x,\theta)\,,\ \frac{\partial^3}{\partial\theta^3}\log f(x,\theta),\ \theta \in \Theta$$

ii) Υπάρχει μια μετρήσιμη συνάρτηση $H:\mathbb{R} \to (0,+\infty)$ τέτοια ώστε για κάθε $\theta \in \Theta$:

α) $\left|\dfrac{\partial^3}{\partial\theta^3}\log f(x,\theta)\right| < H(x)$

β) $E[H(x)] < M \quad (<\infty)$

και το Μ είναι ανεξάρτητο του $\theta \in \Theta$.

iii) $E\left[\dfrac{\partial}{\partial\theta}\log f(x,\theta)\right] = 0$, για κάθε $\theta \in \Theta$.

iv) Για κάθε $\theta \in \Theta$

α) $-E\left[\dfrac{\partial^2}{\partial\theta^2}\log f(x,\theta)\right] = E\left[\dfrac{\partial}{\partial\theta}\log f(x,\theta)\right]^2 = I_x^F(\theta)$

όπου $I_x^F(\theta)$ το μέτρο πληροφορίας του Fisher.

β) $0 < I_x^F(\theta) < \infty$.

(Όσον αφορά τις παραπάνω συνθήκες ομαλότητας και το θεώρημα που ακολουθεί βλέπε Rohatgi et Saleh (2001))

Θεώρημα 3.1

Έστω ότι ισχύουν οι παραπάνω συνθήκες ομαλότητας και επιπλέον για κάθε n υπάρχει ένας μονοσήμαντα ορισμένος ΕΜΠ $\hat{\theta}$ της θ. Τότε ο $\hat{\theta}$ είναι

(i) συνεπής εκτιμητής για το θ

(ii) ασυμπτωτικά κανονικός εκτιμητής για το θ, δηλαδή

$$\sqrt{n}(\hat{\theta}-\theta) \xrightarrow{\kappa} \mathrm{N}\left(0, \frac{1}{I_x^F(\theta)}\right),$$

από την οποία προκύπτει ότι

$$\hat{\theta} \xrightarrow{\kappa} \mathrm{N}\left(\theta, \frac{1}{nI_x^F(\theta)}\right) \qquad (3.38)$$

όπου, το $\frac{1}{nI_x^F(\theta)}$ είναι το κάτω φράγμα της ανισότητας Cramer-Rao για την παράμετρο θ με $E\left(\hat{\theta}\right)=\theta$.

Στα περισσότερα προβλήματα όμως απαιτείται η εκτίμηση κάποιας συνάρτησης της θ, έστω $g(\theta)$. Στην περίπτωση αυτή ισχύει το παρακάτω.

Θεώρημα 3.2 (ΕΜΠ συνάρτησης της θ)

Έστω $X_1,\ldots,X_n$ τ.δ. με σ.π.π. $f(x,\boldsymbol{\theta})$ ή σ.π. $p(x,\boldsymbol{\theta})$, $\boldsymbol{\theta} \in \Theta \subseteq \mathbb{R}^\kappa$ και $\hat{\boldsymbol{\theta}} = \left(\hat{\theta}_1,\ldots,\hat{\theta}_\kappa\right)'$ ο ΕΜΠ της $\boldsymbol{\theta}=(\theta_1,\ldots,\theta_\kappa)'$ όπου $\hat{\theta}_j = \hat{\theta}_j\left(X_1,\ldots,X_n\right)$.

Εάν $\mathbf{g}(\boldsymbol{\theta})=(g_1(\boldsymbol{\theta}),\ldots,g_r(\boldsymbol{\theta}))'$ για $1 \le r \le \kappa$ είναι ένας μετασχηματισμός του παραμετρικού χώρου Θ, τότε ο ΕΜΠ της $g(\boldsymbol{\theta})$ είναι $\mathbf{g}(\hat{\boldsymbol{\theta}})=(g_1(\hat{\boldsymbol{\theta}}),\ldots,g_r(\hat{\boldsymbol{\theta}}))$, όπου $\phi\left(\hat{\theta}_1,\ldots,\hat{\theta}_\kappa\right)$ είναι ο ΕΜΠ της $\phi_j\left(\theta_1,\ldots,\theta_\kappa\right)$, $j=1,\ldots,r$.

Θεωρώντας ότι η g είναι παραγωγίσιμη αποδεικνύεται ότι η $g(\hat{\theta})$ έχει ασυμπτωτική κανονική κατανομή με μέση τιμή $g(\theta)$ και διακύμανση ίση με $\frac{[g'(\theta)]^2}{nI_x^F(\theta)}$, δηλαδή

$$\hat{g}(\theta) = g(\hat{\theta}) \xrightarrow{\kappa} \mathrm{N}\left(g(\theta), \frac{[g'(\theta)]^2}{nI_x^F(\theta)} \right), \tag{3.39}$$

όπου όπως γνωρίζουμε, η διακύμανση στη σχέση (3.42) ισούται με το κάτω φράγμα της ανισότητας Cramer-Rao για την διακύμανση ενός αμερόληπτου εκτιμητή της συνάρτησης $g(\theta)$. (βλέπε Παπαιωάννου-Φερεντίνος(2000), Mood et al(1974))

Συνοψίζοντας, λοιπόν, τα παραπάνω, καταλήγουμε στο συμπέρασμα ότι γνωρίζοντας το κάτω φράγμα της ανισότητας Cramer-Rao για την διακύμανση της ασυμπτωτικής κατανομής ενός αμερόληπτου εκτιμητή είμαστε σε θέση να προσδιορίσουμε και την διακύμανση της ασυμπτωτικής κατανομής ενός εκτιμητή μεγίστης πιθανοφάνειας. Αυτό θα φανεί καλύτερα και στα παραδείγματα που ακολουθούν.

