ΕΘΝΙΚΟ ΚΑΙ ΚΑΠΟΔΙΣΤΡΙΑΚΟ

ΠΑΝΕΠΙΣΤΗΜΙΟ

ΑΘΗΝΩΝ

ΒΕΛΤΙΣΤΑ ΦΡΑΓΜΑΤΑ ΓΙΑ ΤΗΝ ΑΝΑΜΕΝΟΜΕΝΗ ΜΕΓΙΣΤΗ ΔΙΑΤΕΤΑΓΜΕΝΗ ΠΑΡΑΤΗΡΗΣΗ ΚΑΙ ΤΟ ΔΕΙΓΜΑΤΙΚΟ ΕΥΡΟΣ

Διπλωματική εργασία

του

Μπότση Ιωάννη

Επιβλέπων Καθηγητής

Παπαδάτος Νικόλαος

2014

ΠΕΡΙΕΧΟΜΕΝΑ

ΚΕΦΑΛΑΙΟ 1

ΜΙΑ ΕΙΣΑΓΩΓΗ ΣΤΑ ΦΡΑΓΜΑΤΑ ΤΗΣ ΑΝΑΜΕΝΟΜΕΝΗΣ ΤΙΜΗΣ ΤΗΣ ΜΕΓΙΣΤΗΣ ΔΙΑΤΕΤΑΓΜΕΝΗΣ ΠΑΡΑΤΗΡΗΣΗΣ ΚΑΙ ΤΟΥ ΕΥΡΟΥΣ ΓΙΑ ΤΥΧΑΙΑ ΔΕΙΓΜΑΤΑ.

ΚΕΦΑΛΑΙΟ 2

ΑΝΩ ΦΡΑΓΜΑΤΑ ΤΗΣ ΑΝΑΜΕΝΟΜΕΝΗΣ ΤΙΜΗΣ ΤΗΣ ΜΕΓΙΣΤΗΣ ΔΙΑΤΕΤΑΓΜΕΝΗΣ ΠΑΡΑΤΗΡΗΣΗΣ ΕΝΟΣ ΔΕΙΓΜΑΤΟΣ.

ΚΕΦΑΛΑΙΟ 3

ΒΕΛΤΙΣΤΟ ΑΝΩ ΦΡΑΓΜΑ ΓΙΑ ΤΗΝ ΑΝΑΜΕΝΟΜΕΝΗ ΤΙΜΗ ΤΟΥ ΜΕΓΙΣΤΟΥ.

ΚΕΦΑΛΑΙΟ 4

ΕΙΔΙΚΕΣ ΠΕΡΙΠΤΩΣΕΙΣ ΦΡΑΓΜΑΤΩΝ ΓΙΑ 2 ΔΙΑΤΕΤΑΓΜΕΝΕΣ ΠΑΡΑΤΗΡΗΣΕΙΣ

ΠΑΡΑΡΤΗΜΑ

ΚΑΤΑΣΚΕΥΗ ΦΡΑΓΜΑΤΩΝ

ΚΕΦΑΛΑΙΟ 5

ΕΥΡΕΣΗ ΑΝΩ ΦΡΑΓΜΑΤΩΝ ΓΙΑ ΤΟ ΕΥΡΟΣ n ΔΙΑΤΕΤΑΓΜΕΝΩΝ ΠΑΡΑΤΗΡΗΣΕΩΝ.

ΚΕΦΑΛΑΙΟ 6

ΕΥΡΕΣΗ ΒΕΛΤΙΣΤΩΝ ΑΝΩ ΦΡΑΓΜΑΤΩΝ ΓΙΑ ΤΟ ΕΥΡΟΣ n ΔΙΑΤΕΤΑΓΜΕΝΩΝ ΠΑΡΑΤΗΡΗΣΕΩΝ.

Πρόλογος

Ο όρος «διατεταγμένο δείγμα» (order statistics), εισήχθη μόλις το 1942, από τον Wilks. Ωστόσο, το αντικείμενο αυτό της στατιστικής, παρατηρείται πιο παλιά. Για παράδειγμα οι αστρονόμοι ενδιαφερόντουσαν για τις εκτιμήσεις της θέσης, πέρα από το μέσο όρο του δείγματος. Το 1818, λαμβάνεται (ουσιαστικά) από το Laplace, η κατανομή του r – οστού όρου σε ένα τυχαίο δείγμα. Ο πρώτος που ενδιαφέρθηκε για να βρει φράγματα στις ροπές των διατεταγμένων δειγμάτων, ήταν ο Sir Francis Galton το 1902.

Στην παρούσα εργασία, θα προσπαθήσουμε να βρούμε άνω φράγματα σε ένα διατεταγμένο δείγμα. Ειδικότερα, η εργασία χωρίζεται σε δύο ενότητες. Στην πρώτη ενότητα παρουσιάζονται φράγματα για την αναμενόμενη τιμή της μέγιστης παρατήρησης του διατεταγμένου δείγματος (Κεφάλαιο 2 – 4). Στη δεύτερη ενότητα βρίσκουμε φράγματα για την αναμενόμενη τιμή του εύρους του δείγματος, δηλαδή της διαφοράς της μικρότερης διατεταγμένης παρατήρησης από τη μεγαλύτερη διατεταγμένη παρατήρηση (Κεφάλαια 5 – 6).

Στο πρώτο κεφάλαιο, γίνεται μία γρήγορη αναφορά στα διατεταγμένα δείγματα. Βρίσκουμε ένα άνω φράγμα για την μέγιστη αναμενόμενη τιμή σε τυχαίο δείγμα. Το φράγμα αυτό παρουσιάζεται στα άρθρα των Hartley – David 1954 και Gumbel 1954. Τα οποία δημοσιεύθηκαν στο ίδιο περιοδικό και με διαδοχή. Το πρώτο βρίσκεται στις σελίδες 76 – 84 και το δεύτερο 85 – 99 . Στη συνέχεια χρησιμοποιώντας μια ανισότητα που εμφανίζεται στο άρθρο του Ludwig (1959), βρίσκουμε ένα φράγμα για την αναμενόμενη τιμή του εύρους.

Στο δεύτερο κεφάλαιο βρίσκουμε άνω φράγμα για τη μέγιστη διατεταγμένη παρατήρηση. Οι Arnold – Groeneveld (1979) και ο Papadatos (2001) βρήκαν ένα άνω φράγμα για την παρατήρηση αυτή. Ο δεύτερος, βρήκε ένα καλύτερο άνω φράγμα. Αρχικά βρήκε ένα άνω φράγμα για μια ειδική κατηγορία τ.μ., που όποια τους μετάθεση και να πάρουμε, έχουν κοινή κατανομή. Στη συνέχεια το γενίκευσε σε όλα τα διατεταγμένα δείγματα.

Στο τρίτο κεφάλαιο προσπαθούμε να βρούμε έλα βέλτιστο άνω φράγμα για τη μέγιστη αναμενόμενη παρατήρηση. Οι Bertsimas, Natarajan and Teo (2006) βρήκαν μία κατανομή που επιτυγχάνεται το φράγμα για τη μέγιστη αναμενόμενη παρατήρηση, χρησιμοποιώντας την κυρτότητα κάποιων συναρτήσεων.

Στο τέταρτο κεφάλαιο συγκρίνουμε τα φράγματα των προηγουμένων κεφαλαίων για την ειδική περίπτωση που $n = 2$. Στη συνέχεια διαπιστώνουμε μία άμεση σχέση που έχει το εύρος με τη μέγιστη διατεταγμένη παρατήρηση. Στο συγκεκριμένο κεφάλαιο, βρίσκουμε φράγματα για το έυρος των δύο παρατηρήσεων.

Στα κεφάλαια 5 και 6, βρίσκουμε άνω φράγματα για το εύρος. Στην ουσία χρησιμοποιήσαμε ένα αποτέλεσμα του Isii (1962) για την κατασκευή φραγμάτων. ΟPapadatos (2014) βρήκε άνω φράγματα για το αναμενόμενο εύρος. Ειδικότερα στην εργασία αυτή εμφανίζουμε βέλτιστο άνω φράγμα για το αναμενόμενο εύρος μόνο για την περίπτωση των Arnold – Groeneveld. Περισσότερα βέλτιστα φράγματα βρέθηκαν από τον Papadatos (2014) (πρβλ. βιβλ. **[11]**).

ΚΕΦΑΛΑΙΟ 1

ΜΙΑ ΕΙΣΑΓΩΓΗ ΣΤΑ ΦΡΑΓΜΑΤΑ ΤΗΣ ΑΝΑΜΕΝΟΜΕΝΗΣ ΤΙΜΗΣ ΤΗΣ ΜΕΓΙΣΤΗΣ ΔΙΑΤΕΤΑΓΜΕΝΗΣ ΠΑΡΑΤΗΡΗΣΗΣ ΚΑΙ ΤΟΥ ΕΥΡΟΥΣ ΓΙΑ ΤΥΧΑΙΑ ΔΕΙΓΜΑΤΑ.

1.1. Μια σύντομη αναφορά στα διατεταγμένα δείγματα (order statistics)

Έστω $X_1, X_2, \dots, X_n$ είναι ένα τυχαίο δείγμα με κοινή συνάρτηση κατανομής F, και $X_{1:n}$, $X_{2:n}$, … , $X_{n:n}$, το διατεταγμένο δείγμα που προκύπτει αν διατάξουμε τις $X_1, X_2, \dots, X_n$ κατά αύξουσα σειρά μεγέθους. Στο διατεταγμένο οι τυχαίες μεταβλητές δεν είναι ούτε ανεξάρτητες, ούτε ισόνομες. Η κατανομή μιας διατεταγμένης παρατήρησης καθορίζεται εύκολα. Για i = 1, 2, …, n και $x \in \mathbb{R}$ έχουμε

$$F_{X_{i:n}}(x) = \mathbb{P}(X_{i:n} \le x) =$$

$$= \mathbb{P}(\text{τουλάχιστον i παρατηρήσεις από τις } X_1, X_2, \dots, X_n \text{ είναι } \le x) =$$

$$= \sum_{j=i}^{n} \mathbb{P}(\text{ακριβώς j παρατηρήσεις από τις } X_1, X_2, \dots, X_n \text{ είναι } \le x).$$

Επειδή η πιθανότητα «επιτυχίας» σε κάθε δοκιμή είναι $F(x) = \mathbb{P}(X_i \le x)$, προκύπτει ότι:

$$F_{X_{i:n}}(x) = \sum_{j=i}^{n} \binom{n}{j} [F(x)]^j [1 - F(x)]^{n-j}.$$

1.2. Φράγματα για τυχαίο δείγμα

Έστω $X_{1:n}$, $X_{2:n}$, … , $X_{n:n}$, ένα διατεταγμένο δείγμα (order statistics) του δείγματος των τυχαίων μεταβλητών X_1 , X_2 , … , X_n. Στην περίπτωση που οι τ.μ είναι ανεξάρτητες και ισόνομες με συνάρτηση κατανομής F, έχει αποδειχθεί (Hartley – David 1954 και Gumbel 1954) ότι αν το τυχαίο αυτό δείγμα έχει μέση τιμή μ και πεπερασμένη διασπορά σ^2, τότε για την αναμενόμενη τιμή της μέγιστης παρατήρησης $X_{n:n}$ ισχύει ότι:

$$E(X_{n:n}) \le \mu + \sigma \cdot \frac{n-1}{\sqrt{2n-1}} \quad (1.1).$$

Αποδεικνύεται επίσης ότι για κάθε $\mu \in \mathbb{R}$ και $\sigma^2 > 0$, υπάρχει συνάρτηση κατανομής για την οποία επιτυγχάνεται το φράγμα της (1.1).[1]

Η κατανομή αυτή είναι μοναδική, και δίνεται από τον τύπο

[1] Τα αποτελέσματα εμφανίστηκαν στο άρθρο “The maxima of the mean largest value and of the range” από τον Gumbel το 1954 και στο άρθρο “Universal bounds of mean range and extreme observation” από τους David και Hartley το 1954. Τα οποία δημοσιεύθηκαν στο ίδιο περιοδικό και με διαδοχή. Το πρώτο βρίσκεται στις σελίδες 76 – 84 και το δεύτερο 85 – 99 .

$$F(x) = \left\{\frac{1}{n}\left[\frac{n-1}{\sqrt{2n-1}}\frac{x-\mu}{\sigma}+1\right]\right\}^{1/n-1}, \text{ για } \mu-\sigma\cdot\frac{n-1}{\sqrt{2n-1}} \le x \le \mu+\sigma\cdot\sqrt{2n-1}$$

Αν $X_{j:n}$ και $X_{i:n}$ είναι αντίστοιχα η j – οστή και i – οστή $(j > i)$ διατεταγμένη από n παρατηρήσεις του παραπάνω τυχαίου δείγματος, τότε

$$W_{i,j:n} = X_{j:n} - X_{i:n},\ 1 \le i \le j \le n\,,$$

είναι η απόσταση (spacing) δυο τυχόντων διατεταγμένων παρατηρήσεων. Ισχύει η ανισότητα (Ludwig 1959):

$$E\big(W_{i,j:n}\big) \le \sigma\cdot\left\{\frac{B(2j-1,2n-2j+1)}{[B(j,n-j+1)]^2}+\frac{B(2i-1,2n-2i+1)}{[B(i,n-i+1)]^2}-2\frac{B(j+i-1,2n-i-j+1)}{B(j,n-j+1)B(i,n-i+1)}\right\}^{1/2}. \qquad (1.2)$$

Όταν i = 1 και j = n, η $W_{i,j:n}$ εκφράζει το εύρος (range) του τ.δ και συμβολίζεται με R_n $\big(R_n := W_{i,j:n}\big)$. H (1.2) σε αυτήν την περίπτωση γίνεται :

$$\begin{aligned}E(R_n) &\le \sigma\cdot\left\{\frac{B(2n-1,2n-2n+1)}{[B(n,n-n+1)]^2}+\frac{B(2-1,2n-2+1)}{[B(1,n-1+1)]^2}-2\frac{B(n+1-1,2n-1-n+1)}{B(n,n-n+1)B(1,n-1+1)}\right\}^{1/2}\\ &= \sigma\cdot\left\{\frac{B(2n-1,1)}{[B(n,1)]^2}+\frac{B(1,2n-1)}{[B(1,n)]^2}-2\frac{B(n,n)}{B(n,1)B(1,n)}\right\}^{1/2}\\ &= \sigma\cdot\frac{1}{B(n,1)}\{B(2n-1,1)+B(1,2n-1)-2B(n,n)\}^{1/2}\\ &= \sigma\cdot n\left\{\frac{1}{2n-1}+\frac{1}{2n-1}-2\frac{(n-1)!(n-1)!}{(2n-1)!}\right\}^{1/2}\\ &= \sigma\cdot n\frac{\sqrt{2}}{\sqrt{2n-1}}\left\{1-\frac{(n-1)!(n-1)!}{(2n-2)!}\right\}^{1/2}.\end{aligned}$$

Δηλαδή

$$E(R_n) \le \sigma\frac{n\sqrt{2}}{\sqrt{2n-1}}\left\{1-\frac{1}{\binom{2n-2}{n-1}}\right\}^{1/2}. \qquad (1.3).$$

Μάλιστα το φράγμα της αναμενόμενης απόστασης (1.2) επιτυγχάνεται μόνο στην περίπτωση της (1.3). Δηλαδή επιτυγχάνεται μόνο για το εύρος του δείγματος.

Στην εργασία αυτή θα ασχοληθούμε με την εύρεση άνω και κάτω φραγμάτων για τη μέγιστη παρατήρηση και για το εύρος, δειγμάτων όχι κατ' ανάγκη ανεξαρτήτων. Στόχος της είναι να προσδιορίσουμε τα βέλτιστα δυνατά φράγματα, δηλαδή να προσδιοριστούν οι ποσότητες

$$\max E(R_n), \qquad \min E(R_n), \quad \max E(X_{n:n}), \quad \min E(X_{n:n})$$

και να βρούμε της κατανομές των X_i, i=1, …, n, ώστε να μπορούν να επιτευχθούν τα φράγματα αυτά.

ΚΕΦΑΛΑΙΟ 2

ΑΝΩ ΦΡΑΓΜΑΤΑ ΤΗΣ ΑΝΑΜΕΝΟΜΕΝΗΣ ΤΙΜΗΣ ΤΗΣ ΜΕΓΙΣΤΗΣ ΔΙΑΤΕΤΑΓΜΕΝΗΣ ΠΑΡΑΤΗΡΗΣΗΣ ΕΝΟΣ ΔΕΙΓΜΑΤΟΣ.

Έστω ένα δείγμα από n διατεταγμένες παρατηρήσεις. Στο κεφάλαιο αυτό, θα προσπαθήσουμε να βρούμε φράγματα για την αναμενόμενη τιμή της μέγιστης διατεταγμένης παρατήρησης. Ειδικότερα, στην ενότητα 2.1 παρουσιάζονται τέτοια φράγματα όταν οι παρατηρήσεις του δείγματος έχουν ίδια μέση τιμή και διασπορά. Τα φράγματα αυτά κατασκευάστηκαν από τους Arnold και Groeneveld το 1979. Όταν οι παρατηρήσεις δεν έχουν ίδια μέση τιμή και διασπορά, και η συνδιακύμανσή τους δεν είναι μηδέν, τότε για να βρεθούν φράγματα για την $E(X_{n:n})$, θα πρέπει πρώτα να βρεθούν φράγματα, για την αναμενόμενη αυτή τιμή, όταν οι τυχαίες μεταβλητές είναι ανταλλάξιμες (ο ορισμός τους βρίσκεται στην ενότητα 2.2). Στη συνέχεια γενικεύουμε, για όλες τις διατεταγμένες παρατηρήσεις. (ενότητα 2.3).

2.1. Βέλτιστα φράγματα για την αναμενόμενη τιμή της μέγιστης παρατήρησης του δείγματος με κοινή μέση τιμή και διασπορά.

Έστω $X_1, X_2, \dots, X_n$, έστω ένα σύνολο από τυχαίες μεταβλητές (οι οποίες δεν είναι απαραίτητα ανεξάρτητες), με αντίστοιχες μέσες τιμές $\mu_i = E(X_i)$ και πεπερασμένες διασπορές $\sigma_i^2 = Var(X_i)$ $i = 1, 2, \dots, n$ και θεωρούμε της αντίστοιχες διατεταγμένες παρατηρήσεις $X_{1:n} \le X_{2:n} \le \dots \le X_{n:n}$ της παραγράφου 1.1. Τότε προκύπτει το εξής:

Θεώρημα 2.1.1. (Arnold – Groeneveld 1979)

Για οποιεσδήποτε σταθερές $\lambda_1, \lambda_2, \dots, \lambda_n$,

$$\left|\sum_{i=1}^n \lambda_i (E(X_{i:n}) - \bar{\mu})\right| \le \left(\sum_{i=1}^n (\lambda_i - \bar{\lambda})^2\right)^{1/2} \left(\sum_{i=1}^n \left\{(\mu_i - \bar{\mu})^2 + \sigma_i^2\right\}\right)^{1/2} \quad (2.1)$$

όπου $\bar{\mu} = \frac{1}{n}\sum_{i=1}^n \mu_i$ και $\bar{\lambda} = \frac{1}{n}\sum_{i=1}^n \lambda_i$. (Ανισότητα A – G)

Απόδειξη

Έχουμε

$$\left|\sum_{i=1}^n \lambda_i (E(X_{i:n}) - \bar{\mu})\right| = \left|\sum_{i=1}^n \lambda_i (E(X_{i:n}) - \bar{\mu}) - \sum_{i=1}^n \bar{\lambda}(E(X_{i:n}) - \bar{\mu})\right|$$

$$= \left|\sum_{i=1}^n (\lambda_i - \bar{\lambda})(E(X_{i:n}) - \bar{\mu})\right| \quad (2.2)$$

διότι

$$\sum_{i=1}^{n} \bar{\lambda}(E(X_{i:n}) - \bar{\mu}) = \bar{\lambda}E\left(\sum_{i=1}^{n} X_{i:n}\right) - n\bar{\lambda}\bar{\mu}$$

$$= \bar{\lambda}E(\textstyle\sum_{i=1}^{n} X_i) - n\bar{\lambda}\bar{\mu}$$

$$= \bar{\lambda}\textstyle\sum_{i=1}^{n} E(X_i) - n\bar{\lambda}\bar{\mu}$$

$$= n\bar{\lambda}\bar{\mu} - n\bar{\lambda}\bar{\mu} = 0.$$

Εφαρμόζοντας την ανισότητα Cauchy – Schwarz στην (2.2) έχουμε:

$$\left|\textstyle\sum_{i=1}^{n} \lambda_i(E(X_{i:n}) - \bar{\mu})\right| \leq \left[\textstyle\sum_{i=1}^{n}\left(\lambda_i - \bar{\lambda}\right)^2\right]^{1/2} \left[\textstyle\sum_{i=1}^{n}(E(X_{i:n}) - \bar{\mu})^2\right]^{1/2}. \qquad (2.3)$$

Όμως

$$\sum_{i=1}^{n}(E(X_{i:n}) - \bar{\mu})^2 = \sum_{i=1}^{n} E(X_{i:n})^2 - n\bar{\mu}^2$$

$$\leq \sum_{i=1}^{n} E\left({X_{i:n}}^2\right) - n\bar{\mu}^2$$

$$= \textstyle\sum_{i=1}^{n} E\left({X_i}^2\right) - n\bar{\mu}^2$$

$$= \textstyle\sum_{i=1}^{n}\left[Var(X_i) + E^2(X_i)\right] - n\bar{\mu}^2$$

$$= \textstyle\sum_{i=1}^{n} {\sigma_i}^2 + \sum_{i=1}^{n} {\mu_i}^2 - n\bar{\mu}^2$$

$$= \textstyle\sum_{i=1}^{n}\left\{{\sigma_i}^2 + \left(\mu_i - \bar{\mu}\right)^2\right\} \qquad (2.4)$$

$$\overset{(2.3)}{\underset{(2.4)}{\Rightarrow}} \quad \left|\textstyle\sum_{i=1}^{n} \lambda_i(E(X_{i:n}) - \bar{\mu})\right| \leq \left(\textstyle\sum_{i=1}^{n}\left(\lambda_i - \bar{\lambda}\right)^2\right)^{1/2} \left(\textstyle\sum_{i=1}^{n}\left\{\left(\mu_i - \bar{\mu}\right)^2 + {\sigma_i}^2\right\}\right)^{1/2}$$

$$= \left[\textstyle\sum_{i=1}^{n}\left(\lambda_i - \bar{\lambda}\right)^2 \sum_{i=1}^{n}\left\{\left(\mu_i - \bar{\mu}\right)^2 + {\sigma_i}^2\right\}\right]^{1/2}.$$

■

<u>Εφαρμογή 2.1.2:</u> Μεμονωμένες διατεταγμένες παρατηρήσεις $X_{k:n}$.

Για τις διατεταγμένες παρατηρήσεις, ισχύει ότι $X_{i:n} \leq X_{k:n}$, $i = 1,2, \dots, k$, οπότε $\sum_{i=1}^{k} E(X_{i:n}) \leq \sum_{i=1}^{k} E(X_{k:n})$

δηλαδή

$$E(X_{k:n}) \geq \frac{1}{k}\textstyle\sum_{i=1}^{k} E(X_{i:n}). \qquad (2.5)$$

Όμοια προκύπτει ότι

$$E(X_{k:n}) \leq \frac{1}{n-k+1}\sum_{i=k}^{n} E(X_{i:n}) \, , \tag{2.6}$$

αφού $X_{k:n} \leq X_{i:n} \quad i = k, k+1, \dots, n$.