Παράδειγμα 3.4

Έστω X τ.μ. με σ.π.π.

$$f(x,\theta) = \frac{x}{\theta} e^{-x^2/2\theta}. \tag{3.40}$$

Ο ΕΜΠ της παραμέτρου θ για $n = 1$ είναι ο

$$\hat{\theta} = \frac{X^2}{2} \tag{3.41}$$

Το κάτω φράγμα της ανισότητας Cramer-Rao για την διακύμανση ενός αμερόληπτου εκτιμητή της θ δίνεται :

$$\frac{1}{I_x^F(\theta)} = \frac{1}{\theta^2}. \tag{3.42}$$

Επομένως σύμφωνα και με τα παραπάνω, η κατανομή του $\hat{\theta}$ προσεγγίζεται ως εξής:

$$\hat{\theta} \overset{\pi\rho o\sigma.}{\sim} N(\theta, \frac{1}{I_x^F(\theta)})$$

$$\hat{\theta} \overset{\pi\rho o\sigma.}{\sim} N(\theta, \theta^2). \tag{3.43}$$

Έστω τώρα, ότι το ζητούμενο ήταν η ασυμπτωτική κατανομή μιας συνάρτησης του ΕΜΠ $\hat{\theta}$.

Έστω από αυτή την συνάρτηση $g(\theta)=\theta^2$, διαφορίσιμη. Από το θεώρημα (3.2) έχουμε:

$$\hat{g}(\theta)=g(\hat{\theta})=\hat{\theta}^2=\frac{\mathrm{X}^4}{4}. \tag{3.44}$$

Το κάτω φράγμα της ανισότητας Cramer – Rao δίνεται από τη σχέση:

$$\frac{[g'(\theta)]^2}{I_x^F(\theta)}=\frac{4\theta^2}{\frac{1}{\theta^2}}=4\theta^4 . \tag{3.45}$$

Επομένως η ασυμπτωτική κατανομή της $\hat{g}(\theta)$ έχει ως εξής:

$$g(\hat{\theta}) \overset{\pi\rho o\sigma.}{\sim} N\left(g(\theta),\frac{[g'(\theta)]^2}{I_x^F(\theta)}\right)$$

$$g(\hat{\theta}) \overset{\pi\rho o\sigma.}{\sim} N(\theta^2,4\theta^4). \tag{3.46}$$

Παράδειγμα 3.5

Έστω $X_1,\ldots,X_n$ ανεξάρτητες τ.μ. με σ.π.π.

$$f(x,\theta)=\theta e^{-\theta x} \quad (x>0,\theta>0). \tag{3.47}$$

Ο ΕΜΠ της θ δίνεται από την σχέση

$$\hat{\theta}=n\frac{1}{\sum_{i=1}^{n}X_i}=\frac{1}{\overline{\mathrm{X}}}. \tag{3.48}$$

και

$$I_x^F(\theta)=\frac{1}{\theta^2}. \tag{3.49}$$

Άρα η κατανομή του $\hat{\theta}$ προσεγγίζεται ως εξής:

$$\hat{\theta} \overset{\pi\rho o\sigma.}{\sim} N(\theta,\frac{1}{nI_x^F(\theta)})$$

$$\hat{\theta} \overset{\pi\rho o\sigma.}{\sim} N(\theta,\frac{\theta^2}{n}). \tag{3.50}$$

Έστω η διαφορίσιμη συνάρτηση $g(\theta)=\theta^2+\theta$ και ζητούμενο είναι η ασυμπτωτική κατανομή του ΕΜΠ της $g(\theta)$. Δουλεύοντας ανάλογα με το προηγούμενο παράδειγμα προκύπτει :

$$\hat{g(\theta)}=g(\hat{\theta})=\hat{\theta}^2+\hat{\theta}=\left(\frac{1}{\overline{\overline{X}}}\right)^2+\frac{1}{\overline{\overline{X}}}. \tag{3.51}$$

Το κάτω φράγμα της ανισότητας Cramer-Rao έχει ως εξής:

$$\frac{[g'(\theta)]}{nI_x^F(\theta)}=\frac{(2\theta+1)^2}{\frac{n}{\theta^2}}=\frac{1}{n}\theta^2(2\theta+1)^2. \tag{3.52}$$

Άρα

$$g(\hat{\theta})\overset{\pi\rho o\sigma.}{\sim}N\left(g(\theta),\frac{[g'(\theta)]^2}{nI_x^F(\theta)}\right)$$

$$g(\hat{\theta})\overset{\pi\rho o\sigma.}{\sim}N\left(\theta^2+\theta,\frac{1}{n}(2\theta+1)^2\right). \tag{3.53}$$

Παράδειγμα 3.6

Έστω $X_1,\ldots,X_n$ ανεξάρτητες τ.μ. από κατανομή $N\left(-\sigma^2/2,\sigma^2\right)$.

Αρχικά θα βρεθεί ο ΕΜΠ της σ^2 για αυτό θέτουμε $\theta=\sigma^2$ και έχουμε τα εξής:

$$f(x_i,\theta)=\frac{1}{\sqrt{2\pi\theta}}e^{-\frac{1}{2\theta}\left(1+\frac{\theta}{2}\right)^2}. \tag{3.54}$$

Επομένως η συνάρτηση πιθανοφάνειας είναι:

$$L(\theta|\mathbf{x})=\prod_{i=1}^{n}f(x_i,\theta)=\left(\frac{1}{\sqrt{2\pi\theta}}\right)^n e^{-\frac{1}{2\theta}\sum_{i=1}^{n}\left(x_i+\frac{\theta}{2}\right)^2}. \tag{3.55}$$

Έχουμε ότι $\frac{\partial \ln L}{\partial\theta}=0$ και μετά από πράξεις καταλήγουμε στην εξίσωση $\theta^2+4\theta-\frac{4}{n}\sum_{i=1}^{n}X_i^2=0$

από όπου βρίσκεται ο ΕΜΠ της $\theta=\sigma^2$.

Έστω $\hat{\theta}$ ο ΕΜΠ της $\theta=\sigma^2$ τότε

$$I_x^F(\theta)=\frac{1}{2\theta^2}+\frac{1}{4\theta}=\frac{2+\theta}{4\theta^2} \tag{3.56}$$

και επομένως η κατανομή της $\hat{\sigma}^2$ προσεγγίζεται ως εξής:

$$\hat{\sigma}^2 \overset{\pi\rho o\sigma.}{\sim} N\left(\sigma^2, \frac{1}{nI_x^F(\theta)}\right)$$

$$\hat{\sigma}^2 \overset{\pi\rho o\sigma.}{\sim} N\left(\sigma^2, \frac{4\sigma^4}{n(2+\sigma^2)}\right). \tag{3.57}$$

Έστω τώρα η διαφορίσιμη συνάρτηση $g(\sigma^2)=\sigma^4-2\sigma^2$ και ζητούμενο είναι η ασυμπτωτική κατανομή του ΕΜΠ της $g(\sigma^2)$. Εργαζόμενοι ανάλογα με τα προηγούμενα παραδείγματα προκύπτει:

$$\widehat{g(\sigma^2)}=g(\hat{\sigma}^2)=\hat{\sigma}^4-2\hat{\sigma}^2,\ \ (g(\theta)=(\theta^2-2\theta)). \tag{3.58}$$

Το κάτω φράγμα της ανισότητας Cramer-Rao έχει ως εξής:

$$\frac{[g'(\theta)]}{nI_x^F(\theta)}=\frac{(2\theta-2)^2}{n\left(\frac{2+\theta}{4\theta^2}\right)}=\frac{1}{n}\frac{4\theta^2(2\theta-2)^2}{2+\theta},\left(\theta=\sigma^2\right) \tag{3.59}$$

Έτσι

$$g(\hat{\theta}) \overset{\pi\rho o\sigma.}{\sim} N\left(g(\theta), \frac{[g'(\theta)]^2}{nI_x^F(\theta)}\right)$$

ή

$$g(\hat{\theta}) \overset{\pi\rho o\sigma.}{\sim} N\left(\theta^2-2\theta, \frac{1}{n}\frac{4\theta^2(2\theta-2)^2}{(2+\theta)}\right). \tag{3.60}$$

Πολυδιάστατη περίπτωση

Οι εκτιμητές μεγίστης πιθανοφάνειας έχουν παρόμοιες βέλτιστες ιδιότητες για μεγάλα δείγματα και στην περίπτωση που η παράμετρος είναι πολυδιάστατη. Μόνο που τώρα η ασυμπτωτική κατανομή είναι η πολυδιάστατη κανονική με διακύμανση ίση με το κάτω φράγμα της ανισότητας Cramer-Rao για την πολυδιάστατη περίπτωση.

Έστω η παράμετρος $\boldsymbol{\theta}=(\theta_1,\ldots,\theta_\kappa)'$ και $\mathbf{x}_1,\ldots,\mathbf{x}_n$ ανεξάρτητα τυχαία διανύσματα με σ.π.π. $f(\mathbf{x},\boldsymbol{\theta})$. Η συνάρτηση πιθανοφάνειάς τους, δηλαδή η από κοινού σ.π.π. τους θα συμβολίζεται με $L(\boldsymbol{\theta}\mid\mathbf{x})$. Ορίζουμε τις παρακάτω συνθήκες ομαλότητας:

i) $\log(f(\mathbf{x},\boldsymbol{\theta}))$ μετρήσιμη ως προς το $\mathbf{x}$ για κάθε $\boldsymbol{\theta}\in\Theta$, όπου $\Theta\subseteq\mathbb{R}^\kappa$.

ii) Υπάρχει μια $\mathbf{x}$ μετρήσιμη συνάρτηση $H(\mathbf{x})$ τέτοια ώστε $|\log(f(\mathbf{x},\boldsymbol{\theta}))|\le H(\mathbf{x})$ όπου $\int\limits_{\mathbf{x}} H(\mathbf{x})f(\mathbf{x},\boldsymbol{\theta})d\mathbf{x}<\infty$.

iii) Υπάρχει ένα ανοιχτό υποσύνολο $V_\varepsilon(\boldsymbol{\theta}_0)$ του Θ που περιέχει την πραγματική τιμή $\boldsymbol{\theta}_0$ της παραμέτρου $\boldsymbol{\theta}$ έτσι ώστε για σχεδόν όλα τα $\mathbf{x}$, η σ.π.π. έχει όλες της πρώτης και δεύτερης τάξης μερικές παραγώγους ως προς τις συνιστώσες του $\boldsymbol{\theta}\in V_\varepsilon(\boldsymbol{\theta}_0)$.

iv) Οι πρώτης τάξης μερικές παράγωγοι της $\log(f(\mathbf{x},\boldsymbol{\theta}))$ ικανοποιούν την σχέση

$$E\left(\frac{\partial\log(f(\mathbf{X},\boldsymbol{\theta}))}{\partial\boldsymbol{\theta}}\right)=\frac{\partial}{\partial\boldsymbol{\theta}}\mathrm{E}\left(\log(f(\mathbf{X},\boldsymbol{\theta}))\right)=0^{\kappa x1}\quad, \tag{3.61}$$

όπου $0^{\kappa x1}$ ο μηδενικός πίνακας και ο θετικά ορισμένος πίνακας πληροφορίας του Fisher ικανοποιεί τη σχέση

$$I_{\mathbf{x}}^{F}(\boldsymbol{\theta})=-E\left(\frac{\partial^2\log(f(\mathbf{X},\boldsymbol{\theta}))}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'}\right)=E\left(\frac{\partial\log(f(\mathbf{X},\boldsymbol{\theta}))}{\partial\boldsymbol{\theta}}\frac{\partial\log(f(\mathbf{X},\boldsymbol{\theta}))}{\partial\boldsymbol{\theta}'}\right) \tag{3.62}$$

v) Για $\boldsymbol{\theta}_0\in\Theta$ υπάρχει μια $\mathbf{x}$ μετρήσιμη συνάρτηση $K(\mathbf{x})$ τέτοια ώστε

$$\left\|\left.\frac{\partial^2\log(f(\mathbf{X},\boldsymbol{\theta}))}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'}\right|_{\boldsymbol{\theta}_0}\right\|\le K(\mathbf{x}),$$

$$\int_{\mathbf{x}} K(\mathbf{x}) f(\mathbf{X},\boldsymbol{\theta}) d\mathbf{x} < \infty$$

και

$$\frac{\partial^2 \log(f(\mathbf{X},\boldsymbol{\theta}))}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'}$$

είναι μετρήσιμη ως προς $\mathbf{x}$.(βλέπε Klotz (2004))