Οπότε :

$$\frac{1}{k}\sum_{i=1}^{k} E(X_{i:n}) \leq E(X_{k:n}) \leq \frac{1}{n-k+1}\sum_{i=k}^{n} E(X_{i:n}). \tag{2.7}$$

Αν θεωρήσουμε ότι οι X_i, $i = 1,2,\dots,n$ έχουν μέση τιμή μ και διασπορά σ^2, τότε θέτοντας στην (2.1) $\lambda_1 = \lambda_2 = \cdots = \lambda_{\kappa-1} = 0$ και $\lambda_\kappa = \lambda_{\kappa+1} = \cdots = \lambda_n = 1$ προκύπτει η ανισότητα

$$\sum_{i=\kappa}^{n}(E(X_{i:n}) - \mu) \leq \left(\sum_{i=\kappa}^{n}\left(1 - \frac{n-\kappa+1}{n}\right)^2 + \sum_{i=1}^{k-1}\left(0 - \frac{n-\kappa+1}{n}\right)^2\right)^{1/2} \left(\sum_{i=1}^{n}\{(\mu-\mu)^2 + \sigma^2\}\right)^{1/2},$$

η οποία γράφεται ως

$$\sum_{i=\kappa}^{n} E(X_{i:n}) - (n-\kappa+1)\mu \leq \left(\left(\frac{\kappa-1}{n}\right)^2 (n-\kappa+1) + \left(\frac{n-k+1}{n}\right)^2 (\kappa-1)\right)^{1/2} (n\sigma^2)^{1/2}$$

$$\Leftrightarrow \sum_{i=\kappa}^{n} E(X_{i:n}) - (n-\kappa+1)\mu \leq \frac{[(\kappa-1)(n-\kappa+1)]^{1/2}}{n} n^{1/2} \cdot n^{1/2} \cdot \sigma$$

$$\Leftrightarrow \sum_{i=\kappa}^{n} E(X_{i:n}) \leq (n-\kappa+1)\mu + [(\kappa-1)(n-\kappa+1)]^{1/2} \cdot \sigma.$$

Άρα από τη σχέση (2.7) προκύπτει

$$E(X_{k:n}) \leq \mu + \sigma\left\{\frac{k-1}{n-k+1}\right\}^{1/2}. \tag{2.8}$$

Ενώ με τον ίδιο τρόπο θέτοντας στην (2.1)

$\lambda_1 = \lambda_2 = \cdots = \lambda_\kappa = 1$ και $\lambda_{\kappa+1} = \lambda_{\kappa+2} = \cdots = \lambda_n = 0$ προκύπτει

$$\mu - \sigma\left\{\frac{(n-k)}{k}\right\}^{1/2} \leq E(X_{k:n}). \tag{2.9}$$

Καταλήξαμε λοιπόν στο εξής

Πόρισμα 2.1.3

Αν $X_1, X_2, \dots, X_n$ είναι οποιεσδήποτε τυχαίες μεταβλητές με $E(X_i) = \mu$, $Var(X_i) = \sigma^2$ (και οποιαδήποτε συσχέτιση μεταξύ τους) τότε ισχύει η διπλή ανισότητα

$$\mu - \sigma\sqrt{\frac{n-\kappa}{n}} \leq E(X_{\kappa:n}) \leq \mu + \sigma\sqrt{\frac{\kappa-1}{n+1-\kappa}}.$$

■

Παρατήρηση 2.1.4

Αν εφαρμόσουμε το Θεώρημα 2.1.1 για $\lambda_i = \begin{cases} 0, & i \neq \kappa \\ 1, & i = \kappa \end{cases}$, παίρνουμε:

$\sum_{i=1}^{n} \lambda_i (E(X_{i:n}) - \bar{\mu}) = E(X_{\kappa:n}) - \bar{\mu}$, και η ανισότητα A – G γράφεται ως:

$E(X_{\kappa:n}) \leq \bar{\mu} + \sqrt{\left(1 - \frac{1}{n}\right)} \sqrt{\sum_{i=1}^{n}(\mu_i - \bar{\mu})^2 + \sigma_i^2}.$

Στην περίπτωση που $\mu_i = \mu$ και $\sigma_i^2 = \sigma^2$, παίρνουμε χειρότερο αποτέλεσμα από αυτό του πορίσματος:

$$E(X_{\kappa:n}) \leq \mu + \sqrt{\left(1 - \frac{1}{n}\right)} \sqrt{\sum_{i=1}^{n}(\mu - \mu)^2 + n\sigma^2} = \mu + \sigma\sqrt{n-1}\ .$$

Για $k = n$ οδηγούμαστε στην ανισότητα:

$$E(X_{n:n}) \leq \mu + \sigma(n-1)^{\frac{1}{2}} \qquad (2.10).$$

Δηλαδή οδηγηθήκαμε σε ένα άνω φράγμα για τη μέγιστη παρατήρηση, χωρίς να χρειάζεται η υπόθεση της ανεξαρτησίας όπως στην (1.1).

Επίσης από την (2.9) προκύπτει ανισότητα:

$$E(X_{n:n}) \geq \mu$$

Το κάτω φράγμα αυτό επιδέχεται βελτίωση την οποία την πέτυχε ο Hawkins το 1971 λαμβάνοντας δείγματα χωρίς επανάθεση από πεπερασμένους πληθυσμούς, και χρησιμοποιώντας ένα κλασσικό αποτέλεσμα των Pearson και Chandra Sekar (1936). Πράγματι στην περίπτωση δειγμάτων χωρίς επανάθεση από πεπερασμένους πληθυσμούς, ο Hawkins έδειξε ότι

$$E(X_{n:n}) \geq \mu + \frac{\sigma}{\sqrt{n-1}}\ . \qquad (2.11)$$

Όμως στη γενική περίπτωση προκύπτει ότι $\min E(X_{n:n}) = \mu$

2.2. Άνω φράγματα για την αναμενόμενη τιμή της μέγιστης παρατήρησης σε εξαρτημένο δείγμα από ανταλλάξιμες τυχαίες μεταβλητές.

Τώρα θα δείξουμε ένα γενικότερο αποτέλεσμα Έστω ένα σύνολο από $n \geq 2$ ανταλλάξιμες τυχαίες μεταβλητές [2] $X_1, X_2, \ldots, X_n$ με $E(X_1) = \mu$, $Var(X_1) = \sigma^2$ (την υποθέτουμε πεπερασμένη) και $Cov(X_1, X_2) = c$. Τότε ισχύει η ανισότητα:

$$E(X_{n:n}) \leq \mu + \frac{n-1}{\sqrt{n}}\sqrt{\sigma^2 - c}\,. \tag{2.12}$$

Αυτή προκύπτει εύκολα από το παρακάτω Λήμμα:

<u>Λήμμα 2.2.1 (Papadatos 2000)</u> Έστω ένα τυχαίο διάνυσμα $\mathbf{X} = \left(X_1, X_2, \ldots, X_n\right)'$, $n \geq 2$, αποτελούμενο από ανταλλάξιμες τυχαίες μεταβλητές με $E(X_1) = \mu$, $Var(X_1) = \sigma^2$ (την υποθέτουμε πεπερασμένη) και $Cov(X_1, X_2) = c$. Τότε για κάθε πραγματικές σταθερές $\lambda_1, \lambda_2, \ldots, \lambda_n$, ισχύει η ανισ'οτητα

$$\left|\sum_{i=1}^n \lambda_i (E(X_{i:n}) - \mu)\right| \leq \left(\sum_{i=1}^n (\lambda_i - \bar{\lambda})^2\right)^{1/2} \left((n-1)(\sigma^2 - c)\right)^{1/2} \tag{2.13}$$,

όπου $\bar{\lambda} = \frac{1}{n}\sum_{i=1}^n \lambda_i$ και $X_{1:n} \leq X_{2:n} \leq \ldots \leq X_{n:n}$ είναι οι διατεταγμένες παρατηρήσεις που αντιστοιχούν στο διάνυσμα **X**. Επιπρόσθετα το φράγμα στην (2.13) είναι το καλύτερο δυνατό (για κάθε μ, σ^2, και c γνωστά) εφόσον η πεπερασμένη ακολουθία $\lambda_1, \lambda_2, \ldots, \lambda_n$ είναι μονότονη[3].

<u>Απόδειξη</u>

Αρχικά διαπιστώνουμε ότι $-\sigma^2/(n-1) \leq c \leq \sigma^2$ [4], οπότε το άνω φράγμα είναι καλά ορισμένο. Μετά παρατηρούμε ότι $E(\bar{X}) = \mu$, οπότε :

$$\begin{aligned}\sum_{i=1}^n E^2(X_{i:n} - \mu) &= \sum_{i=1}^n E^2(X_{i:n} - \bar{X}) \\ &\leq \sum_{i=1}^n E(X_{i:n} - \bar{X})^2\end{aligned}$$

---

[2] Οι τ.μ. είναι ανταλλάξιμες, δηλαδή $(X_1, X_2, \ldots, X_n) =_d (X_{i_1}, X_{i_2}, \ldots, X_{i_n})$ για κάθε επιλογή των $i_1, i_2, \ldots, i_n$ από τα 1, 2, …, n, οπότε και οι περιθώριες τ.μ θα έχουν την ίδια κατανομη. Δηλαδή $X_1 =_d X_i$ και $(X_1, X_2) =_d (X_i, X_j)$ κ.ο.κ.

[3] Για να ισχύει η ισότητα στην ανισότητα του Cauchy – Schwarz, θα πρέπει τα λ_i να είναι πολλαπλάσια του $E(X_{i:n}) - \mu$. Αφού η ακολουθία $E(X_{i:n})$ είναι μη φθίνουσα, τότε η ακολουθία των λ_i θα είναι μονότονη.

[4] $E(X_i) = E(X_1) = \mu$ και $Var(X_i) = Var(X_1) = \sigma^2$, $cov(X_1, X_2) = cov(X_i, X_j) = c$, i, j = 1, 2, …, n. Άρα
$E(\bar{X}) = E\left(\frac{1}{n}\sum_{i=1}^n X_i\right) = \frac{1}{n}\sum_{i=1}^n E(X_i) = \frac{1}{n} n \cdot \mu = \mu$,
$[cov(X_1, X_2)]^2 \leq Var(X_1)Var(X_2) \Leftrightarrow c \leq \sigma^2$ και
$\sum_{i=1}^n Var(X_i) + 2\sum\sum_{1 \leq i < j \leq n} cov(X_i, X_j) = Var(\sum_{i=1}^n X_i) \geq 0 \Leftrightarrow$
$n\sigma^2 + 2n(n-1)c \geq 0 \Leftrightarrow -\sigma^2/(n-1) \leq c$.

$$= E(\sum_{i=1}^{n}(X_{i:n} - \overline{X})^2)$$

$$= E(\sum_{i=1}^{n}(X_i - \overline{X})^2)$$

$$= (n-1)(\sigma^2 - c). \qquad (2.14)$$

Όπως στην απόδειξη της (2.1) στις σχέσεις (2.2), (2.3), αν αντικαθιστούσαμε το $\overline{\mu}$ με το μ, θα προκύψει:

$$\left|\sum_{i=1}^{n}\lambda_i(E(X_{i:n}) - \mu)\right| = \left|\sum_{i=1}^{n}\lambda_i(E(X_{i:n}) - \mu) - \sum_{i=1}^{n}\overline{\lambda}(E(X_{i:n}) - \mu)\right|$$

$$= \left|\sum_{i=1}^{n}(\lambda_i - \overline{\lambda})(E(X_{i:n}) - \mu)\right|$$

$$\leq \left[\sum_{i=1}^{n}\left(\lambda_i - \overline{\lambda}\right)^2\right]^{1/2}\left[\sum_{i=1}^{n}(E(X_{i:n}) - \overline{\mu})^2\right]^{1/2}$$

$$= \left(\sum_{i=1}^{n}\left(\lambda_i - \overline{\lambda}\right)^2\right)^{1/2}\left((n-1)(\sigma^2 - c)\right)^{1/2}$$

(η τελευταία προκύπτει λόγω της (2.14)).

Αν τα $\lambda_1 = \lambda_2 = \cdots = \lambda_n = \lambda$, τότε παίρνουμε 0=0.

Θα δείξουμε τώρα ότι η ισότητα επιτυγχάνεται για οποιεσδήποτε τιμές των μ, σ^2, c. Αν τα λ_i είναι μια μη φθίνουσα ακολουθία με $\lambda_1 < \lambda_n$, τότε θεωρούμε μια αυθαίρετη τ.μ. Z με $E(Z) = \mu$ και $Var(Z) = (\sigma^2 + (n-1)c)/n$, και θέτουμε $Y_j = Z + A\left(\lambda_j - \overline{\lambda}\right)$, $j = 1,2,\ldots,n$ όπου $A = \left(\frac{(n-1)(\sigma^2 - c)}{\sum_{i=1}^{n}\left(\lambda_i - \overline{\lambda}\right)^2}\right)^{1/2}$ [5].

Τώρα θα δείξουμε ότι το τυχαίο διάνυσμα

$$\left(X_1, X_2, \ldots, X_n\right)' = \left(Y_{\pi(1)}, Y_{\pi(2)}, \ldots, Y_{\pi(n)}\right)',$$

ικανοποιεί τις προϋποθέσεις του Λήμματος, και πετυχαίνει την ισότητα στην (2.13).

Εδώ το διάνυσμα $(\pi(1), \pi(2), \ldots, \pi(n))'$ είναι στοχαστικά ανεξάρτητο του Z και ομοιόμορφακατανεμημένο στις n! μεταθέσεις του $\{1,2,\ldots,n\}$ [6], ικανοποιεί τις προϋποθέσεις του Λήμματος, οπότε πετυχαίνεται η ισότητα στην (2.13).

Ας θεωρήσουμε την τ.μ. $g\left(\pi(n)\right) = E\left(Y_{\pi(n)}|\pi(n)\right)$. Έχουμε

---

[5] Στο Λήμμα 1.1, το φράγμα της ανισότητας (2.13) επιτυγχάνεται όταν ισχύει η ισότητα. Δηλαδή όταν ισχύει η ισότητα στην Cauchy – Schwarz . Οπότε το φράγμα επιτυγχάνεται όταν υπάρχει σταθερά A τέτοια ώστε $[(n-1)(\sigma^2 - c)]^{1/2} = A\left[\sum_{i=1}^{n}\left(\lambda_i - \overline{\lambda}\right)^2\right]^{1/2}$.

[6] Το $(\pi(1), \pi(2), \ldots, \pi(n))'$ έχει ομοιόμορφη κατανομή στις n! μεταθέσεις του $\{1,2,\ldots,n\}$, δηλαδή $\mathbb{P}(\pi(1) = j_1, \pi(2) = j_2, \ldots, \pi(n) = j_n) = \frac{1}{n!}$ για $j_1, j_2, \ldots, j_n \in \{1,2,\ldots,n\}$ με $j_r \neq j_s$ για $r \neq s$. Προφανώς $\mathbb{P}(\pi(i) = j) = \frac{1}{n}$, για κάθε i και για κάθε j.

$$
\begin{aligned}
\mathrm{E}\left(\mathrm{Y}_{\pi(1)}\right) &= \mathrm{E}\left(\mathrm{E}\left(\mathrm{Y}_{\pi(1)}|\pi(1)\right)\right) \\
&= \mathrm{E}\left(\mathrm{g}\left(\pi(1)\right)\right) \\
&= \sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{E}\left(\mathrm{g}(\pi(1)|\pi(1)=\mathrm{j})\right) \cdot \mathbb{P}(\pi(1)=\mathrm{j}) \\
&= \frac{1}{\mathrm{n}} \sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{E}\left(\mathrm{g}(\mathrm{j})\right) = \frac{1}{\mathrm{n}} \sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{E}\left(\mathrm{Y}_{\mathrm{j}}\right) \\
&= \frac{1}{\mathrm{n}} \mathrm{n}\mu + \frac{1}{\mathrm{n}} \mathrm{A} \sum_{\mathrm{i}=1}^{\mathrm{n}} \left(\lambda_j - \bar{\lambda}\right) \\
&= \mu + 0 \\
&= \mu.
\end{aligned}
$$

Επίσης

$$
\begin{aligned}
\mathrm{Var}\left(\mathrm{Y}_{\pi(1)}\right) &= \mathrm{E}\left({\mathrm{Y}_{\pi(1)}}^2\right) - \left[\mathrm{E}\left(\mathrm{Y}_{\pi(1)}\right)\right]^2 \\
&= \frac{1}{\mathrm{n}} \sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{E}\left({\mathrm{Y}_{\mathrm{i}}}^2\right) - \left(\frac{1}{\mathrm{n}} \sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{E}(\mathrm{Y}_{\mathrm{i}})\right)^2 \\
&= \frac{1}{\mathrm{n}} \sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{E}(\mathrm{Z}^2) + \frac{1}{\mathrm{n}} \mathrm{A}^2 \sum_{\mathrm{i}=1}^{\mathrm{n}} \left(\lambda_{\mathrm{i}} - \bar{\lambda}\right)^2 - \mu^2 \\
&= \frac{1}{\mathrm{n}} \sum_{\mathrm{i}=1}^{\mathrm{n}} \left[\mathrm{Var}(\mathrm{Z}) + \left(\mathrm{E}(\mathrm{Z})\right)^2\right] + \frac{(\mathrm{n}-1)(\sigma^2-\mathrm{c})}{\mathrm{n}} - \mu^2 \\
&= \frac{1}{\mathrm{n}} \sum_{\mathrm{i}=1}^{\mathrm{n}} \left[\frac{(\sigma^2+(\mathrm{n}-1)\mathrm{c})}{\mathrm{n}} + \mu^2\right] + \frac{(\mathrm{n}-1)(\sigma^2-\mathrm{c})}{\mathrm{n}} - \mu^2 \\
&= \frac{\sigma^2+(\mathrm{n}-1)\mathrm{c}}{\mathrm{n}} + \mu^2 + \frac{(\mathrm{n}-1)(\sigma^2-\mathrm{c})}{\mathrm{n}} - \mu^2 = \sigma^2.
\end{aligned}
$$

Όμοια προκύπτει ότι $\mathrm{cov}\left(\mathrm{Y}_{\pi(1)}, \mathrm{Y}_{\pi(2)}\right) = \mathrm{c}$.

Επομένως, δείξαμε ότι $\mathrm{E}(\mathrm{X}_{\mathrm{i}}) = \mu$, $\mathrm{Var}(\mathrm{X}_{\mathrm{i}}) = \sigma^2$ (για κάθε i), $\mathrm{cov}\left(\mathrm{X}_{\mathrm{i}}, \mathrm{X}_{\mathrm{j}}\right) = \mathrm{c}$ (για κάθε i, j) και **X** είναι ανταλλάξιμο.

Θα δείξω ότι για το συγκεκριμένο διάνυσμα **X,** πετυχαίνεται το φράγμα ως ισότητα.

Πρώτα παρατηρούμε ότι

$$
\left(\mathrm{X}_1, \mathrm{X}_{2,} \dots, \mathrm{X}_{\mathrm{n}}\right)' = \left(\mathrm{Y}_{\pi(1)}, \mathrm{Y}_{\pi(2)}, \dots, \mathrm{Y}_{\pi(\mathrm{n})}\right)'
$$

και επειδή $\mathrm{Y}_1 \le \mathrm{Y}_2 \le \cdots \le \mathrm{Y}_{\mathrm{n}}$ (από κατασκευή) έχουμε:

$$
\left(\mathrm{X}_{1:\mathrm{n}}, \mathrm{X}_{2:\mathrm{n},} \dots, \mathrm{X}_{\mathrm{n}:\mathrm{n}}\right)' = (\mathrm{Y}_1, \mathrm{Y}_2, \dots, \mathrm{Y}_{\mathrm{n}})'.
$$

Τώρα

$$\left|\sum_{i=1}^{n}\lambda_i(E(X_{i:n})-\mu)\right| = \left|\sum_{i=1}^{n}\lambda_i(E(Y_i)-\mu)\right|$$

$$= \left|\sum_{i=1}^{n}\lambda_i(E(Y_i)-\mu)-\sum_{i=1}^{n}\bar{\lambda}(E(Y_i)-\mu)\right|$$

$$= \left|\sum_{i=1}^{n}(\lambda_i-\bar{\lambda})(E(Y_i)-\mu)\right|$$

$$= \left|\sum_{i=1}^{n}(\lambda_i-\bar{\lambda})\big(E(Z)+A(\lambda_i-\bar{\lambda})-\mu\big)\right|$$

$$= \left|\sum_{i=1}^{n}(\lambda_i-\bar{\lambda})\big(A(\lambda_i-\bar{\lambda})-\mu\big)\right|$$

$$= \left|\sum_{i=1}^{n}(\lambda_i-\bar{\lambda})\left(\left(\frac{(n-1)(\sigma^2-c)}{\sum_{i=1}^{n}(\lambda_i-\bar{\lambda})^2}\right)^{1/2}(\lambda_i-\bar{\lambda})\right)\right|$$

$$= \left|\sum_{i=1}^{n}\left(\left(\frac{(n-1)(\sigma^2-c)}{\sum_{i=1}^{n}(\lambda_i-\bar{\lambda})^2}\right)^{1/2}(\lambda_i-\bar{\lambda})^2\right)\right|$$

$$= \left|\left(\frac{(n-1)(\sigma^2-c)}{\sum_{i=1}^{n}(\lambda_i-\bar{\lambda})^2}\right)^{1/2}\sum_{i=1}^{n}(\lambda_i-\bar{\lambda})^2\right|$$

$$= \left|\big((n-1)(\sigma^2-c)\big)^{1/2}\Big(\sum_{i=1}^{n}(\lambda_i-\bar{\lambda})^2\Big)^{1/2}\right|$$

$$= \big((n-1)(\sigma^2-c)\big)^{1/2}\Big(\sum_{i=1}^{n}(\lambda_i-\bar{\lambda})^2\Big)^{1/2}.$$

Τελικά

$$\left|\sum_{i=1}^{n}\lambda_i(E(X_{i:n})-\mu)\right| = \big((n-1)(\sigma^2-c)\big)^{1/2}\Big(\sum_{i=1}^{n}(\lambda_i-\bar{\lambda})^2\Big)^{1/2},$$

και το φράγμα (2.13) γίνεται ισότητα.

Στην περίπτωση που η ακολουθία των λ_i είναι μη αύξουσα αποδεικνύεται με τον ίδιο τρόπο ■

2.3. Άνω φράγματα για την αναμενόμενη τιμή της μέγιστης παρατήρησης σε εξαρτημένο δείγμα τυχαίων μεταβλητών.

Στην ενότητα αυτή θα δείξουμε πως μπορεί το παραπάνω Λήμμα, να επεκταθεί και στην περίπτωση που δεν έχουμε ανταλλάξιμες τυχαίες μεταβλητές. Η γενική περίπτωση δίνεται από το παρακάτω Θεώρημα

Θεώρημα 2.3.2. Έστω ένα τυχαίο διάνυσμα $\mathbf{X} = (X_1, X_2, \dots, X_n)'$ με $E(X_i) = \mu_i$ και πίνακα συνδιακύμανσης (σ_{ij}), όπου i, j = 1, 2, …,n. Τότε για αυθαίρετες τιμές των πραγματικών σταθερών $\lambda_1, \lambda_2, \dots, \lambda_n$, ισχύει η ανισότητα

$$\left|\sum_{i=1}^{n} \lambda_i (E(X_{i:n}) - \bar{\mu})\right| \leq \left(\sum_{i=1}^{n} (\lambda_i - \bar{\lambda})^2\right)^{1/2} \left(\sum_{i=1}^{n} \left\{(\mu_i - \bar{\mu})^2 + \sigma_i^2\right\} - n\mathrm{Var}\bar{X}\right)^{1/2}, \tag{2.15}$$

όπου $X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{n:n}$ είναι οι διατεταγμένες παρατηρήσεις του διανύσματος **X,** $\sigma_{ii} = \sigma_i^2$, i = 1, 2, …,n, $\bar{\mu} = \frac{1}{n}\sum_{i=1}^{n} \mu_i$, $\bar{\lambda} = \frac{1}{n}\sum_{i=1}^{n} \lambda_i$ και $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$.

Απόδειξη

Για n = 1, η ανισότητα είναι προφανής.

Για $n \geq 2$, θέτουμε

$$(Y_1, Y_2, \dots, Y_n)' = (X_{\pi(1)}, X_{\pi(2)}, \dots, X_{\pi(n)})',$$

όπου $(\pi(1), \pi(2), \dots, \pi(n))'$ είναι ένα διάνυσμα ανεξάρτητο του **X**, και έχει κατάνομή όπως στο Λήμμα 1.4.1. Οι Y_i, i = 1, 2, …,n είναι ανταλλάξιμες τ.μ.[7], με

$$\begin{aligned} E(Y_1) &= E(X_{\pi(1)}) \\ &= \sum_{i=1}^{n} E(X_{\pi(1)} | \pi(1) = i) \mathbb{P}(\pi(1) = i) \\ &= \frac{1}{n}\sum_{i=1}^{n} E(X_i) \\ &= \bar{\mu}. \end{aligned}$$

Εφαρμόζοντας με τον ίδιο τρόπο το Θεώρημα Ολικής Πιθανότητας και χρησιμοποιώντας τις σχέσεις

$$\mathrm{Var}(Y_1) = E(Y_1^2) - E(Y_1)^2 \;\; \text{και} \;\; \mathrm{Cov}(Y_1, Y_2) = E(Y_1 Y_2) - E(Y_1)E(Y_2)$$

για τη διακύμανση και τη συνδιακύμανση αντίστοιχα, προκύπτουν οι σχέσεις

$$\mathrm{Var}(Y_1) = \frac{1}{n}\sum_{i=1}^{n} \left\{(\mu_i - \bar{\mu})^2 + \sigma_i^2\right\}$$

και

$$\mathrm{Cov}(Y_1, Y_2) = \frac{1}{n(n-1)}\left(2\sum_{1\leq i<j\leq n} \sigma_{ij} - \sum_{i=1}^{n} (\mu_i - \bar{\mu})^2\right)$$

7 Είναι ανταλλάξιμες γιατί το διάνυσμα $(\pi(1), \pi(2), \dots, \pi(n))'$ είναι μεταθέσεις των 1, 2, …, n.

$$= \frac{n}{n-1}\mathrm{Var}(\overline{X}) - \frac{1}{n-1}\mathrm{Var}(Y_1).$$

Εφαρμόζοντας τους παραπάνω τύπους στην ανισότητα (2.13)., και χρησιμοποιώντας πλέον την υπόθεση της ανταλλαξιμότητας για τις Y_i, προκύπτει η ανισότητα

$$\left|\sum_{i=1}^{n}\lambda_i\big(E(X_{i:n}) - E(Y_1)\big)\right| \leq \left(\sum_{i=1}^{n}(\lambda_i - \overline{\lambda})^2\right)^{1/2}\Big((n-1)\big(\mathrm{Var}(Y_1) - \mathrm{Cov}(Y_1, Y_2)\big)\Big)^{1/2}$$

$$= \left(\sum_{i=1}^{n}(\lambda_i - \overline{\lambda})^2\right)^{1/2}\big(n\mathrm{Var}(Y_1) - n\mathrm{Var}(\overline{X})\big)^{1/2}$$

$$= \left(\sum_{i=1}^{n}(\lambda_i - \overline{\lambda})^2\right)^{1/2}\left(\sum_{i=1}^{n}\left\{(\mu_i - \overline{\mu})^2 + \sigma_i^2\right\} - n\mathrm{Var}(\overline{X})\right)^{1/2}.$$

■

Παρατηρήσεις

1. Το φράγμα που προκύπτει όπως φαίνεται στην (2.15) είναι καλύτερο από το φράγμα των Arnold – Goeneveld στη σχέση (2.1). Τα φράγματα αυτά συμπίπτουν μόνο στην ειδική περίπτωση που η $\overline{X}$ είναι σταθερή.
2. Αν στη σχέση (2.13) θέσω $\lambda_n = 1$ και $\lambda_1 = \lambda_2 = \cdots = \lambda_{n-1} = 0$, προκύπτει

$$E(X_{n:n}) \leq \left[\sum_{i=1}^{n-1}\left(-\frac{1}{n}\right)^2 + \left(1 - \frac{1}{n}\right)^2\right]^{1/2}\big((n-1)(\sigma^2 - c)\big)^{1/2}$$

$$= \left[\frac{n-1}{n^2} + \frac{(n-1)^2}{n^2}\right]^{1/2}\big((n-1)(\sigma^2 - c)\big)^{1/2}$$

$$= \left[\frac{n^2-n}{n^2}\right]^{1/2}\big((n-1)(\sigma^2 - c)\big)^{1/2}$$

$$= \left[\frac{n(n-1)}{n^2}\right]^{1/2}\big((n-1)(\sigma^2 - c)\big)^{1/2}$$

$$= \frac{n-1}{\sqrt{n}}(\sigma^2 - c)^{1/2}.$$

Δηλαδή παίρνουμε την ανισότητα (2.12) η οποία μας δίνει ένα φράγμα για την αναμενόμενη τιμή της μέγιστης διατεταγμένης παρατήρησης. Γνωρίζουμε ότι αυτό επιτυγχάνεται πάντα αφού η $(\lambda_1, \lambda_2, \ldots, \lambda_n) = (0, 0, \ldots, 0, 1)$ είναι αύξουσα.

Πότε ένα ανταλλάξιμο διάνυσμα επεκτείνεται σε άπειρη ακολουθία ανταλλάξιμων τυχαίων μεταβλητών.

Έστω οι ανταλλάξιμες τ.μ $X_1, X_2, \ldots, X_n$ με $E(X_i) = \mu$, $\mathrm{Var}(X_i) = \sigma^2$ και $\mathrm{cov}(X_i, X_j) = c$. Τότε

$$\mathrm{Var}(\overline{X}) = \frac{1}{n^2}\mathrm{Var}\left(\sum_{i=1}^{n} X_i\right)$$

$$= \frac{1}{n^2}\sum_{i=1}^{n} \mathrm{Var}(X_i) + \frac{2}{n^2}\sum\sum_{1\le i<j\le n} \mathrm{cov}(X_i, X_j)$$

$$= (n\sigma^2 + n(n-1)c)/n^2$$

$$= (\sigma^2 + (n-1)c)/n.$$

Έστω ότι η $(Xi)_{i=1}^{n}$ επεκτείνεται σε άπειρη ακολουθία $X_1, X_2, \ldots$ τότε

$\mathrm{Var}(\overline{X}) = (\sigma^2 + (n-1)c)/n \longrightarrow c$, όταν $n \longrightarrow \infty$. Επειδή $\mathrm{Var}(\overline{X}) \ge 0$, αν ήταν $c < 0$, δε θα μπορούσε να επεκταθεί σε άπειρη ακολουθία από ανταλλάξιμες τ.μ. Αν $c \ge 0$ δεν μπορούμε εύκολα να γνωρίζουμε αν επεκτείνεται ένα δεδομένο σύνολο ανταλλαξί-μων $(Xi)_{i=1}^{n}$ σε άπειρη ακολουθία από ανταλλάξιμες τ.μ.

Αν μια ακολουθία $\{X_i\}$ είναι ανταλλάξιμη, τότε υπάρχει[8] μια τ.μ V τέτοια ώστε οι X_i, i=1, 2, …, δεδομένης της V να είναι ανεξάρτητες και ισόνομες. Η σχέση (1.1) οδηγεί στην ανισότητα:

$$E[X_{n:n}|V] \le E[X_1|V] + \frac{n-1}{\sqrt{2n-1}}\sqrt{\mathrm{Var}[X_1|V]} \quad \sigma.\pi. \tag{2.16}$$

Παίρνοντας μέσες τιμές στην (2.16),

$$E\big[E[X_{n:n}|V]\big] \le E\left[E[X_1|V] + \frac{n-1}{\sqrt{2n-1}}\sqrt{\mathrm{Var}[X_1|V]}\right].$$

Επομένως,

$$E[X_{n:n}] \le E\big[E[X_1|V]\big] + \frac{n-1}{\sqrt{2n-1}} E\left[\sqrt{\mathrm{Var}[X_1|V]}\right]$$

$$= E[X_1] + \frac{n-1}{\sqrt{2n-1}} E\left[\sqrt{\mathrm{Var}[X_1|V]}\right]. \tag{2.17}$$

Από την ανισότητα Jensen, έχουμε ότι $E\left[\sqrt{\mathrm{Var}[X_1|V]}\right] \le \sqrt{E\big[\mathrm{Var}[X_1|V]\big]}$ (διότι η $f(x) = \sqrt{x}$ είναι κοίλη συνάρτηση).

Επίσης έχουμε ότι

$$E\big[\mathrm{Var}[X_1|V]\big] = \mathrm{Var}(X_1) - \mathrm{Var}(E[X_1|V])$$

$$= \mathrm{Var}(X_1) - \left[E\big(E(X_1|V)\big)^2 - \left[E\big(E(X_1|V)\big)\right]^2\right]$$

$$= \mathrm{Var}(X_1) - \left[E\big(E(X_1|V)E(X_1|V)\big) - E\big(E(X_1|V)\big)E\big(E(X_1|V)\big)\right].$$

---

[8] Θεώρημα DeFinetti: Έστω $X_1, X_2, \ldots$, μια ακολουθία από ανταλλάξιμες τυχαίες μεταβλητές. Τότε για κάθε πραγματικό αριθμό x, υπάρχει μια τυχαία μεταβλητή $Y(x) = Y(x,\omega)$ με $0 \le Y(x) \le 1$ και τέτοια ώστε $\mathbb{P}\big(X_{i_s} < x_{i_s}, 1 \le s \le k\big) = E\big(Y(X_1), Y(X_2), \ldots, Y(X_k)\big)$.Επιπρόσθετα, η $Y(x,\omega)$ είναι μια συνάρτηση ως προς x για κάθε σημείο ω.

Όμως $X_1|V =_d X_2|V$,

$$E\big[Var[X_1|V]\big] = Var(X_1) - \big[E\big(E(X_1|V)E(X_2|V)\big) - E\big(E(X_1|V)\big)E\big(E(X_2|V)\big)\big].$$

Οι X_1 και X_2 κάτω από τη δέσμευση της V είναι ανεξάρτητες. Άρα

$$\begin{aligned} Cov(X_1, X_2|V) = 0 &\Rightarrow E\big(Cov(X_1, X_2|V)\big) = 0 \\ &\Rightarrow E\big(E(X_1X_2|V) - E(X_1|V)E(X_2|V)\big) = 0 \\ &\Rightarrow E\big(E(X_1X_2|V)\big) - E\big(E(X_1|V)E(X_2|V)\big) = 0 \\ &\Rightarrow E(X_1X_2) = E\big(E(X_1|V)E(X_2|V)\big). \end{aligned}$$

Οπότε προκύπτει ότι

$$\begin{aligned} E\big[Var[X_1|V]\big] &= Var(X_1) - [E(X_1X_2) - E(X_1)E(X_2)] \\ &= \sigma^2 - c \end{aligned}$$

και η (2.17) γράφεται:

$$E[X_{n:n}] \leq \mu + \frac{n-1}{\sqrt{2n-1}}\sqrt{\sigma^2 - c}\,. \tag{2.18}$$

Η ανισότητα (2.18) μας δίνει καλύτερο άνω φράγμα από την (2.12). Δεν είναι βέλτιστο άνω φράγμα για την $E[X_{n:n}]$ αλλά είναι το καλύτερο όταν η X_1, X_2, … μπορεί να επεκταθεί σε ανταλλάξιμη ακολουθία.

Καταλήξαμε λοιπόν στο εξής κριτήριο ανταλλαξιμότητας:

Θεώρημα 2.3.3 Αν $\big(X_1, X_2, \dots, X_n\big)$ είναι ανταλλάξιμο τυχαίο διάνυσμα με

$$E(X_i) = \mu,\ Var(X_i) = \sigma^2 \text{ και } cov\big(X_i, X_j\big) = c \geq 0,$$

και αν

$$E[X_{n:n}] \leq \mu + \frac{n-1}{\sqrt{2n-1}}\sqrt{\sigma^2 - c},$$

τότε το $\big(X_1, X_2, \dots, X_n\big)$ δεν επεκτείνεται σε άπειρη ακολουθία ανταλλάξιμων τ.μ.

∎

ΚΕΦΑΛΑΙΟ 3

ΒΕΛΤΙΣΤΟ ΑΝΩ ΦΡΑΓΜΑ ΓΙΑ ΤΗΝ ΑΝΑΜΕΝΟΜΕΝΗ ΤΙΜΗ ΤΟΥ ΜΕΓΙΣΤΟΥ.

3.1. Βέλτιστο άνω φράγμα για την αναμενόμενη τιμή του μεγίστου.

Στο κεφάλαιο αυτό ενδιαφερόμαστε για τη μέγιστη δυνατή τιμή της μέσης τιμής του μεγίστου $X_{n:n} = \max\{X_1, X_2, \dots, X_n\}$, όταν όλες οι τυχαίες μεταβλητές X_i, i=1, 2, …, n του διανύσματος $\boldsymbol{X} = (X_1, X_2, \dots, X_n)$ έχουν γνωστές μέσες τιμές και διασπορές: $\mu_i \in \mathbb{R}$, $\sigma_i > 0$.

Έστω $\mathcal{F}_n(\boldsymbol{\mu}, \boldsymbol{\sigma}) = \mathcal{F}$, η οικογένεια των τυχαίων διανυσμάτων $\mathbf{X} = (X_1, X_2, \dots, X_n)$ με:

$$E(X_i) = \mu_i, \text{και } Var(X_i) = \sigma_i^2, \quad i = 1,2,3,\dots,n. \tag{3.1}$$

Το βασικό αποτέλεσμα που θα αποδείξουμε γράφεται ως εξής:

<u>**Θεώρημα 3.1.1.**</u> **(Bertsimas, Natarajan and Teo (2006))** Αν $\mathbf{X} \in \mathcal{F}_n(\boldsymbol{\mu}, \boldsymbol{\sigma})$, τότε

$$E(X_{n:n}) \le t_o + \frac{1}{2}\sum_{i=1}^{n}\left\{\mu_i - t_o + \sqrt{(\mu_i - t_o)^2 + \sigma_i^2}\right\} \tag{3.2}$$

όπου $t = t_o \in \mathbb{R}$ είναι η μοναδική λύση της εξίσωσης

$$\sum_{i=1}^{n} \frac{t-\mu_i}{\sqrt{(t-\mu_i)^2 + \sigma_i^2}} = n - 2, \quad -\infty < t < \infty .$$

<u>**Θεώρημα 3.1.2.**</u> Η ισότητα στην (3.2) χαρακτηρίζει μία (και μόνο μία) κατανομή στον $\mathbb{R}^n$. Συγκεκριμένα η ισότητα στην (3.2) επιτυγχάνεται αν και μόνο αν:

$$\mathbb{P}\begin{bmatrix} X_1 = t_o - \sqrt{(\mu_1 - t_o)^2 + \sigma_1^2} \\ \vdots \\ X_i = t_o + \sqrt{(\mu_i - t_o)^2 + \sigma_i^2} \\ \vdots \\ X_n = t_o - \sqrt{(\mu_n - t_o)^2 + \sigma_n^2} \end{bmatrix} = \frac{1}{2}\left(1 + \frac{\mu_i - t_o}{\sqrt{(\mu_i - t_o)^2 + \sigma_i^2}}\right), \quad i = 1,2,3,\dots,n.$$

(το μοναδικό «+» βρίσκεται στη i – οστή θέση).

<u>**Απόδειξη του Θεωρήματος 3.1.1**</u>

Έστω $t \in \mathbb{R}$, τυχόν.

Είναι:

$$X_i = t + X_i - t \le t + \max\{X_i - t, 0\} = t + (X_i - t)_+ \le t + \sum_{i=1}^{n}(X_i - t)_+.$$

Άρα, αφού το δεξιό μέλος δεν εξαρτάται από το i,

$$X_{n:n} = \max\{X_1, X_2, \dots, X_n\} \le t + \sum_{i=1}^{n}(X_i - t)_+ , \ t \in \mathbb{R}.$$

Προφανώς $(X_i - t)_+ = \frac{(X_i - t) + |X_i - t|}{2}$, οπότε

$$X_{n:n} \le t + \tfrac{1}{2}\sum_{i=1}^{n}(X_i - t) + \tfrac{1}{2}\sum_{i=1}^{n}|X_i - t|. \quad (3.3)$$

Παίρνοντας μέσες τιμές στην (3.3):

$$E(X_{n:n}) \le t + \tfrac{1}{2}\sum_{i=1}^{n}(\mu_i - t) + \tfrac{1}{2}\sum_{i=1}^{n} E|X_i - t|. \quad (3.4)$$

Τώρα έχουμε από την Cauchy – Schwarz:

$$\begin{aligned} E|X_i - t| &\le \sqrt{E(X_i - t)^2} \\ &= \sqrt{(\mu_i - t)^2 + \sigma_i^2}. \quad (3.5) \end{aligned}$$

Συνεπώς, η (3.4) δίνει:

$$E(X_{n:n}) \le t + \tfrac{1}{2}\sum_{i=1}^{n} E(X_i - t) + \tfrac{1}{2}\sum_{i=1}^{n}\sqrt{(\mu_i - t)^2 + \sigma_i^2}. \quad (3.6)$$

Το άνω φράγμα (3.6) μπορεί να υπολογιστεί από τα $\boldsymbol{\mu}$, $\boldsymbol{\sigma^2}$ (που δίνονται) και το t (που είναι αυθαίρετο).

Για να πετύχω <u>το μικρότερο δυνατό</u> άνω φράγμα στην (3.6) θα πρέπει να επιλέξω το κατάλληλο t για να ελαχιστοποιήσω το δεξί μέλος.

Θέτουμε λοιπόν

$$g(t) = t + \frac{1}{2}\sum_{i=1}^{n}\left[(\mu_i - t) + \sqrt{(\mu_i - t)^2 + \sigma_i^2}\right], \qquad t \in \mathbb{R}.$$

Η g(t) είναι γνήσια κυρτή διότι

$$\begin{aligned} \frac{\partial^2}{\partial t^2} g(t) &= \frac{1}{2}\sum_{i=1}^{n}\left[\frac{\partial^2}{\partial t^2}\sqrt{(\mu_i - t)^2 + \sigma_i^2}\right] \\ &= \frac{1}{2}\sum_{i=1}^{n}\frac{\sqrt{(\mu_i - t)^2 + \sigma_i^2} - \frac{(\mu_i - t)^2}{\sqrt{(\mu_i - t)^2 + \sigma_i^2}}}{(\mu_i - t)^2 + \sigma_i^2} \end{aligned}$$

$$\Rightarrow \frac{\partial^2}{\partial t^2} g(t) = \frac{1}{2}\sum_{i=1}^{n}\frac{(\mu_i - t)^2 + \sigma_i^2 - (\mu_i - t)^2}{[(\mu_i - t)^2 + \sigma_i^2]^{3/2}}$$

$$= \frac{1}{2}\sum_{i=1}^{n} \frac{\sigma_i^2}{\left[(\mu_i - t)^2 + \sigma_i^2\right]^{3/2}} > 0.$$

Αφού η g είναι κυρτή, για να δείξουμε ότι έχει ελάχιστο, αρκεί να δείξουμε ότι $\lim_{t\to\pm\infty} g(t) = +\infty$ (για $n \geq 2$).

Απόδειξη ότι $\lim_{t\to\pm\infty} g(t) = +\infty$

Είναι $|x| \leq \sqrt{x^2 + \sigma_i^2}$, για κάθε $x \in \mathbb{R}$.

Άρα

$$|\mu_i - t| \leq \sqrt{(\mu_i - t)^2 + \sigma_i^2},$$

οπότε

$$\mu_i - t \geq -\sqrt{(\mu_i - t)^2 + \sigma_i^2}$$

$$\Longrightarrow \mu_i - t + \sqrt{(\mu_i - t)^2 + \sigma_i^2} \geq 0.$$

$$g(t) = t + \tfrac{1}{2}\textstyle\sum_{i=1}^{n}\left[(\mu_i - t) + \sqrt{(\mu_i - t)^2 + \sigma_i^2}\right] \geq t.$$

Οπότε $g(t) \longrightarrow \infty$ (για $t \longrightarrow \infty$).

Αν $t \longrightarrow -\infty$ θέτω $y = -t$ ($y \longrightarrow \infty$)

$$g(t) = g(-y) = -y + \frac{1}{2}\sum_{i=1}^{n}\left[(\mu_i + y) + \sqrt{(\mu_i + y)^2 + \sigma_i^2}\right]$$

$$\geq -y + \tfrac{1}{2}\textstyle\sum_{i=1}^{n} 2(\mu_i + y)$$

$$=(n-1)y + \textstyle\sum_{i=1}^{n}\mu_i \longrightarrow \infty \quad \text{(καθώς } y \longrightarrow \infty \text{ και } n \geq 2\text{).}$$

Αν $n = 1$ έχουμε

$$g(-y) = -y + \tfrac{1}{2}\left[(\mu_1 + y) + \sqrt{(\mu_1 + y)^2 + \sigma_1^2}\right]$$

$$= \frac{1}{2}\mu_1 + \frac{1}{2}\sqrt{(\mu_1 + y)^2 + \sigma_1^2} - \frac{1}{2}y$$

Τώρα:

$$\sqrt{(\mu_1 + y)^2 + \sigma_1^2} - y = \frac{(\mu_1+y)^2+\sigma_1^2-y^2}{y+\sqrt{(\mu_1+y)^2+\sigma_1^2}}$$

$$= \frac{\mu_1^2+\sigma_1^2+2y\mu_1}{y+\sqrt{(\mu_1+y)^2+\sigma_1^2}}$$

$$= \frac{\mu_1+\sigma_1{}^2}{y+\sqrt{(\mu_1+y)^2+\sigma_1{}^2}} + 2\mu_1 \frac{1}{1+\sqrt{\frac{\sigma_1{}^2}{y}+\left(\frac{\mu_1}{y}+1\right)^2}} \longrightarrow 0 + \frac{2\mu_1}{1} = 2\mu_1 \text{ (όχι 0)}$$

διότι $\frac{\mu_1+\sigma_1{}^2}{y+\sqrt{(\mu_1+y)^2+\sigma_1{}^2}} \longrightarrow 0$ και $\frac{1}{1+\sqrt{\frac{\sigma_1{}^2}{y}+\left(\frac{\mu_1}{y}+1\right)^2}} \longrightarrow 1$ καθώς $y \longrightarrow \infty$.

Άρα, $g(-\infty) = -\frac{1}{2}\mu_1$ (όχι 0) και $g(\infty) = 0$. Και πάλι έχει ελάχιστο.

Υπάρχει λοιπόν μοναδικό t_o τέτοιο ώστε

$$g(t) \geq g(t_o), \qquad \text{για κάθε } t \in \mathbb{R}.$$

Επίσης $t \neq t_o \Longrightarrow g(t) > g(t_o)$ (διότι η g' είναι γνησίως αύξουσα αφού $g'' > 0$).