Θεώρημα 3.3

Έστω $\mathbf{x}_1,\ldots,\mathbf{x}_n$ ανεξάρτητα τυχαία διανύσματα με σ.π.π. $f(\mathbf{x},\boldsymbol{\theta})$, $\boldsymbol{\theta} \in \Theta$, και ικανοποιούνται οι συνθήκες ομαλότητας (i) έως (iv) με Θ να είναι ανοιχτό υποσύνολο του $\mathbb{R}^{\kappa}$. Αν ο εκτιμητής μεγίστης πιθανοφάνειας της $\boldsymbol{\theta} = (\theta_1,\ldots,\theta_\kappa)'$ είναι ο $\hat{\boldsymbol{\theta}} = \left(\hat{\theta}_1,\hat{\theta}_2,\ldots,\hat{\theta}_\kappa\right)'$, τότε, καθώς $n \to \infty$

$$\sqrt{n}\left(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}\right) \xrightarrow{\kappa} Z \sim N_\kappa\left(\mathbf{0}_{\kappa x1},\left(I_{\mathbf{x}}^F(\boldsymbol{\theta})\right)^{-1}\right) \qquad (3.63)$$

όπου $\mathbf{0}_{\kappa x1}$ είναι ο μηδενικός πίνακας.

Από τη σχέση αυτή προκύπτει ότι

$$\hat{\boldsymbol{\theta}} \xrightarrow{\kappa} Z \sim N_\kappa\left(\boldsymbol{\theta},\frac{1}{n}\left(I_{\mathbf{x}}^F(\boldsymbol{\theta})\right)^{-1}\right), \qquad (3.64)$$

όπου το $\left(I_{\mathbf{x}}^F(\boldsymbol{\theta})\right)^{-1}$ είναι ο αντίστροφός του πίνακα πληροφορίας του Fisher.

Όπως όμως και στην μονοδιάστατη περίπτωση συνήθως ζητούμενο είναι ο ΕΜΠ όχι της παραμέτρου $\boldsymbol{\theta}$ αλλά κάποιας συνάρτησης, έστω $\mathbf{g}(\boldsymbol{\theta})$, αυτής. Ισχύουν αντίστοιχα θεωρήματα με την μονοδιάστατη περίπτωση.

Θεώρημα 3.4

Αν $\hat{\boldsymbol{\theta}}$ είναι ο ΕΜΠ της $\boldsymbol{\theta}$ τότε για κάθε συνάρτηση $\mathbf{g}(\boldsymbol{\theta})$, ο ΕΜΠ της $\mathbf{g}(\boldsymbol{\theta})$ είναι $\mathbf{g}(\hat{\boldsymbol{\theta}})$.

Θεώρημα 3.5

Έστω ότι ισχύουν οι προϋποθέσεις του Θεωρήματος 3.3 και

$$\sqrt{n}(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}) \xrightarrow{\kappa} Z \sim N_\kappa\left(0,\left(I_{\mathbf{x}}^F(\boldsymbol{\theta})\right)^{-1}\right) \qquad (3.65)$$

Τότε, αν η $\mathbf{g}=(g_1,\ldots,g_r)'$ έχει συνεχείς παραγώγους όπου $r \le \kappa$, έχουμε

$$\sqrt{n}\left(\mathbf{g}(\hat{\boldsymbol{\theta}})-\mathbf{g}(\boldsymbol{\theta})\right) \xrightarrow{\kappa} W \sim N_r\left(0,\boldsymbol{\Lambda}(\boldsymbol{\theta})\right), \qquad (3.66)$$

όπου

$$\boldsymbol{\Lambda}(\boldsymbol{\theta})^{r x r} = \mathbf{G}\left(I_{\mathbf{x}}^F(\boldsymbol{\theta})\right)^{-1}\mathbf{G}' \qquad (3.67)$$

με

$$\mathbf{G}=\left\|\frac{\partial g_i(\boldsymbol{\theta})}{\partial \theta_j}\right\|_{r x \kappa}, \qquad i=1,\ldots,r \ , \ j=1,\ldots,\kappa .$$

Όπως γνωρίζουμε το κάτω φράγμα της ανισότητας Cramer-Rao για την διακύμανση αμερόληπτου εκτιμητή της $\mathbf{g}(\boldsymbol{\theta})$ είναι ο πίνακας $\boldsymbol{\Lambda}(\boldsymbol{\theta})$ όπως ορίζεται στη σχέση (3.67).

Παράδειγμα 3.7

Έστω

$$\mathbf{X}_i=\begin{pmatrix} X_{i1} \\ X_{i2} \end{pmatrix} \sim N_2\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix},\begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}\right)$$

για $i=1,2,\ldots,n$ να είναι ανεξάρτητα τυχαία διανύσματα από την κανονική διδιάστατη κατανομή με $\boldsymbol{\theta}=(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)' \in \Theta$ όπου :

$$\Theta=\left\{\left(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho\right)' : -\infty<\alpha_{j1}<\mu_j<\alpha_{j2}<\infty, 0<b_{j1}<\sigma_j^2<b_{j2}<\infty, -1<c_1<\rho<c_2<1, j=1,2\right\}$$

Ο εκτιμητής μεγίστης πιθανοφάνειας είναι ο

$$\hat{\boldsymbol{\theta}} = \begin{pmatrix} \hat{\mu}_1 \\ \hat{\mu}_2 \\ \hat{\sigma}_1^2 \\ \hat{\sigma}_2^2 \\ \hat{p} \end{pmatrix} = \begin{pmatrix} \frac{1}{n}\sum_{i=1}^{n} X_{i1} \\ \frac{1}{n}\sum_{i=1}^{n} X_{i2} \\ \frac{1}{n}\sum_{i=1}^{n}\left(X_{i1}-\overline{X_1}\right)^2 \\ \frac{1}{n}\sum_{i=1}^{n}\left(X_{i2}-\overline{X_2}\right)^2 \\ \frac{1}{n\hat{\sigma}_1\hat{\sigma}_2}\sum_{i=1}^{n}\left(X_{i1}-\overline{X_1}\right)\left(X_{i2}-\overline{X_2}\right) \end{pmatrix}$$

όπου $\overline{\mathrm{X}}_j = \frac{1}{n}\sum_{i=1}^{n} X_{ij}$, $j=1,2$.