Συνεπώς, η εξίσωση $g'(t) = 0$ έχει μοναδική λύση t_o, και

$$E(X_{n:n}) \leq g(t_o) = t_o + \frac{1}{2}\sum_{i=1}^{n}\left[(\mu_i - t_o) + \sqrt{(\mu_i - t_o)^2 + \sigma_i{}^2}\right]$$

όπου t_o: $g'(t_o) = 0$, δηλαδή

$$\sum_{i=1}^{n} \frac{t_o-\mu_i}{\sqrt{(t_o-\mu_i)^2+\sigma_i{}^2}} = n - 2. \tag{3.7}$$

■

Απόδειξη του Θεωρήματος 3.1.2

Πηγαίνουμε ανάποδα στις ανισότητες ώστε να δούμε αν μπορούν όλες να γίνουν ισότητες. Η τελευταία ανισότητα ήταν

$$E|X_i - t_o| \leq \sqrt{E(X_i - t_o)^2}, \; i = 1,2,3,\dots,n \tag{3.8}$$

Η (3.8) γίνεται ισότητα αν και μόνο αν $\mathrm{Var}(|X_i - t_o|) = 0$. Επομένως, αρκεί $\mathbb{P}(|X_i - t_o| = c_i) = 1, \;\; i = 1,2,3,\dots,n$ για κάποιες σταθερές $c_i > 0$. (πρέπει $c_i > 0$, αλλιώς θα ήταν $\mathbb{P}(X_i = t_o) = 1 \Longrightarrow \mathrm{Var}(X_i) = 0$).

Όμως,

$\mathbb{P}(B_i) = 1$, για $i = 1,2,3,\dots,n \Leftrightarrow \mathbb{P}(\bigcap_{i=1}^{n} B_i) = 1$ όπου $B_i \in \mathcal{A}$ ($\sigma -$ άλγεβρα).
Άρα για να έχω ισότητα στην (3.8) πρέπει και αρκεί

$$\mathbb{P}[|X_1 - t_o| = c_1, \dots, |X_n - t_o| = c_n] = 1 \tag{3.9}$$

για κάποιες σταθερές $c_1, c_2, \dots, c_n$ με $c_i > 0$.

Πάμε τώρα στην προηγούμενη ανισότητα:

Από την (3.3) έχουμε ότι:

$$X_{n:n} \leq t_o + \frac{1}{2}\sum_{i=1}^{n}(X_i - t_o) + \frac{1}{2}\sum_{i=1}^{n}|X_i - t_o|.$$

Από την οποία πήραμε την (3.4):

$$\mathrm{E}(\mathrm{X}_{n:n}) \le t_o + \frac{1}{2}\sum_{i=1}^{n}(\mu_i - t_o) + \frac{1}{2}\sum_{i=1}^{n} \mathrm{E}|\mathrm{X}_i - t_o|.$$

Όμως αυτή ισοδυναμεί με:

$$\mathrm{E}\left(\mathrm{X}_{n:n} - \left(t_o + \frac{1}{2}\sum_{i=1}^{n}\big((\mathrm{X}_i - t_o) - |\mathrm{X}_i - t_o|\big)\right)\right) \le 0.$$

Αυτή είναι της μορφής $\mathrm{E}(\mathrm{Y}) \le 0$, όπου

$$\mathrm{Y} = \mathrm{X}_{n:n} - \left(t_o + \tfrac{1}{2}\textstyle\sum_{i=1}^{n}\big((\mathrm{X}_i - t_o) - |\mathrm{X}_i - t_o|\big)\right) \le 0 \text{ από την (3.3).}$$

Οι σχέσεις $\mathrm{E}(\mathrm{Y}) \le 0$ και $\mathrm{Y} \le 0$ συνεπάγονται ότι αν $\mathrm{E}(\mathrm{Y}) = 0$ τότε $\mathbb{P}(\mathrm{Y} = 0) = 1$. Άρα για να έχω ισότητα στην (3.3) πρέπει και αρκεί

$$\mathrm{X}_{n:n} = t_o + \tfrac{1}{2}\textstyle\sum_{i=1}^{n}[(\mathrm{X}_i - t_o) + |\mathrm{X}_i - t_o|] = t_o + \sum_{i=1}^{n}(\mathrm{X}_i - t_o)_+ \qquad (*)$$

με πιθανότητα 1.

Παρατηρούμε ότι η $(*)$ ισοδυναμεί με

$$\mathrm{X}_{1:n} \le \cdots \le \mathrm{X}_{n-1:n} \le t_o \le \mathrm{X}_{n:n}. \text{ [9]} \qquad (**)$$

Άρα πρέπει ακριβώς n – 1 τ.μ. από τις $\mathrm{X}_1, \mathrm{X}_2, \dots, \mathrm{X}_n$ να είναι μικρότερες ή ίσες του t_o, και μία μεγαλύτερη ή ίση του t_o.

Αφού $\mathbb{P}(\mathrm{X}_1 \in \{t_o - c_1, t_o + c_1\}, \dots, \mathrm{X}_n \in \{t_o - c_n, t_o + c_n\}) = 1$ λόγω της (3.6), προκύπτει ότι οι μόνες δυνατές τιμές του διανύσματος $(\mathrm{X}_1, \mathrm{X}_2, \dots, \mathrm{X}_n)$ που ικανοποιούν τη $(**)$ είναι οι:

$$\mathrm{S} = \bigcup_{i=1}^{n}\{t_o - c_1, \dots, t_o - c_{i-1},\ t_o + c_i,\ t_o - c_{i+1}, \dots,\ t_o - c_n\}$$

[9] Για να ισχύει η ισότητα, θα πρέπει να ισχύει $\mathrm{X}_{n:n} = t_o + \sum_{i=1}^{n}(\mathrm{X}_i - t_o)_+$. Δηλαδή θα πρέπει να υπάρχει κάποιο i_o τέτοιο ώστε $(\mathrm{X}_{i_o} - t_o)_+ = \mathrm{X}_{i_o} - t_o$ και για όλα τα υπόλοιπα $i \neq i_o$ να ισχύει $(\mathrm{X}_i - t_o)_+ = 0$. Οπότε αν $\mathrm{X}_{i_o} - t_o \ge 0$ τότε $\mathrm{X}_{n:n} = t_o + \mathrm{X}_{i_o} - t_o$. Το οποίο σημαίνει $\mathrm{X}_{n:n} \ge t_o$. Για όλα τα υπόλοιπα i ισχύει $\mathrm{X}_i - t_o \le 0$, άρα και για την προτελευταία διατεταγμένη παρατήρηση $\mathrm{X}_{n-1:n} \le t_o$.

Αντίστροφα, έστω ότι ισχύει $\mathrm{X}_{n-1:n} \le t_o \le \mathrm{X}_{n:n}$, θα ισχύει η ισότητα στην (3.1), διότι αφού $\mathrm{X}_{n-1:n} \le t_o$ τότε $\mathrm{X}_i \le t_o$ για όλα τα i που είναι μικρότερα ή ίσα από την προτελευταία διατεταγμένη παρατήρηση. Δηλαδή $(\mathrm{X}_i - t_o)_+ = 0$ για όλα τα i εκτός της μέγιστης διατεταγμένης παρατήρησης. Τέλος αφού $t_o \le \mathrm{X}_{n:n}$, $(\mathrm{X}_{n:n} - t_o)_+ = \mathrm{X}_{n:n} - t_o$. Άρα $t_o + \sum_{i=1}^{n}(\mathrm{X}_i - t_o)_+ = t_o + \mathrm{X}_{n:n} - t_o = \mathrm{X}_{n:n}$.

που περιέχει n σημεία (στο n – διάστατο χώρο).

Άρα το $\mathbf{X}$ είναι υποχρεωμένο να πάρει μία από τις τιμές του συνόλου $S \subseteq \mathbb{R}^n$, με $N(S) = n$.

Έστω ότι για κάποιο i, $\mathbb{P}(X_i = t_o + c_i) = p_i$. Τότε $\mathbb{P}(X_i = t_o - c_i) = 1 - p_i$. Επομένως, αφού η X_i είναι δίτιμη με τιμές $t_o \pm c_i$, έχουμε

$$\begin{aligned} E(X_i) = \mu_i \quad &\Leftrightarrow (t_o + c_i)p_i + (t_o - c_i)(1 - p_i) = \mu_i \\ &\Leftrightarrow 2c_i p_i + t_o - c_i = \mu_i \\ &\overset{c_i \neq 0}{\Longleftrightarrow} p_i = \frac{\mu_i + c_i - t_o}{2c_i}. \end{aligned}$$

$$\begin{aligned} \mathrm{Var}(X_i) = {\sigma_i}^2 &\Leftrightarrow E({X_i}^2) - \left(E(X_i)\right)^2 = {\sigma_i}^2 \\ &\Leftrightarrow E({X_i}^2) = {\mu_i}^2 + {\sigma_i}^2 \\ &\Leftrightarrow (t_o + c_i)^2 p_i + (t_o - c_i)^2(1 - p_i) = {\mu_i}^2 + {\sigma_i}^2 \\ &\Leftrightarrow {c_i}^2 = {\mu_i}^2 + {\sigma_i}^2 - 2t_o\mu_i + {t_o}^2 \\ &\Leftrightarrow {c_i}^2 = (\mu_i - t_o)^2 + {\sigma_i}^2 \\ &\overset{c_i > 0}{\Longleftrightarrow} c_i = \sqrt{(\mu_i - t_o)^2 + {\sigma_i}^2}\,. \end{aligned}$$

Άρα,

$$c_i = \sqrt{(\mu_i - t_o)^2 + {\sigma_i}^2}, \qquad p_i = \frac{1}{2}\left(1 - \frac{t_o - \mu_i}{\sqrt{(\mu_i - t_o)^2 + {\sigma_i}^2}}\right), \text{ για } i = 1,\ 2,\ 3, \dots, n$$

Βλέπουμε ότι το σύνολο S (το support του τ.δ. $\mathbf{X}$) καθορίζεται πλήρως, έχει n ακριβώς σημεία, και πρέπει

$$\mathbb{P}(\mathbf{X} = \mathbf{s_i}) = p_i \qquad (***)$$

με p_i όπως παραπάνω και $\mathbf{s}_i = \begin{pmatrix} t_o - c_1 \\ \vdots \\ t_o + c_i \\ \vdots \\ t_o - c_n \end{pmatrix}$.

Το τελευταίο που μένει να δείξουμε είναι ότι η (∗∗∗) ορίζει μια συνάρτηση κατανομής. Γι' αυτό αρκεί $p_i > 0$ [10]και $\sum_{i=1}^n p_i = 1$. Όμως η σχέση $\sum_{i=1}^n p_i = 1$ ισοδυναμεί με

10

$0 < p_i \Leftrightarrow 0 < \frac{\mu_i - t_o + \sqrt{(\mu_i - t_o)^2 + {\sigma_i}^2}}{2\sqrt{(\mu_i - t_o)^2 + {\sigma_i}^2}} \Leftrightarrow 0 < \mu_i - t_o + \sqrt{(\mu_i - t_o)^2 + {\sigma_i}^2} \Leftrightarrow$

$$\frac{1}{2}\sum_{i=1}^{n}\frac{t_o-\mu_i}{\sqrt{(\mu_i-t_o)^2+\sigma_i^2}}-1=\sum_{i=1}^{n}\frac{1}{2}\quad\Leftrightarrow\sum_{i=1}^{n}\frac{t_o-\mu_i}{\sqrt{(\mu_i-t_o)^2+\sigma_i^2}}=n-2$$

$$\Leftrightarrow g'(t_o)=0 \text{ (που ισχύει).}$$

Αποδείξαμε λοιπόν, ότι η <u>μοναδική</u> n – διάστατη κατανομή για το **X** που δίνει ισότητα στο φράγμα

$$E(X_{n:n})\le g(t_o)$$

είναι η

$$\mathbb{P}(\mathbf{X}=\mathbf{s_i})=p_i,\quad i=1,2,3,\dots,n$$

όπου
$$p_i=\frac{1}{2}\left(1-\frac{t_o-\mu_i}{\sqrt{(\mu_i-t_o)^2+\sigma_i^2}}\right)>0,\quad i=1,\ 2,\ 3,\dots,n$$

και
$$\mathbf{s}_i=\begin{pmatrix}t_o-c_1\\ \vdots\\ t_o+c_i\\ \vdots\\ t_o-c_n\end{pmatrix},\quad i=1,\ 2,\ 3,\dots,n$$

με
$$c_i=\sqrt{(\mu_i-t_o)^2+\sigma_i^2}>0,\ \ i=1,\ 2,\ 3,\dots,n$$

■

3.2. Το πρόβλημα της μεγιστοποίησης της $\mathbf{E(\max\{\beta,X_1,\dots,X_n\})}$ $(\boldsymbol{\beta}\in\mathbb{R})$

Αρχικά θα εξετάσουμε την απλή περίπτωση που n = 1.

- <u>Το πρόβλημα της μεγιστοποίησης της $E(\max\{\beta,X_1\})$ $(\beta\in\mathbb{R})$</u>

Για n = 1 ισχύει η παρακάτω (τετριμμένη) πρόταση

<u>Πρόταση 3.2.1</u>

$$E(\max\{X,c\})\le\frac{1}{2}\left\{(\mu+c)+\sqrt{(\mu-c)^2+\sigma^2}\right\}.$$

$$0<\mu_i-t_o+\sqrt{(\mu_i-t_o)^2+\sigma_i^2}\Leftrightarrow$$

$$-\mu_i+t_o<\sqrt{(\mu_i-t_o)^2+\sigma_i^2}$$

Πράγματι ισχύει, διότι:

$$t_o-\mu_i\le|\mu_i-t_o|=\sqrt{(\mu_i-t_o)^2}<\sqrt{(\mu_i-t_o)^2+\sigma_i^2}$$

Η ισότητα επιτυγχάνεται αν και μόνο αν

$$\mathbb{P}\left(\mathrm{X} = \mathrm{c} + \sqrt{(\mu - \mathrm{c})^2 + \sigma^2}\right) = \frac{1}{2}\left(1 + \frac{\mu - \mathrm{c}}{\sqrt{(\mu - \mathrm{c})^2 + \sigma^2}}\right),$$

$$\mathbb{P}\left(\mathrm{X} = \mathrm{c} - \sqrt{(\mu - \mathrm{c})^2 + \sigma^2}\right) = \frac{1}{2}\left(1 - \frac{\mu - \mathrm{c}}{\sqrt{(\mu - \mathrm{c})^2 + \sigma^2}}\right).$$

Απόδειξη:

Αρχικά θα αποδείξουμε την ανισότητα.

Γνωρίζουμε ότι

$$\max\{\mathrm{c}, \mathrm{X}\} = \mathrm{c} + \max\{0, \mathrm{X} - \mathrm{c}\} = \mathrm{c} + (\mathrm{X} - \mathrm{c})_+ = \mathrm{c} + \frac{(\mathrm{X} - \mathrm{c}) + |\mathrm{X} - \mathrm{c}|}{2}$$

Άρα:

$$\begin{aligned}\mathrm{E}\max\{\mathrm{c}, \mathrm{X}\} &= \mathrm{c} + \frac{1}{2}\mathrm{E}(\mathrm{X} - \mathrm{c}) + \frac{1}{2}\mathrm{E}|\mathrm{X} - \mathrm{c}| \\ &= \mathrm{c} + \frac{1}{2}(\mu - \mathrm{c}) + \frac{1}{2}\mathrm{E}|\mathrm{X} - \mathrm{c}| \\ &= \frac{\mu + \mathrm{c}}{2} + \frac{1}{2}\mathrm{E}|\mathrm{X} - \mathrm{c}| \\ &\leq \frac{\mu + \mathrm{c}}{2} + \frac{1}{2}\sqrt{\mathrm{E}(\mathrm{X} - \mathrm{c})^2} \\ &= \frac{\mu + \mathrm{c}}{2} + \frac{1}{2}\sqrt{(\mu - \mathrm{c})^2 + \sigma^2}.\end{aligned}$$

Ισότητα ισχύει αν και μόνο αν

$$\mathrm{E}|\mathrm{X} - \mathrm{c}| = \sqrt{\mathrm{E}(\mathrm{X} - \mathrm{c})^2}$$

$$\Leftrightarrow \mathrm{Var}|\mathrm{X} - \mathrm{c}| = 0$$

$$\Leftrightarrow |\mathrm{X} - \mathrm{c}| = \alpha > 0 \text{ } \mu\varepsilon \text{ } \pi\iota\theta\alpha\nu\acute{o}\tau\eta\tau\alpha \text{ } 1.$$

Δηλαδή

$$\mathrm{X} =_{\mathrm{d}} \mathrm{c} + \alpha(2\mathbb{I}\mathrm{p} - 1),$$

όπου $\mathbb{I}\mathrm{p} = \begin{cases} 1, & \text{με πιθανότητα p} \\ 0, & \text{με πιθανότητα } 1 - \mathrm{p} \end{cases}$·

Ελέγχοντας τη συνθήκη για τις ροπές, έχουμε

$$E(X) = c + \alpha(2E(\mathbb{I}p) - 1)$$

$$\Leftrightarrow \; \mu = c + \alpha(2p - 1) \qquad (*)$$

$$\mathrm{Var}(X) = 4\alpha^2 \mathrm{Var}(\mathbb{I}p)$$

$$\Leftrightarrow \; \sigma^2 = 4\alpha^2 p(1-p)$$

$$\Rightarrow \; \alpha = \frac{\sigma}{2\sqrt{p(1-p)}} \qquad (**)$$

Αντικαθιστώντας την $(**)$ στην $(*)$, προκύπτει η σχέση

$$\mu = c + \frac{\sigma}{2\sqrt{p(1-p)}}(2p-1)$$

$$\Rightarrow \; \sigma^2(2p-1)^2 = 4(\mu - c)^2 p(1-p)$$

$$\Leftrightarrow \; 4[(\mu-c)^2 + \sigma^2]p^2 - 4[(\mu-c)^2+\sigma^2]p + \sigma^2 = 0$$

$$\Leftrightarrow \; p^2 - p = \frac{-\sigma^2}{4[(\mu-c)^2+\sigma^2]}$$

$$\Leftrightarrow \; p(1-p) = \frac{\sigma^2}{4[(\mu-c)^2+\sigma^2]}$$

$$(**) \; \Rightarrow \; \alpha = \sqrt{(\mu-c)^2+\sigma^2}$$

$$(*) \; \Rightarrow \; p = \frac{1}{2}\left(1 + \frac{\mu - c}{\alpha}\right)$$

$$\overset{(**)}{\Leftrightarrow} \; p = \frac{1}{2}\left(1 + \frac{\mu-c}{\sqrt{\sigma^2+(\mu-c)^2}}\right).$$

Αντίστροφα, για τη συγκεκριμένη X έχουμε

$$E(\max\{X,c\})^{11} = c \cdot \frac{1}{2}\left(1 - \frac{\mu - c}{\sqrt{(\mu-c)^2+\sigma^2}}\right) + \left(c + \sqrt{(\mu-c)^2+\sigma^2}\right)\frac{1}{2}\left(1 + \frac{\mu - c}{\sqrt{(\mu-c)^2+\sigma^2}}\right)$$

$$= \frac{c}{2} - \frac{c}{2}\frac{\mu-c}{\sqrt{(\mu-c)^2+\sigma^2}} + \frac{c}{2} + \frac{c}{2}\frac{\mu-c}{\sqrt{(\mu-c)^2+\sigma^2}} + \frac{1}{2}\sqrt{(\mu-c)^2+\sigma^2}$$

$$+ \frac{(\mu-c)\sqrt{(\mu-c)^2+\sigma^2}}{2\sqrt{(\mu-c)^2+\sigma^2}}$$

$$= \frac{1}{2}\left((\mu+c) + \sqrt{(\mu-c)^2+\sigma^2}\right).$$

Παρατήρηση:

$$\frac{1}{2}\left((\mu+c) + \sqrt{(\mu-c)^2+\sigma^2}\right) > \frac{1}{2}\left((\mu+c) + \sqrt{(\mu-c)^2}\right)$$

$$= \frac{1}{2}\left((\mu+c) + |\mu - c|\right)$$

$$= \frac{1}{2}(\mu + 2c - c + |\mu - c|)$$

$$= c + (\mu - c)_+$$

$$= c + \max\{0, \mu - c\}$$

$$= \max\{c, \mu\}.$$

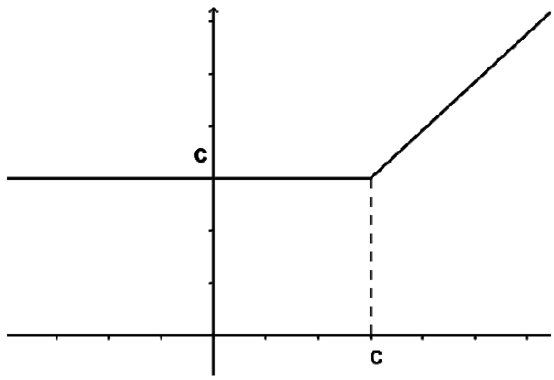
11

$$E(X) = (c+\alpha)p + (c-\alpha)(1-p) = c + \alpha(2p-1) = c + \sqrt{(\mu-c)^2+\sigma^2}\left[\frac{\mu - c}{\sqrt{(\mu-c)^2+\sigma^2}}\right] = \mu$$

και

$$E(X-c)^2 = (\mu-c)^2 + \sigma^2 \Longrightarrow Var(X) = E(X-\mu)^2 = E(X-c)^2 - (c-\mu)^2 = \sigma^2.$$

Η συνάρτηση $g(x) = \max\{x, c\}$ είναι κυρτή.



Άρα από την ανισότητα Jensen $E\big(g(x)\big) \geq g(E(x)) \Longrightarrow E(\max\{X, c\}) \geq \max\{\mu, c\}$.

Από την ανισότητα της Πρότασης 3.2.1 και από την προηγούμενη ανισότητα, προκύπτει ότι:

$$\max\{\mu, c\} \leq E(\max\{X, c\}) \leq \frac{1}{2}\Big((\mu + c) + \sqrt{(\mu - c)^2 + \sigma^2}\Big).$$

- Το γενικό αποτέλεσμα για την $E(\max\{\beta, X_1, \dots, X_n\})$ ($n \geq 2$)

Δουλεύοντας όπως πριν βρίσκουμε:

$$\begin{aligned} X_i &= t + (X_i - t) \\ &\leq t + (X_i - t)_+ \\ &\leq t + (\beta - t)_+ + \textstyle\sum_{i=1}^{n}(X_i - t)_+, \ \ i = 1, 2, 3, \dots, n. \end{aligned}$$

$$\begin{aligned} \beta &= t + (\beta - t) \\ &\leq t + (\beta - t)_+ \\ &\leq t + (\beta - t)_+ + \textstyle\sum_{i=1}^{n}(X_i - t)_+, \ \ i = 1, 2, 3, \dots, n. \end{aligned}$$

$\Rightarrow \max\{\beta,\ X_i\} \leq t + (\beta - t)_+ + \sum_{i=1}^{n}(X_i - t)_+$.

Οπότε

$$E(\max\{\beta, X_i\}) \leq (\beta - t)_+ + g(t) := h(t)$$

με

$$g(t) = t + \frac{1}{2}\sum_{i=1}^{n}\left[(\mu_i - t) + \sqrt{(\mu_i - t)^2 + \sigma_i{}^2}\right]$$

όπως πριν.

Η $h(t) = g(t) + (\beta - t)_+$ είναι άθροισμα μιας γνήσια κυρτής (της $g(t)$) και μιας κυρτής (της $(\beta - t)_+$) άρα είναι γνήσια κυρτή. Αφού $h(t) \geq g(t)$ προκύπτει ότι $\min_t h(t) = g(t_o)$ όταν $\beta \leq t_o$.

Όταν $\beta > t_o$ τα πράγματα αλλάζουν.

Έστω $\beta > t_o$.

Έχουμε

$$h(t) = \begin{cases} \beta - t + g(t), & t \leq \beta \\ \quad g(t), & t \geq \beta. \end{cases}$$

Επίσης, $h(\beta) = g(\beta)$ και $g'(x) > 0$ σε μια περιοχή του β. Θα δείξουμε ότι το ελάχιστο της h επιτυγχάνεται στο $t = \beta$. Αρκεί να δείξουμε ότι $h'(\beta -) < 0$, (αφού $h'(\beta +) > 0$). Δηλαδή $g'(\beta) < 1$.

Γενικά, $g'(t) = 1 + \frac{1}{2}\sum_{i=1}^{n}\left[(-1) + \frac{t-\mu_i}{\sqrt{(\mu_i-t)^2+\sigma_i{}^2}}\right] < 1 \Leftrightarrow \sum_{i=1}^{n}\left[\frac{t-\mu_i}{\sqrt{(\mu_i-t)^2+\sigma_i{}^2}}\right] < n.$

Αυτό ισχύει για κάθε $t \in \mathbb{R}$ επειδή

$$t - \mu_i \leq |t - \mu_i| < \sqrt{(\mu_i - t)^2 + \sigma_i{}^2}$$

$$\begin{aligned} \Rightarrow \sum_{i=1}^{n}\left[\frac{t-\mu_i}{\sqrt{(\mu_i-t)^2+\sigma_i{}^2}}\right] &\leq \left|\sum_{i=1}^{n}\left[\frac{t-\mu_i}{\sqrt{(\mu_i-t)^2+\sigma_i{}^2}}\right]\right| \\ &\leq \sum_{i=1}^{n}\left[\frac{|t-\mu_i|}{\sqrt{(\mu_i-t)^2+\sigma_i{}^2}}\right] \\ &< \sum_{i=1}^{n} 1 \\ &= n. \end{aligned}$$

Δείξαμε λοιπόν ότι $g'(\beta) < 1$. Συνεπώς $h'(\beta -) < 0$, και το ελάχιστο επιτυγχάνεται για $\widetilde{t_o} = \beta$.

Γενικά,

$$\widetilde{t_o} = \max\{t_o, \beta\} \text{ (πάντα).}$$

Για ισότητα (όταν $\beta > t_o$):

Η από κοινού κατανομή των $(X_1, X_2, \dots, X_n)$, για να έχουμε $\widetilde{t_o} = \beta$, πρέπει να είναι τέτοια ώστε:

- είτε όλα τα X_i να είναι μικρότερα του β,
- είτε ακριβώς ένα να είναι μεγαλύτερο του β, και όλα τα άλλα μικρότερα του β.

Πρόταση 3.2.2 Έστω $g(t) = t + \frac{1}{2}\sum_{i=1}^{n}\left\{(\mu_i - t) + \sqrt{(\mu_i - t)^2 + \sigma_i^2}\right\}$. Αν $\beta > t_o$ τότε $E\max\{\beta, X_1, X_2, \dots, X_n\} \le g(\beta)$. Η ισότητα επιτυγχάνεται αν και μόνο αν:

$$\max\{\beta, X_1, X_2, \dots, X_n\} = \begin{cases} \beta + \sqrt{(\mu_i - \beta)^2 + \sigma_i^2}, & \text{με πιθανότητα } p_i \\ \beta, & \text{με πιθανότητα } 1 - \sum_{i=1}^{n} p_i \end{cases}$$

όπου $p_i = \frac{1}{2}\left(1 + \frac{\mu_i - \beta}{\sqrt{\sigma_i^2 + (\mu_i - \beta)^2}}\right)$.

Απόδειξη

Ξέρουμε ότι $g'(\beta) > 0$ ($\textit{επειδή } \beta > t_o$).

Όμως

$$g'(\beta) > 0$$

$$\Leftrightarrow \sum_{i=1}^{n}\left[\frac{\beta - \mu_i}{\sqrt{(\mu_i - \beta)^2 + \sigma_i^2}}\right] > \sum_{i=1}^{n}\left[\frac{t_o - \mu_i}{\sqrt{(\mu_i - t_o)^2 + \sigma_i^2}}\right]$$

$$= n - 2. \tag{3.10}$$

Δηλαδή:

$$n - 2 < \sum_{i=1}^{n}\left[\frac{\beta - \mu_i}{\sqrt{(\mu_i - \beta)^2 + \sigma_i^2}}\right] < n. \tag{3.11}$$

Θέλουμε

$$|X_i - \beta| = c_i > 0 \; \textit{με πιθανότητα } 1.$$

$$\mathbb{P}(X_i = \beta + c_i) = p_i,$$

$$\mathbb{P}(X_i = \beta - c_i) = 1 - p_i.$$

Τότε

$$\begin{aligned}\mu_i &= E(X_i)\\ &= p_i(\beta + c_i) + (1 - p_i)(\beta - c_i)\\ &= \beta + c_i(2p_i - 1)\end{aligned}$$

$$\Leftrightarrow \quad 2p_i - 1 = \frac{\mu_i - \beta}{c_i}$$

$$\Leftrightarrow \quad p_i = \frac{1}{2}\left(1 + \frac{\mu_i - \beta}{c_i}\right)$$

$$\begin{aligned}{c_i}^2 &= E(X_i - \beta)^2\\ &= E(X_i - \mu_i + \mu_i + \beta)^2\\ &= {\sigma_i}^2 + (\mu_i - \beta)^2\end{aligned}$$

$$\Longrightarrow \quad c_i = \sqrt{{\sigma_i}^2 + (\mu_i - \beta)^2}$$

Οπότε

$$p_i = \frac{1}{2}\left(1 + \frac{\mu_i - \beta}{\sqrt{{\sigma_i}^2 + (\mu_i - \beta)^2}}\right).$$

Βρήκαμε

$$\mathbb{P}\left(X_i = \beta + \sqrt{(\mu_i - \beta)^2 + {\sigma_i}^2}\right) = \frac{1}{2}\left(1 + \frac{\mu_i - \beta}{\sqrt{(\mu_i - \beta)^2 + {\sigma_i}^2}}\right) \in (0,1)$$

$$\mathbb{P}\left(X_i = \beta - \sqrt{(\mu_i - \beta)^2 + {\sigma_i}^2}\right) = \frac{1}{2}\left(1 - \frac{\mu_i - \beta}{\sqrt{(\mu_i - \beta)^2 + {\sigma_i}^2}}\right) \in (0,1)$$

$i = 1, 2, \ldots, n.$

Τώρα

$$\begin{aligned}\sum_{i=1}^n p_i &= \frac{n}{2} + \frac{1}{2}\sum_{i=1}^n \frac{\mu_i - \beta}{\sqrt{{\sigma_i}^2 + (\mu_i - \beta)^2}}\\ &= \frac{n}{2} - \frac{1}{2}\sum_{i=1}^n \frac{\beta - \mu_i}{\sqrt{{\sigma_i}^2 + (\mu_i - \beta)^2}}\end{aligned}$$

$$\overset{(3.10)}{\Longleftrightarrow} \quad \sum_{i=1}^n p_i < \frac{n}{2} - \frac{n-2}{2}$$

$$\Leftrightarrow \qquad \sum_{i=1}^{n} p_i < 1.$$

Τότε

$$1 - \sum_{i=1}^{n} p_i = 1 - \frac{n}{2} + \frac{1}{2}\sum_{i=1}^{n} \frac{\beta - \mu_i}{\sqrt{{\sigma_i}^2 + (\mu_i - \beta)^2}}$$

$$= \frac{1}{2}\left\{\sum_{i=1}^{n} \frac{\beta - \mu_i}{\sqrt{{\sigma_i}^2 + (\mu_i - \beta)^2}} - (n-2)\right\}$$

$$\overset{(3.10)}{\Longleftrightarrow} \qquad 1 - \sum_{i=1}^{n} p_i > 0$$

$$\Leftrightarrow \qquad \sum_{i=1}^{n} p_i > 1$$

Καταλήγουμε στο εξής:

$$\mathbb{P}\begin{pmatrix} X_1 = \beta - \sqrt{(\mu_1 - \beta)^2 + {\sigma_1}^2} \\ \vdots \\ X_i = \beta + \sqrt{(\mu_i - \beta)^2 + {\sigma_i}^2} \\ \vdots \\ X_n = \beta - \sqrt{(\mu_n - \beta)^2 + {\sigma_n}^2} \end{pmatrix} = \frac{1}{2}\left(1 + \frac{\mu_i - \beta}{\sqrt{(\mu_i - \beta)^2 + {\sigma_i}^2}}\right), i = 1,2,\dots,n.$$

$$\mathbb{P}\left(X_i = \beta - \sqrt{(\mu_i - \beta)^2 + {\sigma_i}^2},\ i = 1,2,\dots,n\right) = \frac{1}{2}\left\{\sum_{i=1}^{n} \frac{\beta - \mu_i}{\sqrt{{\sigma_i}^2 + (\mu_i - \beta)^2}} - (n-2)\right\}$$

$$= 1 - \sum_{i=1}^{n} \frac{1}{2}\left(1 + \frac{\mu_i - \beta}{\sqrt{{\sigma_i}^2 + (\mu_i - \beta)^2}}\right)$$

$$> 0.$$

Τότε,

$$\max\{\beta, X_1, X_2, \dots, X_n\} = \begin{cases} \beta + \sqrt{(\mu_i - \beta)^2 + {\sigma_i}^2}, & \text{με πιθανότητα } p_i \\ \beta, & \text{με πιθανότητα } 1 - \sum\limits_{i=1}^{n} p_i \end{cases}$$

και

$$E\max\{\beta, X_1, X_2, \dots, X_n\} = \sum_{i=1}^{n} p_i\left(\beta + \sqrt{(\mu_i - \beta)^2 + {\sigma_i}^2}\right) + \beta\left(1 - \sum_{i=1}^{n} p_i\right)$$

$$= \beta + \sum_{i=1}^{n} \frac{1}{2}\left(1 + \frac{\mu_i - \beta}{\sqrt{{\sigma_i}^2 + (\mu_i - \beta)^2}}\right)\sqrt{{\sigma_i}^2 + (\mu_i - \beta)^2}$$

$$= \beta + \frac{1}{2}\sum_{i=1}^{n}\left(\sqrt{{\sigma_i}^2 + (\mu_i - \beta)^2} + (\mu_i - \beta)\right) = g(\beta).$$

■

Το τελικό συμπέρασμα είναι το εξής:

Έστω

$$g(t) = t + \frac{1}{2}\sum_{i=1}^{n}\left\{(\mu_i - t) + \sqrt{(\mu_i - t)^2 + {\sigma_i}^2}\right\}.$$

Η g(t) είναι γνήσια κυρτή και ελαχιστοποιείται στο $t = t_o$.

1. Αν $\beta \le t_o$ τότε $\mathrm{E}\max\{\beta, X_1, X_2, \dots, X_n\} \le g(t_o)$ με ισότητα όπως πριν (χωρίς περιορισμό).
2. Αν $\beta > t_o$ τότε $\mathrm{E}\max\{\beta, X_1, X_2, \dots, X_n\} \le g(\beta)$ με ισότητα ανν

$$\max\{\beta, X_1, X_2, \dots, X_n\} = \begin{cases} \beta + \sqrt{(\mu_i - \beta)^2 + {\sigma_i}^2}, & \text{με πιθανότητα } p_i \\ \beta, & \text{με πιθανότητα } 1 - \sum_{i=1}^{n} p_i \end{cases}$$

όπου $p_i = \frac{1}{2}\left(1 + \frac{\mu_i - \beta}{\sqrt{{\sigma_i}^2 + (\mu_i - \beta)^2}}\right)$.

Πιο γενικά ισχύει το παρακάτω Θεώρημα:

Θεώρημα 3.2.3

Αν $\gamma = \max\{t_o, \beta\}$ τότε

$$\sup_{\mathcal{F}} \mathrm{E}\max\{\beta, X_1, X_2, \dots, X_n\} = g(\max\{t_o, \beta\}) = g(\gamma).$$

Σημειώνεται ότι το γ είναι συνάρτηση των **μ**, $\boldsymbol{\sigma^2}$, και β.

Η ισότητα ισχύει αν και μόνο αν:

$$\mathbb{P}\begin{pmatrix} X_1 = \gamma - \sqrt{(\mu_1 - \gamma)^2 + {\sigma_1}^2} \\ \vdots \\ X_i = \gamma + \sqrt{(\mu_i - \gamma)^2 + {\sigma_i}^2} \\ \vdots \\ X_n = \gamma - \sqrt{(\mu_n - \gamma)^2 + {\sigma_n}^2} \end{pmatrix} = \frac{1}{2}\left(1 + \frac{\mu_i - \gamma}{\sqrt{(\mu_i - \gamma)^2 + {\sigma_i}^2}}\right), i = 1,2, \dots, n.$$

και

$$\mathbb{P}\left(X_i = \gamma - \sqrt{(\mu_i - \gamma)^2 + {\sigma_i}^2},\ i = 1,2, \dots, n\right) = 1 - \sum_{i=1}^{n} \frac{1}{2}\left(1 + \frac{\mu_i - \gamma}{\sqrt{{\sigma_i}^2 + (\mu_i - \gamma)^2}}\right) \ge 0.$$

■

Ας πάρουμε σαν παράδειγμα την ειδική περίπτωση n = 2, $\mu_1 = \mu_2 = 0$, ${\sigma_1}^2 = {\sigma_2}^2 = 1$.

Τώρα $g(t) = \sqrt{1 + t^2}$, οπότε $t_o = 0$.

(A) $\mathrm{E}\max\{X_1, X_2\} \le 1$ με ισότητα αν και μόνο αν

$$\mathbb{P}(X_1 = -1, X_2 = 1) = \mathbb{P}(X_1 = 1, X_2 = -1) = \frac{1}{2} \qquad (*)$$

(Β) Αν $\beta \in \mathbb{R}$ τότε:

$$\mathrm{E}\max\{\beta, X_1, X_2\} \leq g(\gamma) = \sqrt{1+\gamma^2} \text{ όπου } \gamma = \max\{0, \beta\}$$

Αν $\beta \leq 0$ η ισότητα επιτυγχάνεται μόνο στην $(*)$

Αν $\beta > 0$, τότε $\gamma = \beta$ και $\mathrm{E}\max\{\beta, X_1, X_2\} \leq \sqrt{1+\beta^2}$ (> 1)

με ισότητα αν και μόνο αν

$$\mathbb{P}\left(X_1 = \beta + \sqrt{1+\beta^2},\ X_2 = \beta - \sqrt{1+\beta^2}\right) = \frac{1}{2}\left(1 - \frac{\beta}{\sqrt{1+\beta^2}}\right)$$

$$\mathbb{P}\left(X_1 = \beta - \sqrt{1+\beta^2},\ X_2 = \beta + \sqrt{1+\beta^2}\right) = \frac{1}{2}\left(1 - \frac{\beta}{\sqrt{1+\beta^2}}\right) \qquad (**)$$

$$\mathbb{P}\left(X_1 = \beta - \sqrt{1+\beta^2},\ X_2 = \beta - \sqrt{1+\beta^2}\right) = \frac{\beta}{\sqrt{1+\beta^2}}\ (> 0)$$

Η κατανομή $(**)$ διαφέρει από την $(*)$.

Τα φράγματα που μελετήσαμε στα κεφάλαια 2 και 3, συνοψίζονται στον παρακάτω πίνακα

ΦΡΑΓΜΑΤΑ
1. Arnold – Groeveld (1979 – 80)
$\sum_{i=1}^{n} \lambda_i E(X_{i:n}) \leq n\bar{\lambda}\bar{\mu} + \left(\sum_{i=1}^{n}(\lambda_i - \bar{\lambda})^2\right)^{1/2} \left(\sum_{i=1}^{n}\{(\mu_i - \bar{\mu})^2 + \sigma_i{}^2\}\right)^{1/2} = AG_n$
2. **Bertsimas – Natarajan – Teo (2006)**
$E(X_{n:n}) \leq t_o + \frac{1}{2}\sum_{i=1}^{n}\left\{\mu_i - t_o + \sqrt{(\mu_i - t_o)^2 + \sigma_i{}^2}\right\} = BNT_n$ όπου $t = t_o \in \mathbb{R}$ είναι η μοναδική λύση της εξίσωσης $\sum_{i=1}^{n} \frac{t - \mu_i}{\sqrt{(t - \mu_i)^2 + \sigma_i{}^2}} = n - 2, \qquad -\infty < t < \infty$
3. Papadatos (2001)
$\sum_{i=1}^{n} \lambda_i E(X_{i:n}) \leq n\bar{\lambda}\bar{\mu} + \left(\sum_{i=1}^{n}(\lambda_i - \bar{\lambda})^2\right)^{1/2} \left(\sum_{i=1}^{n}\left\{(\mu_i - \bar{\mu})^2 + \sigma_i{}^2\right\} - n\mathrm{Var}\bar{X}\right)^{1/2} = P_n$

Στο επόμενο κεφάλαιο, και ειδικότερα στην ενότητα 4.1, θα μελετήσουμε τη σχέση που έχουν τα παραπάνω φράγματα για n = 2. Όπως θα διαπιστώσουμε, το φράγμα P_n για n = 2, είναι καλύτερο σε σχέση με τα φράγματα BNT_n και AG_n για n = 2, παρόλο που το φράγμα BNT είναι σχετικά πιο πρόσφατο. Επίσης στην ενότητα 4.2, θα δείξουμε ότι το εύρος δύο διατεταγμένων παρατηρήσεων, έχει γραμμική σχέση με τη μέγιστη παρατήρηση. Οπότε βρίσκοντας ένα φράγμα για τη μέση τιμή της μέγιστης παρατήρησης, βρίσκουμε φράγμα για τη μέση τιμή του εύρους δύο διατεταγμένων παρατηρήσεων.

ΚΕΦΑΛΑΙΟ 4

ΕΙΔΙΚΕΣ ΠΕΡΙΠΤΩΣΕΙΣ ΦΡΑΓΜΑΤΩΝ ΓΙΑ 2 ΔΙΑΤΕΤΑΓΜΕΝΕΣ ΠΑΡΑΤΗΡΗΣΕΙΣ

4.1. Σύγκριση βέλτιστων φραγμάτων στις ανισότητες των Arnold – Groevelt (AG_2), Bertsimas – Natarajan – Teo (BNT_2), Papadatos (P_2)

Ας θεωρήσουμε την περίπτωση n = 2. Χωρίς βλάβη της γενικότητας υποθέτουμε ότι $\mu_1 \le \mu_2$. Αν $\mu_1 = \mu_2$ τότε $t_o = \mu_1 = \mu_2$. Αν $\mu_1 < \mu_2$, τότε η σχέση (3.7) γίνεται:

$$\frac{t_o-\mu_1}{\sqrt{(\mu_1-t_o)^2+{\sigma_1}^2}}+\frac{t_o-\mu_2}{\sqrt{(\mu_2-t_o)^2+{\sigma_2}^2}}=0$$

$$\Leftrightarrow \quad \frac{t_o-\mu_1}{\sqrt{(\mu_1-t_o)^2+{\sigma_1}^2}}=\frac{\mu_2-t_o}{\sqrt{(\mu_2-t_o)^2+{\sigma_2}^2}}$$

$$\Leftrightarrow \quad \frac{1}{\sqrt{1+\frac{{\sigma_1}^2}{(\mu_1-t_o)^2}}}=\frac{1}{\sqrt{1+\frac{{\sigma_2}^2}{(\mu_2-t_o)^2}}}$$

$$\Longrightarrow \quad \frac{{\sigma_1}^2}{(\mu_1-t_o)^2}=\frac{{\sigma_2}^2}{(\mu_2-t_o)^2}$$

$$\Leftrightarrow \quad \left|\frac{\sigma_1}{\mu_1-t_o}\right| = \left|\frac{\sigma_2}{\mu_2-t_o}\right|$$

$$\xRightarrow{\mu_1<t_o<\mu_2} \quad \frac{t_o-\mu_1}{\sigma_1}=\frac{\mu_2-t_o}{\sigma_2}$$

$$\Leftrightarrow \quad t_o=\frac{\sigma_2}{\sigma_1+\sigma_2}\mu_1+\frac{\sigma_1}{\sigma_1+\sigma_2}\mu_2 \qquad (\#)$$

Βλέπουμε λοιπόν ότι $t_o \in [\mu_1, \mu_2]$, όπως αναμενόταν (για n = 2).

Το φράγμα AG για n = 2 γίνεται

$$AG_2=\frac{\mu_1+\mu_2}{2}+\left(\left(\frac{1}{2}\right)^2+\left(1-\frac{1}{2}\right)^2\right)^{1/2}\left(\left(\mu_1-\frac{\mu_1+\mu_2}{2}\right)^2+{\sigma_1}^2+\left(\mu_2-\frac{\mu_1+\mu_2}{2}\right)^2+{\sigma_2}^2\right)^{1/2}$$

$$=\frac{\mu_1+\mu_2}{2}+\frac{1}{2}\sqrt{2{\sigma_1}^2+2{\sigma_2}^2+(\mu_2-\mu_1)^2}. \qquad (4.1)$$

Η BNT για n = 2 γίνεται

$$\mathrm{BNT}_2 = g(t_o)$$

$$= t_o + \frac{1}{2}\left[(\mu_1 - t_o) + \sqrt{(\mu_1 - t_o)^2 + {\sigma_1}^2} + (\mu_2 - t_o) + \sqrt{(\mu_2 - t_o)^2 + {\sigma_2}^2}\right]$$

$$= \frac{\mu_1+\mu_2}{2} + \frac{1}{2}\left\{\sqrt{(\mu_1 - t_o)^2 + {\sigma_1}^2} + \sqrt{(\mu_2 - t_o)^2 + {\sigma_2}^2}\right\}. \tag{4.2}$$

Από τη σχέση (#) προκύπτει:

$$t_o - \mu_1 = \frac{\sigma_2}{\sigma_1 + \sigma_2}\mu_1 + \frac{\sigma_1}{\sigma_1 + \sigma_2}\mu_2 - \frac{\sigma_1 + \sigma_2}{\sigma_1 + \sigma_2}\mu_1 = \frac{\sigma_1}{\sigma_1 + \sigma_2}(\mu_2 - \mu_1).$$

Δηλαδή,

$$(t_o - \mu_1)^2 + {\sigma_1}^2 = \frac{{\sigma_1}^2}{(\sigma_1+\sigma_2)^2}(\mu_2 - \mu_1)^2 + {\sigma_1}^2.$$

Όμοια:

$$(t_o - \mu_2)^2 + {\sigma_2}^2 = \frac{{\sigma_2}^2}{(\sigma_1+\sigma_2)^2}(\mu_2 - \mu_1)^2 + {\sigma_2}^2.$$

Οπότε η BNT για n = 2 γίνεται:

$$\mathrm{BNT}_2 = \frac{\mu_1+\mu_2}{2} + \frac{1}{2}\left\{\sqrt{\frac{{\sigma_1}^2}{(\sigma_1+\sigma_2)^2}\left(\mu_2 - \mu_1\right)^2 + {\sigma_1}^2} + \sqrt{\frac{{\sigma_2}^2}{(\sigma_1+\sigma_2)^2}\left(\mu_2 - \mu_1\right)^2 + {\sigma_2}^2}\right\}$$

$$= \frac{\mu_1+\mu_2}{2} + \frac{1}{2}\left\{\frac{1}{\sigma_1+\sigma_2}\left[\sigma_1\sqrt{(\mu_2 - \mu_1)^2 + (\sigma_1 + \sigma_2)^2} + \sigma_2\sqrt{(\mu_2 - \mu_1)^2 + (\sigma_1 + \sigma_2)^2}\right]\right\}$$

$$= \frac{\mu_1+\mu_2}{2} + \frac{1}{2}\left[\sqrt{\left(\mu_2 - \mu_1\right)^2 + (\sigma_1 + \sigma_2)^2}\right]. \tag{4.3}$$

Τέλος η ανισότητα (2.15) για n = 2 και για $\lambda_1 = 0,\ \lambda_2 = 1$, γίνεται

$$E(X_{2:2}) \le \bar{\mu} + \frac{\sqrt{2}}{2}\left\{{\sigma_1}^2 + {\sigma_2}^2 + \frac{(\mu_1 - \mu_2)^2}{2} - \frac{1}{2}({\sigma_1}^2 + {\sigma_2}^2 + 2c)\right\}^{1/2}$$

$$= \bar{\mu} + \frac{1}{2}\sqrt{(\mu_1 - \mu_2)^2 + {\sigma_1}^2 + {\sigma_2}^2 - 2c},$$

όπου $c = \mathrm{Cov}(X_1, X_2),\ \bar{\mu} = \frac{\mu_1+\mu_2}{2}$.