Ο ΕΜΠ ,καθώς το $n \to \infty$, ικανοποιεί τη σχέση

$$\sqrt{n}(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}) \xrightarrow{\kappa} Z \sim N_5\left(0,\left(I_{\mathbf{x}}^F(\boldsymbol{\theta})\right)^{-1}\right) \tag{3.68}$$

ή

$$\hat{\boldsymbol{\theta}} \xrightarrow{\kappa} Z \sim N_5\left(\boldsymbol{\theta},\frac{1}{n}\left(I_{\mathbf{x}}^F(\boldsymbol{\theta})\right)^{-1}\right).$$

Με

$$I_{\mathbf{x}}^F(\boldsymbol{\theta}) = \frac{1}{\left(1-p^2\right)}\begin{pmatrix} 1/\sigma_1^2 & -\rho/(\sigma_1\sigma_2) & 0 & 0 & 0 \\ -\rho/(\sigma_1\sigma_2) & 1/\sigma_2^2 & 0 & 0 & 0 \\ 0 & 0 & (2-\rho^2)/4(\sigma_1^4) & -\rho^2/4(\sigma_1^2\sigma_2^2) & -\rho/2(\sigma_1^2) \\ 0 & 0 & -\rho^2/4(\sigma_1^2\sigma_2^2) & (2-\rho^2)/4(\sigma_2^4) & -\rho/2(\sigma_2^2) \\ 0 & 0 & -\rho/2(\sigma_1^2) & -\rho/2(\sigma_2^2) & (1+\rho^2)/(1-\rho^2) \end{pmatrix}$$

$$\left(I_{\mathbf{x}}^F(\boldsymbol{\theta})\right)^{-1} = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 & 0 & 0 & 0 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 & 0 & 0 & 0 \\ 0 & 0 & 2\sigma_1^4 & 2\rho^2\sigma^2{}_1\sigma^2{}_2 & \rho(1-\rho^2)\sigma_1^2 \\ 0 & 0 & 2\rho^2\sigma^2{}_1\sigma^2{}_2 & 2\sigma_2^4 & \rho(1-\rho^2)\sigma_2^2 \\ 0 & 0 & \rho(1-\rho^2)\sigma_1^2 & \rho(1-\rho^2)\sigma_2^2 & (1-\rho^2)^2 \end{pmatrix}$$

Αυτό το παράδειγμα αποτελεί εφαρμογή του Θεωρήματος 3.3. Όσον αφορά το Θεώρημα 3.5, εφαρμογή του είναι το παράδειγμα που ακολουθεί.(βλέπε Klotz (2004))

Παράδειγμα 3.8

Έστω $X_1,\dots,X_n$ ανεξάρτητες τ.μ από $N(\mu,\sigma^2)$.

Θεωρούμε $\boldsymbol{\theta}=(\mu,\sigma^2)'$.

Ο ΕΜΠ δίνεται ως εξής:

$$\hat{\boldsymbol{\theta}}=\begin{pmatrix} \overline{\mathrm{X}} \\ \frac{1}{n}\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2 \end{pmatrix}=\begin{pmatrix} \overline{X} \\ S'^2 \end{pmatrix} \tag{3.69}$$

και ικανοποιείται η σχέση

$$\sqrt{n}(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta})\xrightarrow{\kappa} Z\sim N_2\left(\begin{pmatrix}0\\0\end{pmatrix},\left(I_{\mathbf{x}}^F(\boldsymbol{\theta})\right)^{-1}\right)$$

ή

$$\sqrt{n}\begin{pmatrix} \overline{\mathrm{X}}-\mu \\ \frac{1}{n}\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2-\sigma^2 \end{pmatrix}\xrightarrow{\kappa} Z\sim N_2\left(\begin{pmatrix}0\\0\end{pmatrix},\begin{pmatrix}\sigma^2 & 0\\ 0 & 2\sigma^4\end{pmatrix}\right), \tag{3.70}$$

όπου

$$\overline{\mathrm{X}}=\frac{1}{n}\sum_{i=1}^{n}\mathrm{X}_i \quad \text{και} \quad I_{\mathbf{x}}^F(\boldsymbol{\theta})=\begin{pmatrix}\frac{1}{\sigma^2} & 0\\ 0 & \frac{1}{2\sigma^4}\end{pmatrix}.$$

Άρα

$$\hat{\boldsymbol{\theta}}\xrightarrow{\kappa} Z\sim N_2\left(\begin{pmatrix}\mu\\ \sigma^2\end{pmatrix},\begin{pmatrix}\frac{\sigma^2}{n}\\ \frac{2\sigma^4}{n}\end{pmatrix}\right). \tag{3.71}$$

Αν είχαμε $g(\boldsymbol{\theta})=\frac{\mu}{\sigma}$ προκύπτει:

$$\hat{g(\boldsymbol{\theta})}=\frac{\overline{X}}{\frac{1}{n}\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^{1/2}} \tag{3.72}$$

Με εφαρμογή του Θεωρήματος 3.5

$$\sqrt{n}(g(\hat{\boldsymbol{\theta}})-g(\boldsymbol{\theta}))\xrightarrow{K}W\sim N_1\left(0,\boldsymbol{\Lambda}(\boldsymbol{\theta})\right) \tag{3.73}$$

ή

$$\hat{g}(\boldsymbol{\theta})=g(\hat{\boldsymbol{\theta}})\xrightarrow{K}W\sim N_1\left(g(\boldsymbol{\theta}),\frac{1}{n}\left(1+\frac{\mu^2}{2\sigma^2}\right)\right),$$

όπου $\boldsymbol{\Lambda}(\boldsymbol{\theta})$ να είναι 1x1 καθώς

$$\boldsymbol{\Lambda}(\boldsymbol{\theta})=\begin{pmatrix}\frac{1}{\sigma} & \frac{-\mu}{2\sigma^3}\end{pmatrix}_{1x2}\begin{pmatrix}\sigma^2 & 0\\ 0 & 2\sigma^4\end{pmatrix}_{2x2}\begin{pmatrix}\frac{1}{\sigma}\\ \frac{-\mu}{2\sigma 3}\end{pmatrix}_{2x1}=\left(1+\frac{\mu^2}{2\sigma^2}\right)_{1x1}. \tag{3.74}$$