Επομένως το φράγμα της (2.15) για n = 2, είναι το

$$P_2 = \bar{\mu} + \frac{1}{2}\sqrt{(\mu_1 - \mu_2)^2 + {\sigma_1}^2 + {\sigma_2}^2 - 2c}, \tag{4.4}$$

όπου $c = \mathrm{Cov}(X_1, X_2),\ \bar{\mu} = \frac{\mu_1+\mu_2}{2}$.

Γνωρίζοντας πλέον τα φράγματα για δύο διατεταγμένες παρατηρήσεις, είμαστε σε θέση πλέον να τα συγκρίνουμε μεταξύ τους.

4.1.1. ΣΥΓΚΡΙΣΗ BNT_2 ΚΑΙ P_2

Από τη σχέση (4.4), είναι προφανές ότι το φράγμα σχετίζεται με την ποσότητα

$$2(P_2 - \bar{\mu}) = \sqrt{(\mu_1 - \mu_2)^2 + {\sigma_1}^2 + {\sigma_2}^2 - 2c}. \qquad (4.5)$$

Από την ανισότητα Cauchy – Schwarz ισχύει

$$c^2 \leq {\sigma_1}^2 {\sigma_2}^2,$$

με ισότητα αν και μόνο αν $X_1 = \alpha X_2 + \beta$ με πιθανότητα 1 (και $\alpha \neq 0$).

Επομένως,

$$c \geq -\sigma_1 \sigma_2$$

από την οποία προκύπτει η $-2c \leq 2\sigma_1\sigma_2$. Άρα η (4.5) γίνεται

$$\begin{aligned} 2(P_2 - \bar{\mu}) &\leq \sqrt{(\mu_1 - \mu_2)^2 + {\sigma_1}^2 + {\sigma_2}^2 + 2\sigma_1\sigma_2} \\ &= \sqrt{(\mu_1 - \mu_2)^2 + (\sigma_1 + \sigma_2)^2} \\ &= 2(BNT_2 - \bar{\mu}). \end{aligned}$$

Η ισότητα στα φράγματα P_2 και BNT_2 επιτυγχάνεται μόνο στην τετριμμένη περίπτωση που $\rho = -1$.(Όπου $\rho = \frac{c}{\sigma_1\sigma_2}$ είναι ο συντελεστής συσχέτισης των X_1 και X_2). Αλλιώς είναι καλύτερο το P_2.

Πρόταση 4.1.1

Για n = 2 και οποιεσδήποτε τιμές των $\mu_1, \mu_2 \in \mathbb{R},\ \sigma_1, \sigma_2 > 0\ \textit{και}\ c \in [-\sigma_1\sigma_2,\ \sigma_1\sigma_2]$,

$$BNT_2 - P_2 = \frac{\sigma_1\sigma_2(1+\rho)}{\sqrt{(\mu_1 - \mu_2)^2 + (\sigma_1 + \sigma_2)^2} + \sqrt{(\mu_1 - \mu_2)^2 + {\sigma_1}^2 + {\sigma_2}^2 - 2c}}, \qquad (4.6)$$

όπου $\rho = \frac{c}{\sigma_1\sigma_2} \in [-1,\ 1]$.

Απόδειξη

$$\begin{aligned} &[2(BNT_2 - \bar{\mu})]^2 - [2(P_2 - \bar{\mu})]^2 \\ &\qquad = (\mu_1 - \mu_2)^2 + (\sigma_1 + \sigma_2)^2 - (\mu_1 - \mu_2)^2 - ({\sigma_1}^2 + {\sigma_2}^2 - 2c) \end{aligned}$$

$$= 2\sigma_1\sigma_2 + 2c$$

$$= 2\sigma_1\sigma_2(1+\rho),$$

όπου $\rho = \frac{c}{\sigma_1\sigma_2}$ είναι ο συντελεστής συσχέτισης των X_1 και X_2.

Επομένως,

$$(BNT_2 - \bar{\mu})^2 - (P_2 - \bar{\mu})^2 = \tfrac{1}{2}\sigma_1\sigma_2(1+\rho).$$

Δηλαδή

$$[(BNT_2 - \bar{\mu}) - (P_2 - \bar{\mu})][(BNT_2 - \bar{\mu}) + (P_2 - \bar{\mu})]$$

$$= (BNT_2 - P_2)(BNT_2 + P_2 - \mu_1 - \mu_2)$$

$$= \frac{1}{2}\sigma_1\sigma_2(1+\rho).$$

Όμως,

$$BNT_2 + P_2 - \mu_1 - \mu_2$$

$$= \tfrac{1}{2}\left[\sqrt{\left(\mu_2 - \mu_1\right)^2 + (\sigma_1 + \sigma_2)^2}\right] + \tfrac{1}{2}\sqrt{(\mu_1 - \mu_2)^2 + {\sigma_1}^2 + {\sigma_2}^2 - 2c}$$

$$= \frac{1}{2}\left[\sqrt{\left(\mu_2 - \mu_1\right)^2 + (\sigma_1 + \sigma_2)^2}\right] + \frac{1}{2}\sqrt{(\mu_1 - \mu_2)^2 + {\sigma_1}^2 + {\sigma_2}^2 - 2\sigma_1\sigma_2(1+\rho)}$$

Οπότε

$$(BNT_2 - P_2)\left[\tfrac{1}{2}\left[\sqrt{(\mu_2 - \mu_1)^2 + (\sigma_1 + \sigma_2)^2}\right] + \tfrac{1}{2}\sqrt{(\mu_1 - \mu_2)^2 + {\sigma_1}^2 + {\sigma_2}^2 - 2\sigma_1\sigma_2(1+\rho)}\right]$$

$$= \tfrac{1}{2}\sigma_1\sigma_2(1+\rho).$$

Από την παραπάνω ισότητα, προκύπτει η ισότητα που θέλουμε να αποδείξουμε.

■

Τώρα είμαστε σε θέση να ελέγξουμε ποιο από τα παραπάνω φράγματα τελικά είναι καλύτερο όταν έχουμε δύο διατεταγμένες παρατηρήσεις. Την απάντηση τη δίνει το παρακάτω πόρισμα

Πόρισμα 4.1.2. Για τα φράγματα BNT_2 και P_2 ισχύει η ανισότητα

$$BNT_2 \geq P_2.$$

Η ισότητα επιτυγχάνεται αν και μόνο αν $\rho = -1$.

Απόδειξη

Η απόδειξη είναι προφανής από τη σχέση (4.6), αφού $\rho \geq -1$.

■

Ας δούμε πότε γίνεται να έχουμε $\mathrm{E}(\mathrm{X}_1) = \mu_1$, $\mathrm{Var}(\mathrm{X}_1) = {\sigma_1}^2 > 0$, $\mathrm{E}(\mathrm{X}_2) = \mu_2$, $\mathrm{Var}(\mathrm{X}_2) = {\sigma_2}^2 > 0$ και $\rho = -1$.

Αφού $\rho = -1$ πρέπει

$$\mathrm{X}_2 = -\theta \mathrm{X}_1 + \beta \text{ με πιθανότητα 1 για κάποιο } \theta > 0.$$

Άρα

$$\mu_2 = -\theta\mu_1 + \beta.$$

Επίσης

$$\begin{aligned} {\sigma_2}^2 &= \mathrm{Var}(\mathrm{X}_2) \\ &= \mathrm{Var}(-\theta \mathrm{X}_1 + \beta\ ) \\ &= \theta^2 \mathrm{Var}(\mathrm{X}_1) \\ &= \theta^2 {\sigma_1}^2 \end{aligned}$$

οπότε $\theta = \frac{\sigma_2}{\sigma_1}$.

Επομένως,

$$\mu_2 = -\frac{\sigma_2}{\sigma_1}\mu_1 + \beta$$

δηλαδή

$$\beta = \mu_2 + \frac{\sigma_2}{\sigma_1}\mu_1.$$

Οπότε

$$\mathrm{X}_2 = -\frac{\sigma_2}{\sigma_1}\mathrm{X}_1 + \mu_2 + \frac{\sigma_2}{\sigma_1}\mu_1$$

$$\Leftrightarrow \sigma_1(\mathrm{X}_2 - \mu_2) = -\sigma_2(\mathrm{X}_1 - \mu_1).$$

Άρα

$$\frac{\mathrm{X}_2 - \mu_2}{\sigma_2} = -\frac{\mathrm{X}_1 - \mu_1}{\sigma_1} \text{ με πιθανότητα 1.} \qquad (4.7)$$

Παρατηρούμε τώρα ότι η βέλτιστη κατανομή που πετυχαίνει την ισότητα στην BNT_2 ικανοποιεί αναγκαστικά και την (4.7).

Πράγματι, ας θεωρήσουμε την κατανομή X_i, i= 1, 2 που πετυχαίνει την ισότητα στην BNT_2, η οποία δίνεται στο Θεώρημα 3.1.2. Αντικαθιστώντας την τιμή του t_o που δίνει η σχέση (#) της παραγράφου 4.1. προκύπτει:

$$\mathbb{P}\begin{bmatrix} X_1-\mu_1=\frac{\sigma_1}{\sigma_1+\sigma_2}(\mu_2-\mu_1)+\sqrt{\frac{\sigma_1{}^2}{(\sigma_1+\sigma_2)^2}(\mu_2-\mu_1)^2+\sigma_1{}^2} \\ X_2-\mu_2=\frac{\sigma_2}{\sigma_1+\sigma_2}(\mu_1-\mu_2)-\sqrt{\frac{\sigma_2{}^2}{(\sigma_1+\sigma_2)^2}(\mu_2-\mu_1)^2+\sigma_2{}^2} \end{bmatrix}$$

$$=\frac{1}{2}\left(1+\frac{\mu_1-\mu_2}{\sqrt{(\mu_1-\mu_2)^2+(\sigma_1+\sigma_2)^2}}\right)$$

και

$$\mathbb{P}\begin{bmatrix} X_1-\mu_1=\frac{\sigma_1}{\sigma_1+\sigma_2}(\mu_2-\mu_1)-\sqrt{\frac{\sigma_1{}^2}{(\sigma_1+\sigma_2)^2}(\mu_2-\mu_1)^2+\sigma_1{}^2} \\ X_2-\mu_2=\frac{\sigma_2}{\sigma_1+\sigma_2}(\mu_1-\mu_2)+\sqrt{\frac{\sigma_2{}^2}{(\sigma_1+\sigma_2)^2}(\mu_2-\mu_1)^2+\sigma_2{}^2} \end{bmatrix}$$

$$=\frac{1}{2}\left(1+\frac{\mu_2-\mu_1}{\sqrt{(\mu_1-\mu_2)^2+(\sigma_1+\sigma_2)^2}}\right).$$

Δηλαδή

$$\mathbb{P}\begin{bmatrix} \frac{X_1-\mu_1}{\sigma_1}=\frac{\mu_2-\mu_1}{\sigma_1+\sigma_2}-\sqrt{(\mu_2-\mu_1)^2+(\sigma_1+\sigma_2)^2} \\ \frac{X_2-\mu_2}{\sigma_2}=\frac{\mu_1-\mu_2}{\sigma_1+\sigma_2}+\sqrt{(\mu_2-\mu_1)^2+(\sigma_1+\sigma_2)^2} \end{bmatrix}=\frac{1}{2}\left(1+\frac{\mu_1-\mu_2}{\sqrt{(\mu_1-\mu_2)^2+(\sigma_1+\sigma_2)^2}}\right)$$

και

$$\mathbb{P}\begin{bmatrix} \frac{X_1-\mu_1}{\sigma_1}=\frac{\mu_2-\mu_1}{\sigma_1+\sigma_2}+\sqrt{(\mu_2-\mu_1)^2+(\sigma_1+\sigma_2)^2} \\ \frac{X_2-\mu_2}{\sigma_2}=\frac{\mu_1-\mu_2}{\sigma_1+\sigma_2}-\sqrt{(\mu_2-\mu_1)^2+(\sigma_1+\sigma_2)^2} \end{bmatrix}=\frac{1}{2}\left(1+\frac{\mu_2-\mu_1}{\sqrt{(\mu_1-\mu_2)^2+(\sigma_1+\sigma_2)^2}}\right).$$

Οπότε καταλήγουμε στο συμπέρασμα ότι:

$$\mathbb{P}\left(\frac{X_1-\mu_1}{\sigma_1}=-\frac{X_2-\mu_2}{\sigma_2}\right)=\frac{1}{2}\left(1+\frac{\mu_1-\mu_2}{\sqrt{(\mu_1-\mu_2)^2+(\sigma_1+\sigma_2)^2}}\right)$$

και

$$\mathbb{P}\left(\frac{X_1-\mu_1}{\sigma_1}=-\frac{X_2-\mu_2}{\sigma_2}\right)=\frac{1}{2}\left(1+\frac{\mu_2-\mu_1}{\sqrt{(\mu_1-\mu_2)^2+(\sigma_1+\sigma_2)^2}}\right).$$

Αν συνοψίσουμε τις δύο παραπάνω ισότητες σε μία, προκύπτει ότι

$$\mathbb{P}\left(\frac{X_1-\mu_1}{\sigma_1} = -\frac{X_2-\mu_2}{\sigma_2}\right) = 1\,,$$

δηλαδή η σχέση (4.7).

Επομένως, το BNT_2 ΔΕΝ ΕΙΝΑΙ ΒΕΛΤΙΣΤΟ στη γενική περίπτωση, εκτός αν ρ = – 1. Για οποιαδήποτε άλλη συσχέτιση $\rho \in (-1, 1]$, το P_2 είναι γνήσια μικρότερο του BNT_2.

Σχόλιο: Από τα προηγούμενα προκύπτει ότι το BNT_2 δεν είναι καινούριο, αφού το P_2 είναι γνωστό από το 2001.

4.1.2. ΣΥΓΚΡΙΣΗ BNT_2 ΚΑΙ AG_2

Αν $\mu_1 = \mu_2 = \mu$, έχουμε

$$2(AG_2 - \bar{\mu}) = \sqrt{2\sigma_1{}^2 + 2\sigma_2{}^2}$$

και

$$2(BNT_2 - \bar{\mu}) = \sigma_1 + \sigma_2.$$

Παρατηρούμε ότι $\sigma_1 + \sigma_2 \le \sqrt{2\sigma_1{}^2 + 2\sigma_2{}^2} \Leftrightarrow (\sigma_1 - \sigma_2)^2 \ge 0$

Η ισότητα ισχύει αν και μόνο αν $\sigma_1 = \sigma_2$.

Αν $\sigma_1 = \sigma_2 = \sigma$, τότε, χωρίς να χρειάζεται $\mu_1 = \mu_2$, έχουμε:

$$2(AG_2 - \bar{\mu}) = \sqrt{4\sigma^2 + (\mu_1 - \mu_2)^2}$$

$$2(BNT_2 - \bar{\mu}) = \frac{1}{2\sigma}\left\{\sqrt{4\sigma^4 + \sigma^2(\mu_2 - \mu_1)^2} + \sqrt{4\sigma^4 + \sigma^2(\mu_2 - \mu_1)^2}\right\}$$

$$= \tfrac{1}{2}\sqrt{4\sigma^2 + (\mu_2 - \mu_1)^2} = 2(AG_2 - \bar{\mu}).$$

Δηλαδή $AG_2 = BNT_2$ ακόμα και αν $\mu_1 \neq \mu_2$.

Συμπέρασμα: Αν $\sigma_1 = \sigma_2$ τότε το AG είναι βέλτιστο (για n = 2)

Εδώ μπορούμε να διατυπώσουμε την παρακάτω πρόταση:

Πρόταση 4.1.3 Για όλες τις τιμές $\mu_1, \mu_2, \sigma_1{}^2 > 0, \sigma_2{}^2 > 0$,η διαφορά των φραγμάτων AG και BNT για n = 2 γίνεται

$$AG_2 - BNT_2 = \frac{(\sigma_1 - \sigma_2)^2}{2K},$$

όπου $K = \sqrt{(\mu_1 - \mu_2)^2 + (\sigma_1 + \sigma_2)^2} + \sqrt{(\mu_1 - \mu_2)^2 + 2\sigma_1{}^2 + 2\sigma_2{}^2}$

Απόδειξη:

Έστω AG_2 και BNT_2 είναι τα φράγματα για τη μέγιστη διατεταγμένη παρατήρηση όπως παρουσιάζονται στις σχέσεις (4.1) και (4.3). Η διαφορά AG – BNT_2 διαμορφώνεται ως εξής:

$$AG_2 - BNT_2 = \frac{1}{2}\sqrt{2\sigma_1{}^2 + 2\sigma_2{}^2 + \left(\mu_2 - \mu_1\right)^2} - \frac{1}{2}\sqrt{\left(\mu_2 - \mu_1\right)^2 + (\sigma_1 + \sigma_2)^2}$$

$$= \frac{1}{2}\left[\sqrt{2\sigma_1{}^2 + 2\sigma_2{}^2 + (\mu_2 - \mu_1)^2} - \sqrt{(\mu_2 - \mu_1)^2 + (\sigma_1 + \sigma_2)^2}\right]$$

$$= \frac{1}{2}\frac{2\sigma_1{}^2 + 2\sigma_2{}^2 + (\mu_2 - \mu_1)^2 - (\mu_2 - \mu_1)^2 - (\sigma_1 + \sigma_2)^2}{\sqrt{2\sigma_1{}^2 + 2\sigma_2{}^2 + (\mu_2 - \mu_1)^2} + \sqrt{(\mu_2 - \mu_1)^2 + (\sigma_1 + \sigma_2)^2}}$$

$$= \frac{(\sigma_1 - \sigma_2)^2}{2\left[\sqrt{2\sigma_1{}^2 + 2\sigma_2{}^2 + (\mu_2 - \mu_1)^2} + \sqrt{(\mu_2 - \mu_1)^2 + (\sigma_1 + \sigma_2)^2}\right]}. \qquad \blacksquare$$

Παρατηρούμε ότι $AG_2 - BNT_2 = 0$, αν και μόνο αν $(\sigma_1 - \sigma_2)^2 = 0$. Δηλαδή όταν $\sigma_1 = \sigma_2 = \sigma$.

Στην περίπτωση που το τ.δ είναι το (X_1, X_2), τότε το βέλτιστο φράγμα επιτυγχάνεται όταν $\sigma_1 = \sigma_2 = \sigma$ χωρίς να μας ενδιαφέρει η ισότητα των μ_1, μ_2. Στη γενική περίπτωση ισχύει το παρακάτω θεώρημα.

Θεώρημα 2.2.2. Υπάρχει μία μόνο μία κατανομή που επιτυγχάνει το AG (η οποία είναι η BNT), όταν $\mu_i = \mu$ και $\sigma_i{}^2 = \sigma^2$.

Απόδειξη

Έστω g η συνάρτηση που αναφέρεται στην πρόταση 3.2.2.

$$g'(t_o) = 0 \Leftrightarrow \sum_{i=1}^{n} \frac{t_o - \mu_i}{\sqrt{(t_o - \mu_i)^2 + \sigma_i{}^2}} = n - 2\ .$$

Αν $\mu_i = \mu$ και $\sigma_i{}^2 = \sigma^2$ για κάθε i

$$(3.7) \implies \quad n\,\frac{t_o - \mu}{\sqrt{(t_o - \mu)^2 + \sigma^2}} = n - 2$$

$$\Leftrightarrow \quad \frac{t_o - \mu}{\sqrt{(t_o - \mu)^2 + \sigma^2}} = \frac{n-2}{n}\,. \qquad (*)$$

Αν n = 2 $\overset{(*)}{\Rightarrow} \mathrm{t_o} = 0$.

Αν $\mathrm{n} > 2 \overset{(*)}{\Rightarrow} \mathrm{t_o} > \mu$.

$$(*) \quad \Leftrightarrow \quad \frac{1}{\sqrt{1+\frac{\sigma^2}{(\mathrm{t_o}-\mu)^2}}} = \frac{\mathrm{n}-2}{\mathrm{n}}$$

$$\Leftrightarrow \quad \sqrt{1+\frac{\sigma^2}{(\mathrm{t_o}-\mu)^2}} = \frac{\mathrm{n}}{\mathrm{n}-2}$$

$$\Rightarrow \quad 1+\frac{\sigma^2}{(\mathrm{t_o}-\mu)^2} = \left(\frac{\mathrm{n}}{\mathrm{n}-2}\right)^2$$

$$\Rightarrow \quad \frac{\sigma^2}{(\mathrm{t_o}-\mu)^2} = \frac{\mathrm{n}^2-(\mathrm{n}-2)^2}{(\mathrm{n}-2)^2}$$

$$\Leftrightarrow \quad \frac{\sigma^2}{(\mathrm{t_o}-\mu)^2} = \frac{4(\mathrm{n}-1)}{(\mathrm{n}-2)^2}$$

$$\Leftrightarrow \quad \frac{(\mathrm{t_o}-\mu)^2}{\sigma^2} = \frac{(\mathrm{n}-2)^2}{4(\mathrm{n}-1)}$$

$$\Leftrightarrow \quad |\mathrm{t_o}-\mu| = \frac{\mathrm{n}-2}{2\sqrt{\mathrm{n}-1}}\sigma$$

$$\overset{\mathrm{t_o}>\mu}{\Longrightarrow} \quad \mathrm{t_o} = \mu + \frac{\mathrm{n}-2}{2\sqrt{\mathrm{n}-1}}\sigma.$$

Οπότε αν $\mathrm{t} = \mathrm{t_o}$, η συνάρτηση g γίνεται:

$$\mathrm{g}(\mathrm{t_o}) = \mathrm{t_o} + \frac{\mathrm{n}}{2}(\mu - \mathrm{t_o}) + \frac{\mathrm{n}}{2}\sqrt{(\mu-\mathrm{t_o})^2+\sigma^2}$$

Αντικαθιστώντας τη σχέση $\mathrm{t_o} - \mu = \frac{\mathrm{n}-2}{2\sqrt{\mathrm{n}-1}}\sigma$ στην παραπάνω σχέση έχουμε:

$$\mathrm{g}(\mathrm{t_o}) = \mu + \frac{\mathrm{n}-2}{2\sqrt{\mathrm{n}-1}}\sigma + \frac{\mathrm{n}}{2}\left(\mu-\mu-\frac{\mathrm{n}-2}{2\sqrt{\mathrm{n}-1}}\sigma\right) + \frac{\mathrm{n}}{2}\sqrt{\left(\mu-\mu-\frac{\mathrm{n}-2}{2\sqrt{\mathrm{n}-1}}\sigma\right)^2+\sigma^2}$$

$$= \mu + \frac{\mathrm{n}-2}{2\sqrt{\mathrm{n}-1}}\sigma + \frac{\mathrm{n}}{2}\left(-\frac{\mathrm{n}-2}{2\sqrt{\mathrm{n}-1}}\sigma\right) + \frac{\mathrm{n}}{2}\sqrt{\frac{(\mathrm{n}-2)^2}{4(\mathrm{n}-1)}\sigma^2+\sigma^2}$$

$$= \mu - \frac{\mathrm{n}-2}{2\sqrt{\mathrm{n}-1}}\sigma\left(\frac{\mathrm{n}}{2}-1\right) + \frac{\mathrm{n}}{4\sqrt{\mathrm{n}-1}}\sigma\sqrt{(\mathrm{n}-2)^2+4(\mathrm{n}-1)}$$

$$= \mu - \frac{(\mathrm{n}-2)^2}{4\sqrt{\mathrm{n}-1}}\sigma + \frac{\mathrm{n}\sigma}{4\sqrt{\mathrm{n}-1}}\sqrt{\mathrm{n}^2}$$

$$= \mu - \frac{(\mathrm{n}-2)^2}{4\sqrt{\mathrm{n}-1}}\sigma + \frac{\mathrm{n}^2}{4\sqrt{\mathrm{n}-1}}\sigma$$

$$= \mu + \frac{\sigma}{4\sqrt{n-1}}[n^2 - (n-2)^2]$$

$$= \mu + \frac{\sigma}{4\sqrt{n-1}}4(n-1)$$

$$= \mu + \sigma\sqrt{n-1}.$$

Δηλαδή βρήκαμε το φράγμα της (2.10).