για τον οποίο γνωρίζουμε ότι είναι το κάτω φράγμα της ανισότητας Cramer-Rao για την διακύμανση ενός αμερόληπτου εκτιμητή της συνάρτησης $\hat{g}(\boldsymbol{\theta})$.(βλέπε Klotz (2004))

<u>Παράδειγμα 3.9 (Πολυωνυμική Κατανομή)</u>

Έστω $\mathbf{X}=\left(X_1,......,X_K\right)$ με από κοινού σ.π.π να δίνεται ως εξής

$$P\left(X=x\right)=\frac{n!}{x_1!x_2!......x_K!}p_1^{x_1}p_2^{x_2}.........p_K^{x_K},$$

όπου $\sum_{i=1}^{K}x_i=n, p_i>0, \sum_{i=1}^{K}p_i=1$.

Αποδεικνύεται ότι ο ΕΜΠ για κάθε p_i είναι:

$$\hat{p}_i=\frac{X_i}{n} \text{ με } i=1,...,K.$$

Ο ΕΜΠ, καθώς το $n\rightarrow\infty$, ικανοποιεί τη σχέση

$$\sqrt{n}(\hat{\mathbf{p}}-\mathbf{p})\xrightarrow{K}Z\sim N_K\left(0,\left(I_{\mathbf{x}}^F(\mathbf{p})\right)^{-1}\right) \tag{3.75}$$

με $\hat{\mathbf{p}}=\left(\hat{p}_1,...,\hat{p}_K\right)'$ και $\mathbf{p}=\left(p_1,...,p_K\right)'$.

Ο αντίστροφος του πίνακα πληροφορίας του Fisher δίνεται από τη σχέση:

$$\left(I_{\mathbf{x}}^{F}(\mathbf{p})\right)^{-1}=\begin{bmatrix} p_1(1-p_1) & -p_1p_2 & \cdots & -p_1p_{K-1} \\ -p_1p_2 & p_2(1-p_2) & \cdots & -p_2p_{K-1} \\ \vdots & \vdots & \ddots & \vdots \\ -p_{K-1}p_1 & -p_{K-1}p_2 & \cdots & p_{K-1}(1-p_{K-1}) \end{bmatrix}$$

Άρα

$$\hat{\mathbf{p}}\xrightarrow{\kappa} Z \sim N_K\left(\begin{pmatrix} p_1 \\ \vdots \\ \vdots \\ p_K \end{pmatrix},\begin{pmatrix} \frac{p_1(1-p_1)}{n} \\ \vdots \\ \vdots \\ \frac{p_{K-1}(1-p_{K-1})}{n} \end{pmatrix}\right).$$

(βλέπε Klotz (2004))

Παράδειγμα 3.10

Έστω οι ανεξάρτητες τ.μ με $X \sim P(\lambda), Y_1 \sim P(\lambda_1), Y_2 \sim P(\lambda_2)$ όπου $\lambda_1+\lambda_2=\lambda$.

Η από κοινού σ.π.π δίνεται ως εξής

$$P(X=x,Y_1=y_1,Y_2=y_2)=\frac{e^{-\lambda}\lambda^{x}}{x!}\frac{e^{-\lambda_1}\lambda_1^{y_1}}{y_1!}\frac{e^{-\lambda_2}\lambda_2^{y_2}}{y_2!}.$$

Θέτοντας $\lambda_1+\lambda_2=\lambda$ έχουμε

$$P(X=x,Y_1=y_1,Y_2=y_2)=\frac{e^{-(\lambda_1+\lambda_2)}(\lambda_1+\lambda_2)^{x}}{x!}\frac{e^{-\lambda_1}\lambda_1^{y_1}}{y_1!}\frac{e^{-\lambda_2}\lambda_2^{y_2}}{y_2!}.$$

Ο λογάριθμος της συνάρτησης πιθανοφάνειας είναι

$$\log P(X=x,Y_1=y_1,Y_2=y_2)=-2\lambda_1-2\lambda_2+x\log(\lambda_1+\lambda_2)+y_1\log\lambda_1+y_2\log\lambda_2-\log(x!y_1!y_2!)$$

Παραγωγίζοντας και εξισώνοντας με το 0 προκύπτει

$$\frac{\partial \log P(X=x, Y_1=y_1, Y_2=y_2)}{\partial \lambda_1} = -2 + \frac{x}{\lambda_1+\lambda_2} + \frac{y_1}{\lambda_1} = 0 \tag{3.76}$$

ή

$$\frac{\partial \log P(X=x, Y_1=y_1, Y_2=y_2)}{\partial \lambda_2} = -2 + \frac{x}{\lambda_1+\lambda_2} + \frac{y_2}{\lambda_2} = 0 \tag{3.77}$$

Από τη σχέση αυτή προκύπτουν οι ΕΜΠ των λ_1, λ_2 και λ

$$\hat{\lambda}_1 = \frac{Y_1}{2}\left(\frac{X}{Y_1+Y_2}+1\right),\ \hat{\lambda}_2 = \frac{Y_2}{2}\left(\frac{X}{Y_1+Y_2}+1\right) \text{ και } \hat{\lambda} = \hat{\lambda}_1 + \hat{\lambda}_2. \tag{3.78}$$

Ορίζουμε τη συνάρτηση

$$g(\lambda, \lambda_1, \lambda_2) = \left(\frac{\lambda_1}{2}\right)\left(\frac{\lambda}{\lambda_1+\lambda_2}+1\right),$$

τότε $\hat{g} = g\left(\hat{\lambda}, \hat{\lambda}_1, \hat{\lambda}_2\right)$ και θέλουμε να βρούμε την ασυμπτωτική κατανομή του $\hat{g}$.

Έτσι έχουμε το διάνυσμα των μερικών παραγώγων της g να δίνεται από τη σχέση

$$G = \left(\frac{\lambda_1}{2\lambda}, 1-\frac{\lambda_1}{2\lambda}, -\frac{\lambda_1}{2\lambda}\right).$$

Ο πίνακας πληροφορίας του Fisher δίνεται από τη σχέση

$$I_{\mathbf{x}}^{F}(\boldsymbol{\lambda}) = \begin{pmatrix} \frac{1}{\lambda} & 0 & 0 \\ 0 & \frac{1}{\lambda_1} & 0 \\ 0 & 0 & \frac{1}{\lambda_2} \end{pmatrix}$$

και ο αντίστροφος του είναι

$$\left(I_{\mathbf{x}}^{F}(\boldsymbol{\lambda})\right)^{-1} = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_2 \end{pmatrix}$$

με $\boldsymbol{\lambda} = (\lambda, \lambda_1, \lambda_2)$.

Επομένως από τη σχέση (3.67)

$$\Lambda(\lambda)^{1x1} = G\left(I_{\mathbf{x}}^{F}(\theta)\right)^{-1} G' = \left(\frac{\lambda_1}{2\lambda}, 1-\frac{\lambda_1}{2\lambda}, -\frac{\lambda_1}{2\lambda}\right)\begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_2 \end{pmatrix}\left(\frac{\lambda_1}{2\lambda}, 1-\frac{\lambda_1}{2\lambda}, -\frac{\lambda_1}{2\lambda}\right)' =$$

$$= \lambda_1 - \frac{\lambda_1}{2\lambda}.$$

Από τη σχέση (3.66) ,λοιπόν,

$$\sqrt{n}\left(\mathbf{g}(\hat{\lambda}) - \mathbf{g}(\lambda)\right) \xrightarrow{\kappa} W \sim N_1\left(0, \Lambda(\lambda)\right).$$

Έτσι

$$\mathbf{g}(\hat{\lambda}) \xrightarrow{\kappa} W \sim N_1\left(\mathbf{g}(\lambda), \frac{1}{n}\Lambda(\lambda)\right).$$

Για τον πίνακα $\Lambda(\lambda)$ γνωρίζουμε ότι είναι το κάτω φράγμα της ανισότητας Cramer-Rao για τη διακύμανση ενός αμερόληπτου εκτιμητή της συνάρτησης $\hat{g}(\lambda)$.(βλέπε Ferguson (1996))

Επίλογος-Συμπέρασμα

Σε ότι προηγήθηκε έγινε μια προσπάθεια μελέτης της ανισότητας Cramer-Rao και των εφαρμογών της. Διατυπώθηκε και αποδείχθηκε η ανισότητα Cramer-Rao στην περίπτωση που η παράμετρος είναι μονοδιάστατη αλλά και πολυδιάστατη. Στη συνέχεια έγινε αναφορά στις εφαρμογές της στην εκτιμητική όσον αφορά την εύρεση Λ.Ο.Ε.Δ εκτιμητών. Ακόμη, αναφορά έγινε στη χρήση της ανισότητας για την εύρεση κάτω φράγματος για την διακύμανση συνάρτησης τ.μ που ακολουθεί γνωστή κατανομή. Τέλος μελετήθηκε η εφαρμογή της ανισότητας Cramer-Rao στην εύρεση της ασυμπτωτικής κατανομής ενός ΕΜΠ.

Είναι σαφές ότι στα πλαίσια μιας μεταπτυχιακής διατριβής δεν μπορεί να εξαντληθεί πλήρως ένα θέμα με τόσες πολλές προεκτάσεις και εφαρμογές όσο είναι η ανισότητα Cramer-Rao. Έτσι επιπλέον θα μπορούσε κανείς να αναφερθεί σε μία γενίκευση της όπου βρίσκεται ένα κάτω φράγμα για την κεντρική ροπή p-στής τάξης συνεχούς τυχαίας μεταβλητής χρησιμοποιώντας το γενικευμένο μέτρο πληροφορίας του Fisher. Εργασίες στις οποίες αναπτύσσεται το παραπάνω θέμα είναι των Lutwak, Yang, Zhang (2005) και Vajda (1972).

Μεταξύ των εφαρμογών της ανισότητας Cramer-Rao συγκαταλέγεται και η εύρεση των όρων της ασυμπτωτικής κατανομής του Ε.Μ.Π. Αυτή όμως η προσέγγιση δεν είναι τόσο εύχρηστη για διάφορους λόγους. Καταρχάς το κάτω φράγμα Cramer-Rao βρίσκεται για αμερόληπτους εκτιμητές ενώ οι περισσότεροι εκτιμητές στην πράξη είναι μεροληπτικοί. Επίσης για την εφαρμογή της ανισότητας Cramer-Rao χρειάζονται κάποιες συνθήκες κανονικότητας οι οποίες είναι δύσκολο να ελεγχθούν σε όλες τις περιπτώσεις και το αποτέλεσμα συγκρίνει μόνο εκτιμητές που ικανοποιούν αυτή τη συνθήκη. Για να αποφευχθούν οι παραπάνω δυσκολίες έχουν δημοσιευθεί διάφορες εργασίες όπως είναι η των Gill και Levit (1995) στην οποία δίνεται μια εκδοχή κατά Bayes για το κάτω φράγμα Cramer-Rao που οφείλεται στον Van Trees έτσι ώστε να αποδειχθεί ότι η προσεγγιστική κατανομή οποιουδήποτε κανονικού εκτιμητή δεν μπορεί να έχει διακύμανση μικρότερη από το κλασσικό φράγμα πληροφορίας , κάτω από ελάχιστες συνθήκες κανονικότητας.

Μεγάλο ενδιαφέρον παρουσιάζει η εφαρμογή της ανισότητας Cramer-Rao όταν δεν πληρούνται οι συνθήκες κανονικότητας (non- regular). Σχετικές εργασίες είναι των Vincze (1979), Mori (1983), Bartlett (1982) και άλλων. Επίσης, ενδιαφέρουσες είναι και οι μη παραμετρικές ανισότητες τύπου Cramer-Rao όπως παρουσιάζονται στην εργασία του Vincze (1992).