Αν όλα τα $\mu_i, {\sigma_i}^2$ είναι ίδια, τότε όλες οι p_i είναι ίδιες $\Longrightarrow p_i = \frac{1}{n}$. Επίσης όλα τα $c_i = c$ είναι ίδια, και ισχύει

$$\mathbb{P}(|X_i - t_o| = c_i) = 1.$$

Άρα

$$c = \sqrt{(t_o - \mu)^2 + \sigma^2}$$

$$= \sqrt{\tfrac{(n-2)^2}{4(n-1)}\sigma^2 + \sigma^2}$$

$$= \tfrac{1}{2\sqrt{n-1}}\sigma\sqrt{(n-2)^2 + 4(n-1)}$$

δηλαδή

$$c = \frac{n}{2\sqrt{n-1}}\sigma$$

Επομένως, <u>η μόνη κατανομή</u> που πετυχαίνει το φράγμα AG = BNT (όταν $\mu_i = \mu$ και ${\sigma_i}^2 = \sigma^2$) είναι η εξής:

$$\mathbb{P}\begin{bmatrix} X_1 = \mu - \frac{\sigma}{\sqrt{n-1}} \\ \vdots \\ X_i = \mu + \sigma\sqrt{n-1} \\ \vdots \\ X_n = \mu - \frac{\sigma}{\sqrt{n-1}} \end{bmatrix} = \frac{1}{n}.$$

Αφού

$$t_o - c = \mu + \frac{n-2}{2\sqrt{n-1}}\sigma - \frac{n}{2\sqrt{n-1}}\sigma$$

$$= \mu - \tfrac{\sigma}{\sqrt{n-1}}$$

έπεται ότι

$$t_o + c = \mu + \frac{n-2}{2\sqrt{n-1}}\sigma + \frac{n}{2\sqrt{n-1}}\sigma$$

$$= \mu + \sigma\sqrt{n-1}$$

■

Σχόλιο: Για τα φράγματα που μελετήσαμε στην ενότητα αυτή ισχύει η ανισότητα

$$P_2 \leq BNT_2 \leq AG_2.$$

Στην πραγματικότητα

$$P_2 = \sup_{\mathcal{F}(\mu_1,\mu_2,\sigma_1{}^2,\sigma_2{}^2,\rho)} E(X_{2:2})$$

$$BNT_2 = \sup_{\mathcal{F}(\mu_1,\mu_2,\sigma_1{}^2,\sigma_2{}^2,\rho=-1 \text{ ή χωρίς συσχέτιση})} E(X_{2:2}).$$

Το πρώτο έχει περισσότερους περιορισμούς στο supremum, άρα είναι καλύτερο φράγμα.

4.2. Άνω φράγμα για το εύρος (Range) δύο διατεταγμένων παρατηρήσεων.

Στην περίπτωση που n = 2, η απευθείας μελέτη των φραγμάτων για την αναμενόμενη τιμή του μεγίστου, απλοποιείται σημαντικά επειδή το $\max\{X_1, X_2\}$ συνδέεται με το range $R_2 = X_{2:2} - X_{1:2}$.

Συγκεκριμένα, επειδή

$$X_1 + X_2 = \max\{X_1, X_2\} + \min\{X_1, X_2\}$$

και

$$|X_2 - X_1| = \max\{X_1, X_2\} - \min\{X_1, X_2\} := R_2,$$

έπεται ότι

$$\begin{aligned} X_{2:2} &:= \max\{X_1, X_2\} \\ &= \frac{X_1+X_2}{2} + \frac{|X_1-X_2|}{2} \\ &= \frac{X_1+X_2}{2} + \frac{1}{2}R_2 \,. \end{aligned}$$

Επομένως λαμβάνοντας μέσες τιμές, προκύπτει η εξής γραμμική σχέση

$$2\left(E(X_{2:2}) - \frac{\mu_1+\mu_2}{2}\right) = E(R_2) = E|X_1 - X_2| \qquad (2.17)$$

Η σχέση αυτή δείχνει ότι οι ακραίες τιμές του μεγίστου και του εύρους σχετίζονται γραμμικά μεταξύ τους και μπορούν να βρεθούν απευθείας μεγιστοποιώντας ή ελαχιστοποιώντας την $\mathrm{E}|\mathrm{X}_1 - \mathrm{X}_2|$.

Όμως

$$\mathrm{E}|\mathrm{X}_1 - \mathrm{X}_2| \leq \sqrt{\mathrm{E}(\mathrm{X}_1 - \mathrm{X}_2)^2}$$

με ισότητα να επιτυγχάνεται αν και μόνο αν $\mathrm{Var}|\mathrm{X}_1 - \mathrm{X}_2| = 0$.

Η μέση τιμή της $\mathrm{E}(\mathrm{X}_1 - \mathrm{X}_2)^2$ εκφράζεται συναρτήσει των $\mu_1 \in \mathbb{R}, \mu_2 \in \mathbb{R}, \sigma_1 > 0$, $\sigma_2 > 0$ *και* $\rho \in [-1,\ 1]$ (τα οποία θεωρούνται γνωστά).

Πράγματι έχουμε

$$\begin{aligned}(\mathrm{X}_1 - \mathrm{X}_2)^2 &= [\mathrm{X}_1 - \mu_1 - (\mathrm{X}_2 - \mu_2) + (\mu_1 - \mu_2)]^2 \\ &= (\mu_1 - \mu_2)^2 + [(\mathrm{X}_1 - \mu_1) - (\mathrm{X}_2 - \mu_2)]^2 \\ &\quad +2(\mu_1 - \mu_2)[(\mathrm{X}_1 - \mu_1) - (\mathrm{X}_2 - \mu_2)] \\ &= (\mu_1 - \mu_2)^2 + (\mathrm{X}_1 - \mu_1)^2 + (\mathrm{X}_2 - \mu_2)^2 - 2(\mathrm{X}_1 - \mu_1)(\mathrm{X}_2 - \mu_2) \\ &\quad +2(\mu_1 - \mu_2)[(\mathrm{X}_1 - \mu_1) - (\mathrm{X}_2 - \mu_2)]\end{aligned}$$

Παίρνοντας μέσες τιμές, προκύπτει ότι

$$\mathrm{E}(\mathrm{X}_1 - \mathrm{X}_2)^2 = (\mu_1 - \mu_2)^2 + {\sigma_1}^2 + {\sigma_2}^2 - 2\rho\sigma_1\sigma_2.$$

Εναλλακτικά, θέτοντας στην (2.15) για n = 2, $\lambda_1 = -1$ και $\lambda_2 = 1$ προκύπτει απευθείας η ανισότητα

$$\begin{aligned}\mathrm{E}(\mathrm{R}_2) &\leq \sqrt{(\mu_1 - \mu_2)^2 + {\sigma_1}^2 + {\sigma_2}^2 - 2\rho\sigma_1\sigma_2} \\ &= \sqrt{\mathrm{E}(\mathrm{X}_1 - \mathrm{X}_2)^2}.\end{aligned}$$

Όπως είπαμε και σε προηγούμενη ενότητα, κάθε φράγμα από τα P_2 και BNT_2 είναι το supremum του $\mathrm{E}(\mathrm{X}_{2:2})$. Για το πρώτο δεν έχουμε κάποιους περιορισμούς, ενώ για το δεύτερο, θα πρέπει ρ = – 1, ή να μην υπάρχει συσχέτιση μεταξύ των τιμών.

$$\mathrm{P}_2 = \sup_{\mathcal{F}(\mu_1,\mu_2,\sigma_1\ \ ,\sigma_2\ \ ,\rho)} \mathrm{E}(\mathrm{X}_{2:2})$$

$$\mathrm{BNT}_2 = \sup_{\mathcal{F}(\mu_1,\mu_2,\sigma_1\ \ ,\sigma_2\ \ ,\rho=-1 \text{ ή χωρίς συσχέτιση})} \mathrm{E}(\mathrm{X}_{2:2})$$

Για να πετύχουμε το φράγμα θα πρέπει, όπως έχουμε προαναφέρει άλλωστε,

$$\mathrm{Var}|\mathrm{X}_1 - \mathrm{X}_2| = 0.$$

Το οποίο είναι ισοδύναμο με το να δείξουμε ότι

$$\mathbb{P}(|X_1 - X_2| = c) = 1.$$

Υπάρχουν πράγματι τ.μ $X_i, i = 1, 2$ με

$$E(X_i) = \mu_i, \qquad Var(X_i) = \sigma_i{}^2 \text{ και } Cov(X_1, X_2) = \rho\sigma_1\sigma_2$$

ώστε να επιτυγχάνεται το φράγμα;

Ας ονομάσουμε $\theta = \sqrt{(\mu_1 - \mu_2)^2 + \sigma_1{}^2 + \sigma_2{}^2 - 2\rho\sigma_1\sigma_2} \geq 0$.

Δηλαδή

$$E(R_2) \leq \theta = \sqrt{(\mu_1 - \mu_2)^2 + \sigma_1{}^2 + \sigma_2{}^2 - 2\rho\sigma_1\sigma_2}.$$

Διακρίνουμε τις παρακάτω περιπτώσεις

(A) Όταν $\sigma_1{}^2 + \sigma_2{}^2 > 2\rho\sigma_1\sigma_2$ (δηλαδή $\theta > |\mu_1 - \mu_2|$),

Ορίζουμε

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = \frac{1}{\sigma_1{}^2 + \sigma_2{}^2 - 2\rho\sigma_1\sigma_2} \begin{pmatrix} (\sigma_1{}^2 - \rho\sigma_1\sigma_2)B_p + W \\ (\rho\sigma_1\sigma_2 - \sigma_2{}^2)B_p + W \end{pmatrix}$$

όπου η B_p ακολουθεί μία κατανομή $\mathbb{P}(B_p = \theta) = p$ και $\mathbb{P}(B_p = -\theta) = 1 - p$ με $p = \frac{1}{2}\left(1 + \frac{\mu_1 - \mu_2}{\theta}\right)$ και η W οποιαδήποτε τ.μ ανεξάρτητη της B_p έτσι ώστε

$$E(W) = \mu_1(\sigma_2^2 - \rho\sigma_1\sigma_2) + \mu_1(\sigma_1^2 - \rho\sigma_1\sigma_2)$$

και

$$Var(W) = \sigma_1^2\sigma_2^2(\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2)(1 - \rho^2).$$

Οπότε τελικά έχουμε ότι

$$E(X_1) = \mu_1,\ E(X_2) = \mu_2,\ Var(X_1) = \sigma_1{}^2,\ Var(X_2) = \sigma_2{}^2, cov(X_1, X_2) = \rho\sigma_1\sigma_2 \text{ και}$$

$$E|X_1 - X_2| = E\left|B_p\right| = |\theta|p + |-\theta|(1 - p) = |\theta| = \theta.$$

(B) Όταν $\sigma_1{}^2 + \sigma_2{}^2 = 2\rho\sigma_1\sigma_2$ και $\theta > 0$, τότε $\mu_1 \neq \mu_2$.

$$Var(X_1 - X_2) = \sigma_1{}^2 + \sigma_2{}^2 - 2\rho\sigma_1\sigma_2 = 0$$

$$\Longrightarrow X_1 - X_2 = \gamma \text{ με πιθανότητα 1.}$$

Ο πίνακας

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_1{}^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2{}^2 \end{bmatrix}$$

είναι θετικά ορισμένος, αφού $\det\begin{bmatrix} \sigma_1{}^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2{}^2 \end{bmatrix} \geq 0.$

Θέτω $\sigma_1 = \sigma_2 = \sigma$, οπότε ο πίνακας συνδιακύμανσης γίνεται

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma^2 & \sigma^2 \\ \sigma^2 & \sigma^2 \end{bmatrix}$$

Θέλω να βρω $\begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$ με $E\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = \boldsymbol{\mu}$ και $\mathrm{Var}\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = \boldsymbol{\Sigma}$,

όπου

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \text{ και } \boldsymbol{\Sigma} = \begin{bmatrix} \sigma^2 & \sigma^2 \\ \sigma^2 & \sigma^2 \end{bmatrix}$$

έτσι ώστε $X_1 - X_2 = \gamma$ με πιθανότητα 1.

Έστω W μια τυχαία μεταβλητή ανεξάρτητη από τα X_1, X_2 με $E(W) = \mu_1$ και $\mathrm{Var}(W) = \sigma^2 > 0$.

Θεωρώ την κατανομή

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = \begin{pmatrix} W \\ W + \mu_2 - \mu_1 \end{pmatrix}.$$

Παρατηρούμε ότι όντως ικανοποιεί της συνθήκες που θέλουμε, και μάλιστα $\gamma = \mu_1 - \mu_2 \neq 0$.

Οπότε

$$E|X_1 - X_2| = |\mu_1 - \mu_2| = \theta.$$

(Γ) Αν $\theta = 0$, τότε προκύπτει ότι $\mu_1 = \mu_2 = \mu$. Αν $\sigma_1{}^2 = \sigma_2{}^2 = \sigma^2$ τότε $\rho = 1$.

Θεωρούμε τώρα την κατανομή $W = X_1 = X_2$, τέτοια ώστε

$$E(W) = \mu \quad \text{και } \mathrm{Var}(W) = \sigma^2 > 0.$$

Οπότε και σε αυτήν την τετριμμένη περίπτωση, ικανοποιούνται οι συνθήκες που έχουμε θεωρήσει για την κατανομή.

Το παρακάτω Θεώρημα, μας δείχνει δύο βέλτιστα φράγματα για το Range (R_2).

Θεώρημα 4.2.1. Για το εύρος δύο διατεταγμένων παρατηρήσεων ισχύουν τα φράγματα

$$\inf \mathrm{E}(\mathrm{R}_2) = |\mu_1 - \mu_2|$$

και

$$\sup \mathrm{E}(\mathrm{R}_2) = \theta = \sqrt{(\mu_1 - \mu_2)^2 + {\sigma_1}^2 + {\sigma_2}^2 - 2\rho\sigma_1\sigma_2}.$$

Λήμμα 4.2.2 Έστω $Y \in \mathcal{F}(\mu, \sigma)$, δηλαδή $EY = \mu \in \mathbb{R}$, και $\mathrm{Var}Y = \sigma^2 > 0$.

(α) Αν $\mu = 0$, τότε $E|Y| > 0$ και $\inf_{Y \in \mathcal{F}(\mu,\sigma)} E|Y| = 0$.

(β) Αν $\mu \neq 0$, τότε $E|Y| \geq |\mu|$ και $\min_{Y \in \mathcal{F}(\mu,\sigma)} E|Y| = |\mu|$.

Απόδειξη Λήμματος

(α) Είναι γνωστό ότι $E|Y| \geq |EY| = |\mu|$.

Αν $\mu = 0$, τότε προκύπτει ότι $E|Y| \geq 0$. Έστω ότι $E|Y| = 0$ τότε $Y = 0$ με πιθανότητα 1. Δηλαδή η Y είναι εκφυλισμένη[12], με αποτέλεσμα να έχουμε Var(Y) = 0. Το οποίο είναι άτοπο. Άρα $E|Y| > 0$.

Θεωρούμε την ακολουθία

$$Y_n = \begin{cases} -n\sigma, & \text{με πιθανότητα } \frac{1}{2n^2}, \\ 0, & \text{με πιθανότητα } 1 - \frac{1}{n^2}, \quad n = 1, 2, \ldots \\ n\sigma, & \text{με πιθανότητα } \frac{1}{2n^2}. \end{cases}$$

για την οποία ισχύουν τα εξής:

$E(Y_n) = 0$ γιατί είναι συμμετρική, και $\mathrm{Var}(Y_n) = \sigma^2$. Το οποίο σημαίνει ότι $Y_n \in \mathcal{F}(0, \sigma)$ για κάθε n, και $E|Y_n| = \frac{\sigma}{n} \longrightarrow 0$ καθώς $n \longrightarrow \infty$. Άρα το $\inf_{Y \in \mathcal{F}(0,\sigma)} E|Y| = 0$. Μπορώ να βρω Y, έτσι ώστε να πιάνεται το minimum; Την απάντηση θα τη δώσουμε στο (β).

(β) Αν $\mu > 0$, ορίζουμε την

$$Y = \begin{cases} 0, & \text{με πιθανότητα } \frac{\sigma^2}{\mu^2 + \sigma^2}, \\ \mu + \frac{\sigma^2}{\mu}, & \text{με πιθανότητα } \frac{\mu^2}{\mu^2 + \sigma^2}. \end{cases}$$

Έχουμε υποθέσει ότι $\mu > 0$, άρα $Y > 0$. Η υπόθεση αυτή μας δίνει την ισότητα

---

[12] Μια τ.μ X ονομάζεται εκφυλισμένη, όταν υπάρχει $\alpha \in \mathbb{R}$, τέτοιο ώστε X=α με πιθανότητα 1. Δηλαδή αν η X είναι εκφυλισμένη, τότε EX=α και VarX=0.

$$E|Y| = EY = \mu = |\mu|.$$

Επίσης, $E(Y^2) = \left(\mu + \frac{\sigma^2}{\mu}\right)^2 \left(\frac{\mu^2}{\mu^2+\sigma^2}\right) = \mu^2 + \sigma^2$. Οπότε η διασπορά Var(X) θα είναι

$\mathrm{Var}(X) = E(Y^2) - \big(E(Y)\big)^2 = \sigma^2.$

Στην περίπτωση που το $\mu < 0$, τότε έχουμε ισότητα για την ίδια Y, που τώρα όμως είναι μικρότερη ή ίση από το μηδέν με πιθανότητα 1.

■

Πόρισμα 4.2.3 Αν $\mathbf{X} = (X_1, X_2) \in \mathcal{F}(\mu_1, \mu_2, \sigma_1, \sigma_2, \rho)$, τότε ισχύει ότι

$$\inf_{(X_1,X_2)\in\mathcal{F}(\mu_1,\mu_2,\sigma_1,\sigma_2,\rho)} E|X_1 - X_2| = |\mu_1 - \mu_2|.$$

Απόδειξη

(A) Αν $2\rho\sigma_1\sigma_2 < \sigma_1^2 + \sigma_2^2$ τότε διακρίνουμε τις εξής υποπεριπτώσεις

(A1) Αν $\mu_1 \neq \mu_2$ θεωρούμε την κατανομή

$$Y = \begin{cases} 0, & \text{με πιθανότητα } \frac{\sigma_1^2+\sigma_2^2-2\rho\sigma_1\sigma_2}{(\mu_1-\mu_2)^2+\sigma_1^2+\sigma_2^2-2\rho\sigma_1\sigma_2}, \\ \frac{(\mu_1-\mu_2)^2+\sigma_1^2+\sigma_2^2-2\rho\sigma_1\sigma_2}{\mu_1-\mu_2}, & \text{με πιθανότητα } \frac{(\mu_1-\mu_2)^2}{(\mu_1-\mu_2)^2+\sigma_1^2+\sigma_2^2-2\rho\sigma_1\sigma_2}. \end{cases}$$

$$E(Y) = \frac{(\mu_1-\mu_2)^2+\sigma_1^2+\sigma_2^2-2\rho\sigma_1\sigma_2}{\mu_1-\mu_2} \cdot \frac{(\mu_1-\mu_2)^2}{(\mu_1-\mu_2)^2+\sigma_1^2+\sigma_2^2-2\rho\sigma_1\sigma_2} = \mu_1 - \mu_2,$$

$$E|Y| = \left|\frac{(\mu_1-\mu_2)^2+\sigma_1^2+\sigma_2^2-2\rho\sigma_1\sigma_2}{\mu_1-\mu_2}\right| \cdot \frac{(\mu_1-\mu_2)^2}{(\mu_1-\mu_2)^2+\sigma_1^2+\sigma_2^2-2\rho\sigma_1\sigma_2} = |\mu_1 - \mu_2|, \text{ και}$$

$$\mathrm{Var}(Y) = \left(\frac{(\mu_1-\mu_2)^2+\sigma_1^2+\sigma_2^2-2\rho\sigma_1\sigma_2}{\mu_1-\mu_2}\right)^2 \cdot \frac{(\mu_1-\mu_2)^2}{(\mu_1-\mu_2)^2+\sigma_1^2+\sigma_2^2-2\rho\sigma_1\sigma_2} - (\mu_1 - \mu_2)^2$$

$$= \sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2 > 0.$$

Τέλος ορίζω την κατανομή

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = \begin{pmatrix} \frac{\alpha+1}{2} Y + \frac{1}{2} W \\ \frac{\alpha-1}{2} Y + \frac{1}{2} W \end{pmatrix}$$

όπου $\alpha = \frac{\sigma_1^2 - \sigma_2^2}{\sigma_1^2 + \sigma_2^2 - \rho\sigma_1\sigma_2}$ και Y, W ανεξάρτητες, με κατανομές όπως φαίνεται παραπάνω, δηλαδή

$$\mathbb{P}\left(Y = \frac{\theta^2}{\mu_1 - \mu_2}\right) = \left(\frac{\mu_1 - \mu_2}{\theta}\right)^2 = 1 - \mathbb{P}(Y = 0)$$

όπου $\theta = \sqrt{(\mu_1 - \mu_2)^2 + {\sigma_1}^2 + {\sigma_2}^2 - 2\rho\sigma_1\sigma_2}$,

$$E(W) = \mu_1(\sigma_2^2 - \rho\sigma_1\sigma_2) + \mu_1(\sigma_1^2 - \rho\sigma_1\sigma_2)$$

και

$$\mathrm{Var}(W) = \sigma_1^2\sigma_2^2(\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2)(1 - \rho^2).$$

Επίσης από το προηγούμενο Λήμμα έχουμε ότι, αφού $|\mu_1 - \mu_2| \neq 0$, θα πιάνεται το minimum.

$$E|X_1 - X_2| = E|Y| = |\mu_1 - \mu_2|.$$

(A2) Αν $\mu_1 = \mu_2$, θεωρούμε την κατανομή

$$Y_n = \begin{cases} -n\sqrt{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}, & \text{με πιθανότητα } \frac{1}{2n^2}, \\ 0, & \text{με πιθανότητα } 1 - \frac{1}{n^2}, \\ n\sqrt{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}, & \text{με πιθανότητα } \frac{1}{2n^2}, \end{cases}$$

για την οποία ισχύουν τα παρακάτω

$E(Y_n) = 0,\ \mathrm{Var}(Y_n) = \sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2$ και $E|Y_n| = \frac{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}{n} \longrightarrow 0$. Φτιάχνω

$$\begin{pmatrix} {X_1}^{(n)} \\ {X_2}^{(n)} \end{pmatrix} := \frac{1}{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2} \begin{pmatrix} ({\sigma_1}^2 - \rho\sigma_1\sigma_2)Y_n + W \\ (\rho\sigma_1\sigma_2 - {\sigma_2}^2)Y_n + W \end{pmatrix}$$

με Y_n, W ανεξάρτητες για κάθε n, με

$$E(W) = \mu_1(\sigma_2^2 - \rho\sigma_1\sigma_2) + \mu_1(\sigma_1^2 - \rho\sigma_1\sigma_2)$$

και

$$\mathrm{Var}(W) = \sigma_1^2\sigma_2^2(\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2)(1-\rho^2).$$

Τώρα ισχύει το εξής $\left|X_1^{(n)} - X_2^{(n)}\right| = |Y_n|$. Οι μέσες τιμές οι διασπορές των ακολουθιών και η συδιακύμανσή δίνονται από τους παρακάτω τύπους:

$$E\left(X_1^{(n)}\right) = \mu_1 = \mu = \mu_2 = E\left(X_2^{(n)}\right),$$

$$\mathrm{Var}\left(X_1^{(n)}\right) = \sigma_1^2,\ \ \mathrm{Var}\left(X_2^{(n)}\right) = \sigma_2^2,$$

$$\mathrm{Cov}\left(X_1^{(n)},\ X_2^{(n)}\right) = \rho\sigma_1\sigma_2.$$

για n = 1, 2, …

Τέλος,

$$E\left|X_1^{(n)} - X_2^{(n)}\right| = E|Y_n| = \frac{\sigma_1^2+\sigma_2^2-2\rho\sigma_1\sigma_2}{n} \longrightarrow 0, \text{ καθώς } n \longrightarrow \infty.$$

(B) Τέλος, όταν ${\sigma_1}^2 + {\sigma_2}^2 = 2\rho\sigma_1\sigma_2$, είχαμε δείξει ότι

$$\boldsymbol{\mu} = \begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix} \text{ και } \boldsymbol{\Sigma} = \begin{bmatrix}\sigma^2 & \sigma^2\\ \sigma^2 & \sigma^2\end{bmatrix}$$

και ορίζουμε $X_{\mathbf{1}} \in \mathcal{F}_1(\mu_1, \sigma)$ και $X_2 = X_1 + (\mu_2 - \mu_1)$. Για τις $X_1,\ X_2$ ισχύουν τα εξής

$E(X_i) = \mu_i$, $\mathrm{Var}(X_i) = \sigma^2$, i = 1, 2, και $|X_1 - X_2| = |\mu_1 - \mu_2|$.