Σημαντικές είναι, τέλος, και οι ανισότητες τύπου Cramer-Rao κατά Bayes για έναν εκτιμητή που βασίζονται σε συνάρτηση ζημίας το τετραγωνικό σφάλμα. Τέτοιες εργασίες είναι των Rao P. (1991,1992,1999) που ακολούθησαν αυτή των Borovkov και Sakhanenko (1980).

Από τα όσα αναφέρθηκαν καταδεικνύεται η σημαντικότητα της ανισότητας Cramer-Rao στη στατιστική και κυρίως στη στατιστική συμπερασματολογία. Οι εφαρμογές της είναι πάρα πολλές και συνεχίζει να προσελκύει το ενδιαφέρον των ερευνητών αν και έχουν περάσει τόσα χρόνια από την εμφανισή της.

ΒΙΒΛΙΟΓΡΑΦΙΑ

Anderson, W. (1984). *An Introduction to Multivariate Statistical Analysis*.2nd Edition ,John Wiley, New York.

Apostol, T.M. (1974). *Mathematical Analysis,* Addison-Wesley Publishing Company.

Bartlett, M.S (1982). The Ideal Estimation Equation,*Journal of Applied Probability* 19,Essays in Statistical Science.187-200.

Borovkov, A. A , Sakhanenko A. I. (1980) : On estimates for the average quadratic risk, *Probabability and Mathematical Statistics* ,1, 185-195.

Cacoullos, T. (1982). On upper and lower bounds for the variance of a function of a random variable .*The Annals of Probability*, 10 ,799-809.

Casella, G. and Berger, R.L(2002). *Statistical Inference*, 2nd Edition.Wadsworth Inc, Belmont, California.

DeGroot , M.H (1987). A conversation with C.R.Rao. *Statistical Science*, 2 (1), 53-67.

Drygas, H. (1987). On the multivariate Cramer-Rao inequality. *Statistische Hefte*, 28, 69-71.

Ferentinos, K. (1984). Note of the use of the Cramer-Rao inequality for finding uniformly minimum variance unbiased estimators .*Metron,* XLII (1-2), 127-131.

Ferguson, T.S (1996). *A Course in Large Sample Theory*, Chapman& Hall, London.

Gill , D.R, Levit ,B. Y. (1995) : Applications of the van Trees inequality: A Bayesian Cramer-Rao bound *Bernoulli* 1, 59-79.

Giri, N.C. (2004). *Multivariate Statistical Analysis*.2nd Edition. Marcel Dekker. New York.

Johnson, R.A- Wichern, D.W (1998). *Applied Multivariate Statistical Analysis*, 4th Edition, Prentice Hall.

Joshi, V.M. (1976). On the attainment of the Cramer-Rao lower bound.*The Annals of Statistics*,4,(5),998-1002.

Kagan, A. (2001). Another look at the Cramer-Rao inequality. *The American Statistician*, 55, (3), 211-212.

Klotz, J. (2004). *A Computational Approach to Statistics*, Department of Statistics, University of Wisconsin at Madison.

Lehmann, E.L.(1959). *Testing Statistical Hypotheses*. John Wiley, New York.

Lutwak, E., Yang, D., Zhang G. (2005): Cramer-Rao and moment-entropy inequalities for Renyi entropy and generalized Fisher information. *IEEE Transactions on Information Theory* 51(2): 473-478

Mardia, K.V., Kent, J.T and Bibby, J.M. (1979). *Multivariate Analysis*. Academic Press, New York

Mood, A.M., Graybill, F.A and Boes, D.C. (1974). *Introduction to the Theory of Statistics*.3rd Edition ,McGraw-Hill Book Company, New York.

Mori, T.F (1983). Note on the Cramer-rao inequality in the non-regular case:The family of uniform distributions .*Journal of Statistical Planning and Inference* 7,353-358.

Muirhead, R.J (1982). *Aspects of Multivariate Statistical Theory* .John Wiley .New York.

Rao, R.C. (1965). *Linear Statistical Inference and Its Applications*.1st Edition ,John Wiley, New York.

Rao, R.C. (1973). *Linear Statistical Inference and Its Applications*.2nd Edition ,John Wiley, New York.

Rao ,P.B. (1991). On Cramer-Rao type integral inequalities', Hari Kinkar Nandi Memorial Volume (Ed. S.K. Chatterjee et al.), *Calcutta Statistical Association Bulletin 40* ,183-205 .

Rao, P.B (1992). Cramer-Rao type integral inequalities for estimators of functions of multidimensional parameter ***Sankhya* A** 54 ,53-73 .

Rao, P.B (1999). Cramer-Rao type integral inequalities for general loss functions, *Test 10* ,105-120 .

Rohatgi, V. K. (1976). An *Introduction to Probability Theory and Mathematical Statistics* . John Wiley, New York.

Rohatgi, V. K. and Saleh, E. (2001). *An Introduction to Probability Theory and Mathematical Statistics* .2nd Edition, John Wiley, New York.

Vajda, I. (1972). On the f-divergence and singularity of probability measures . *Periodica Mathematica Hungarica*.2 ,223-234.

Vincze, I. (1979). *On the Cramer-Frechet-Rao inequality in the non-regular case. In: Contributions to Statistics*. The J.Hajek Memorial Volume. Academia, Prague, 253-262.

Vincze, I. (1992). *On non-parametric Cramer-Rao inequalities, Order Statistics and Nonparametrics:Theory and Applications*(eds Sen, P.K and Salama ,I.A.)Elsevier Science Publishers B.V,439-454.

Wijsman, R.A. (1973). On the attainment of the Cramer-Rao lower bound.*The Annals of Statistics*,1,(3),538-542.

Zacks , **S.** (1971). *The theory of Statistical Inference*. John Wiley, New York.

Zografos, K. and Ferentinos, K. (1994). An information theoretic argument for the validity of the exponential model. *Metrika*, 41, 109-119.

Ντούγιας Σ.(1998). *Απειροστικός Λογισμος 1*. Εκδόσεις Πανεπιστημίου Ιωαννίνων.

Παπαιωάννου, Τ. και Φερεντίνος, Κ. (2000). *Μαθηματική Στατιστική, Εκτιμητική-Έλεγχος Υποθέσεων-Εφαρμογές*. Σταμούλης, Αθήνα.