■

<u>Πόρισμα 4.2.4.</u> Ακόμα και όταν δε δίνεται ο συντελεστής συσχέτισης, ο επιτυγχάνεται η ισότητα

$$\inf_{(X_1,X_2)\in\mathcal{F}(\mu_1,\mu_2)} E|X_1 - X_2| = |\mu_1 - \mu_2|.$$

■

Το παραπάνω πόρισμα ισχύει προφανώς, διότι

$$\mathcal{F}(\mu_1, \mu_2) \supseteq \cup_{\sigma_1>0} \cup_{\sigma_2>0} \cup_{\rho\in[-1,1]} \mathcal{F}(\mu_1, \mu_2, \sigma_1, \sigma_2, \rho).$$

Τώρα είμαστε σε θέση να αποδείξουμε το Θεώρημα 4.2.1.

Απόδειξη Θεωρήματος 4.2.1.

Αποδεικνύεται πλέον εύκολα χρησιμοποιώντας τα παραπάνω Λήμματα, τα Πορίσματα, και τις ανισότητες

$$|\mu_1 - \mu_2| = |EX_1 - EX_2| = |E(X_1 - X_2)| \leq E|X_1 - X_2| \leq ER_2 \leq \theta$$

και

$$\max\{\mu_1, \mu_2\} \leq EX_{2:2} \leq \frac{\mu_1+\mu_2}{2} + \frac{1}{2}\theta.$$

■

ΠΑΡΑΡΤΗΜΑ

ΚΑΤΑΣΚΕΥΗ ΦΡΑΓΜΑΤΩΝ

Στο κεφάλαιο αυτό θα βρούμε φράγματα σε ανισότητες δίνοντας μία κοινή μέθοδο. Επίσης θα μελετήσουμε το πρόβλημα της ύπαρξης μιας κατανομής πιθανότητας για να επιτυγχάνεται το φράγμα.

Π.1. Εισαγωγή

Έστω Ω ένας τυχαίος χώρος[13], $\mathcal{A}$ μία σ – άλγεβρα στον Ω και

$f(\omega) = (f_1(\omega),\ f_2(\omega), \dots ,\ f_{\mathrm{m}}(\omega))$ μία Borel μετρήσιμη απεικόνιση του Ω στον $\mathbb{R}^{\mathrm{m}}$. Ακόμη, θεωρούμε $\mathfrak{B}$ μία γνωστή κλάση από κατανομές πιθανότητας στο χώρο $(\Omega, \mathcal{A})$ έτσι ώστε κάθε μία από τις $f(\omega)$ να είναι ολοκληρώσιμη, A ένα μετρήσιμο σύνολο, και $\mathbf{M} = (\mathrm{M}_1, \dots, \mathrm{M}_{\mathrm{m}})$ ένα δοσμένο πραγματικό διάνυσμα. Αν μία ανισότητα της μορφής $\mathbb{P}(\mathrm{A}) \leq \alpha$ (ή $\mathbb{P}(\mathrm{A}) \geq \beta$) ισχύει για όλες τις κατανομές $\mathbb{P}$ στο $\mathfrak{B}$ που ικανοποιεί τη

$$\mathrm{E}_{\mathbb{P}}[\mathrm{f}] \equiv \int f(\omega)\mathrm{d}\mathbb{P}(\omega) = \mathbf{M}, \qquad (\Pi.1)$$

λέγεται Ανισότητα τύπου Tchebycheff. Ο πραγματικός αριθμός α (ή ο β), εξαρτάται από το **M** και το A. Η ανισότητα όπως είδαμε και στα προηγούμενα κεφάλαια, θα είναι βέλτιστη, αν δεν μπορούμε να αντικαταστήσουμε το α με ένα μικρότερο αριθμό α΄ (ή το β με ένα μεγαλύτερο β΄). Ένα τέτοιο α (ή β) λέγεται βέλτιστο άνω (ή κάτω) φράγμα για την $\mathbb{P}(\mathrm{A})$. Βέλτιστα φράγματα είναι, εξ ορισμού, αυτά που εκφράζονται ως εξής

$$\alpha = \sup_{\mathbb{P}\in\mathfrak{B}(\mathrm{M})}\mathbb{P}(\mathrm{A}),$$

$$\beta = \inf_{\mathbb{P}\in\mathfrak{B}(\mathrm{M})}\mathbb{P}(\mathrm{A}),$$

όπου $\mathfrak{B}(\mathrm{M})$ είναι μία υποκλάση του $\mathfrak{B}$ που αποτελείται από όλες τις κατανομές $\mathbb{P}$ στο $\mathfrak{B}$, τέτοιες ώστε

$$\mathrm{E}_{\mathbb{P}}[\mathrm{f}] = \mathbf{M}.$$

Μία από τις βασικές μεθόδους για να βρούμε φράγματα σε ανισώσεις του τύπου $\mathbb{P}(\mathrm{A}) \leq \alpha$, είναι η εξής:

Έστω $\mathbb{I}_{\mathrm{A}}$ η δείκτρια συνάρτηση του συνόλου A, και $\boldsymbol{\alpha} = (\alpha_0, \alpha_1, \dots, \alpha_{\mathrm{m}})$ ένα πραγματικό διάνυσμα στον $\mathbb{R}^{\mathrm{m}+1}$.

[13] Δε μας ενδιαφέρει τι χώρος είναι. Μπορεί να είναι τοπολογικός, δειγματικός, μετρήσιμος κλπ.

Τότε ισχύει η σχέση

$$f = \alpha_0 + \alpha_1 f_1 + \cdots + \alpha_m f_m \geq \mathbb{I}_A \text{ στον } \Omega,$$

από την οποία συνεπάγεται ότι

$$\mathbb{P}(A) = \int \mathbb{I}_A d\mathbb{P} \leq \int f d\mathbb{P} = \alpha_0 + \alpha_1 M_1 + \cdots + \alpha_m M_m \quad \text{για κάθε } \mathbb{P} \in \mathfrak{B}(\mathbf{M}).$$

Έτσι αποκτούμε την ανισότητα

$$\mathbb{P}(A) \leq \inf_{\boldsymbol{\alpha} \in A(A)} (\alpha_0 + \alpha_1 M_1 + \cdots + \alpha_m M_m) \tag{Π.2}$$

όπου

$$A(\mathrm{A}) = \{\boldsymbol{\alpha} : \alpha_0 + \alpha_1 f_1 + \cdots + \alpha_m f_m \geq \mathbb{I}_A\}.$$

Τώρα φυσικό είναι να ρωτήσουμε αν η ανισότητα (Π.2) είναι βέλτιστη ή όχι. Δυστυχώς η μέθοδος αυτή δε μπορεί να εφαρμοστεί σε γενικές περιπτώσεις. Η ανισότητα (Π.2) δεν είναι πάντα βέλτιστη, και ακόμη και αν είναι βέλτιστη, η ισότητα δεν είναι απαραίτητο να επιτυγχάνεται. Στο κεφάλαιο αυτό, θα δώσουμε μια γενική θεωρία για τις βέλτιστες ανισότητες. Τα ερωτήματα που πρέπει να απαντήσουμε είναι τα εξής:

(1) Για ποιο A και **M** η (Π.2) είναι βέλτιστη; Ποια είναι τότε η γενική μορφή των βέλτιστων φραγμάτων;
(2) Υπάρχει κατανομή $\mathbb{P}$ ώστε να επιτυγχάνεται ισότητα στην (Π.2); Υπάρχει **α** στο A(A) ώστε να επιτυγχάνει το infimum στο δεξιό σκέλος της (Π.2);
(3) Πόσο μπορούμε να μικρύνουμε ή να μεγαλώσουμε το μέγεθος της κλάσης $\mathfrak{B}$ χωρίς να αλλάξει η τιμή του βέλτιστου φράγματος;

Τα παραπάνω ερωτήματα θα τα απαντήσουμε στις επόμενες παραγράφους

Π.2. Βέλτιστα φράγματα

Θα διατυπώσουμε το πρόβλημα (1) της ενότητας Π.1. με μία πιο γενική μορφή ως εξης:

Υποθέτουμε ότι υπάρχει μία γνωστή οικογένεια $\mathfrak{B}$ από μη αρνητικά μέτρα[14] σε ένα μετρήσιμο χώρο $(\Omega, \mathfrak{B})$ και θεωρούμε $f(\omega) = (f_0(\omega), f_1(\omega), \ldots, f_m(\omega))$ μία Borel μετρήσιμη απεικόνιση από τον Ω στον $\mathbb{R}^{m+1}$, ώστε κάθε μέλος της οικογένειας $\mathfrak{B}$ να είναι ολοκληρώσιμο. Επιπρόσθετα, θέτουμε $g(\omega)$ μια πραγματική μετρήσιμη συνάρτηση, και ολοκληρώσιμη στο $\mathfrak{B}$. Για κάθε $\mathbf{M} \in \mathbb{R}^{m+1}$ ορίζουμε

[14] Δεν είναι απαραίτητο να είναι μέτρα πιθανότητας.

$$\mathfrak{B}(\mathbf{M}) = \{\mathbb{P} : \mathbb{P} \in \mathfrak{B} \text{ και } \int f(\omega) d\mathbb{P}(\omega) = \mathbf{M}\}.$$

Το πρόβλημα τώρα είναι να αναζητήσουμε

$$\mathrm{U}(\mathrm{g}, \mathbf{M}) \equiv \sup_{\mathbb{P} \in \mathfrak{B}(\mathrm{M})} \int \mathrm{g}(\omega) d\mathbb{P}(\omega),$$

$$\mathrm{L}(\mathrm{g}, \mathbf{M}) \equiv \inf_{\mathbb{P} \in \mathfrak{B}(\mathrm{M})} \int \mathrm{g}(\omega) d\mathbb{P}(\omega).$$

<u>Θεώρημα Π1</u> Αν το $\mathfrak{B}$ είναι κυρτό[15], τότε το $\mathfrak{M} = \{\int f(\omega) d\mathbb{P}(\omega) : \mathbb{P} \in \mathfrak{B}\}$ είναι κυρτό σύνολο στον $\mathbb{R}^{m+1}$. Αν $\mathbf{M} = (\mathrm{M}_0, \dots, \mathrm{M}_m)$ είναι ένα εσωτερικό σημείο του $\mathfrak{M}$, τότε έχουμε

$$\mathrm{U}(\mathrm{g}, \mathbf{M}) = \inf_{\alpha_0, \alpha_1, \dots, \alpha_m} \{\sum_{i=0}^m \alpha_i \mathrm{M}_i + \sup_{\mathbb{P} \in \mathfrak{B}} \int (\mathrm{g} - \sum_{i=0}^m \alpha_i f_i) d\mathbb{P}\}. \quad (\Pi.3)$$

Επιπλέον, αν $\mathrm{U}(\mathrm{g}, \mathbf{M}) < \infty$, τότε υπάρχουν $\alpha_0, \alpha_1, \dots, \alpha_m$ ώστε να επιτυγχάνεται το infimum στο δεξί μέλος της σχέσης (Π.3). Δεν υπάρχει πάντα μέτρο $\mathbb{P}$ στο $\mathfrak{B}(\mathbf{M})$ ώστε να επιτυγχάνεται το φράγμα, αλλά, αν υπάρχει θα πρέπει να επιτυγχάνεται το supremum στην (Π.3) για τα $\alpha_0, \alpha_1, \dots, \alpha_m$ που επιτυγχάνεται το infimum στην (Π.3).

Απόδειξη

Όταν το $\mathfrak{B}$ είναι κυρτό, το $\mathfrak{M}$ είναι προφανώς κυρτό.

Για κάθε $\alpha_0, \alpha_1, \dots, \alpha_m$, έχουμε

$$\sup_{\mathbb{P} \in \mathfrak{B}} \int (\mathrm{g} - \sum_{i=0}^m \alpha_i f_i) d\mathbb{P} \geq \sup_{\mathbb{P} \in \mathfrak{B}(\mathbf{M})} \int (\mathrm{g} - \sum_{i=0}^m \alpha_i f_i) d\mathbb{P}$$

$$= \mathrm{U}(\mathrm{g}, \mathbf{M}) - \sum_{i=0}^m \alpha_i \mathrm{M}_i.$$

Ως εκ τούτου, το αριστερό μέλος της (Π.3) δεν είναι μεγαλύτερο από το δεξί.

Η αντίστροφη ανισότητα αποδεικνύεται ως εξής.

Το σύνολο

$$\mathfrak{N} = \{(\int f_0 d\mathbb{P}, \dots, \int f_m d\mathbb{P}, \int \mathrm{g} d\mathbb{P}) : \mathbb{P} \in \mathfrak{B}\}$$

είναι, για τον ίδιο λόγο όπως την περίπτωση του $\mathfrak{M}$, ένα κυρτό σύνολο στο $\mathbb{R}^{m+1}$. Αν $\mathrm{U} \equiv \mathrm{U}(\mathrm{g}, \mathbf{M}) < \infty$, το σημείο $\mathrm{Q} = (\mathrm{M}_0, \dots, \mathrm{M}_m, \mathrm{U})$ είναι συνοριακό σημείο του $\mathfrak{N}$. Ως εκ τούτου, υπάρχει ένα υπερεπίπεδο Π που στηρίζει το $\mathfrak{N}$ στο Q. Έστω η εξίσωση του Π είναι η

$$\sum_{i=0}^{m+1} \alpha_i x_i + \beta = 0.$$

[15] Αν $\mathbb{P}_1, \mathbb{P}_2 \in \mathfrak{B}$ τότε $\lambda\mathbb{P}_1 + (1-\lambda)\mathbb{P}_2 \in \mathfrak{B}$ για κάθε $\lambda \in [0,1]$.

Αφού το $\mathbf{M}$ το υποθέσαμε να είναι ένα εσωτερικό σημείο του $\mathfrak{M}$, το οποίο είναι η προβολή του $\mathfrak{N}$ στον $\mathbb{R}^{m+1}$, μπορούμε να πάρουμε $\alpha_{m+1} \neq 0$. Οπότε το Π εκφράζεται

$$x_{m+1} = c + \sum_{i=0}^{m} \alpha_i x_i,$$

και κάθε σημείο του $\mathfrak{N}$ ικανοποιεί τη

$$x_{m+1} \leq c + \sum_{i=0}^{m} \alpha_i x_i,$$

το οποίο σημαίνει ότι, για κάθε $\mathbb{P} \in \mathfrak{B}$, έχουμε

$$\int g d\mathbb{P} \leq c + \sum_{i=0}^{m} \alpha_i \int f_i d\mathbb{P}.$$

Αφού το Q βρίσκεται στο Π,

$$U = c + \sum_{i=0}^{m} \alpha_i M_i.$$

Απαλείφοντας το c, έχουμε, για κάθε $\mathbb{P} \in \mathfrak{B}$,

$$U \geq \int g d\mathbb{P} - \sum_{i=0}^{m} \alpha_i \int f_i d\mathbb{P} + \sum_{i=0}^{m} \alpha_i M_i.$$

Έτσι δείξαμε την (Π.3) καθώς και την ύπαρξη των $\alpha_0, \alpha_1, \ldots, \alpha_m$ που επιτυγχάνεται το infimum. Το τελευταίο ζητούμενο του θεωρήματος είναι προφανές.

■

Παρατήρηση Για το κάτω φράγμα $L(g, \mathbf{M})$, θα θέσουμε στην (Π.3) το $-g$, και προκύπτει αμέσως το δυϊκό αποτέλεσμα.

■

Τώρα θα εφαρμόσουμε το Θεώρημα Π1, σε μερικές περιπτώσεις

1. Θεωρούμε την περίπτωση που το $\mathfrak{B}$ αποτελείται από όλα τα μη αρνητικά μέτρα[16] έτσι ώστε οι $f_0, \ldots, f_m$, g να είναι ολοκληρώσιμες.
   Αν σταθεροποιήσω τα $\alpha_0, \alpha_1, \ldots, \alpha_m$,

$$\sup_{\mathbb{P}\in\mathfrak{B}} \int \left(g - \sum_{i=0}^{m} \alpha_i f_i\right) d\mathbb{P} = \begin{cases} 0, & \text{αν } g(\omega) - \sum \alpha_i f_i(\omega) \leq 0 \text{ στον } \Omega \\ \infty, & \text{διαφορετικά} \end{cases}$$

   Οπότε

$$\inf_{\alpha_0,\alpha_1,\ldots,\alpha_m}\left\{\sum_{i=0}^{m} \alpha_i M_i + \sup_{\mathbb{P}\in\mathfrak{B}} \int \left(g - \sum_{i=0}^{m} \alpha_i f_i\right) d\mathbb{P}\right\} = \inf{}' \sum_{i=0}^{m} \alpha_i M_i,$$

[16] Δε χρειάζεται να υποθέσουμε ότι είναι μέτρα πιθανότητας, αφού το $\mathfrak{B}(\mathbf{M})$ τα περιορίζει αυτόματα σε μέτρα πιθανότητας αν υποθέσουμε $f_0(\omega) \equiv 1$ και $M_0=1$. Γι' αυτό το λόγο αναφέρουμε το πρόβλημα στη γενική μορφή.

όπου inf ′ θεωρούμε το infimum ως προς όλα τα διανύσματα $(\alpha_0, \alpha_1, \dots, \alpha_m)$ έτσι ώστε $\sum \alpha_i f_i(\omega) \geq g(\omega)$ στον Ω. Έτσι αποκτάμε τη σχέση

$$U(g, \mathbf{M}) = \inf{}' \sum_{i=0}^{m} \alpha_i M_i \tag{Π.4}$$

Π.3. Εύρεση φραγμάτων για αναμενόμενες τιμές παρατηρήσεων.

Σε αυτή την ενότητα του παραρτήματος, θα «μεταφράσουμε» τα αποτελέσματα που βρήκαμε, στις ενότητες Π1 και Π2 στη γλώσσα των τυχαίων μεταβλητών. Ο ορισμός της μέσης τιμής ως ολοκλήρωμα εξηγεί το λόγο που βρήκαμε φράγματα για ολοκληρώματα στην Π2.

Έστω $h, h_1, h_2, \dots, h_\kappa: \mathbb{R} \longrightarrow \mathbb{R}$, κ+1 Borel συναρτήσεις.

Έστω επίσης

$$\boldsymbol{\mu} = (\mu_1, \mu_2, \dots, \mu_\kappa)^T \in \mathbb{R}^\kappa.$$

Υποθέτουμε τα εξής:

(Υ_1) Υπάρχει ε > 0 τέτοιο ώστε για κάθε $\nu_1, \nu_2, \dots, \nu_\kappa$ με

$$\mu_i - \varepsilon < \nu_i < \mu_i + \varepsilon, \quad i = 1, 2, \dots, \kappa,$$

να μπορώ να βρω μια τυχαία μεταβλητή Y ($Y = Y(\nu_1, \nu_2, \dots, \nu_\kappa)$) τέτοια ώστε

(1) $E|h(Y)| < \infty$,

(2) $E|h_i(Y)| < \infty, \quad i = 1, 2, \dots, \kappa$,

(3) $Eh_i(Y) = \nu_i, \quad i = 1, 2, \dots, \kappa$.

Ενδιαφερόμαστε για το $\sup_{X \in \mathcal{F}_{\boldsymbol{\mu}}} Eh(X)$,

όπου

$$\mathcal{F}_{\boldsymbol{\mu}} = \{\text{οι τ.μ. } X \text{ με } E|h(X)| < \infty, E|h_i(X)| < \infty \;\kappa\alpha\iota\; Eh_i(X) = \mu_i, \; i = 1, 2, \dots, \kappa\}.$$

Θέτουμε $A \subseteq \mathbb{R}^{\kappa+1}$ με

$$A := \left\{ \boldsymbol{\alpha} = \begin{pmatrix} \alpha_{\mathbf{o}} \\ \vdots \\ \alpha_{\boldsymbol{\kappa}} \end{pmatrix} : \; h(x) \leq \alpha_{\mathbf{o}} + \sum_{\mathbf{i}=\mathbf{1}}^{\boldsymbol{\kappa}} \alpha_{\mathbf{i}} h_{\mathbf{i}}(x), \; \forall x \in \mathbb{R} \right\}.$$

Το παρακάτω Θεώρημα είναι μία ανακατασκευή του Θεωρήματος Π1 (Isii 1969).

<u>Θεώρημα Π2</u> Υπό την υπόθεση (Υ_1) ισχύουν τα παρακάτω:

(α)
$$\sup_{X \in \mathcal{F}_{\boldsymbol{\mu}}} Eh(X) = \inf_{\boldsymbol{\alpha} \in A} \{\alpha_{\mathbf{o}} + \sum_{\mathbf{i}=\mathbf{1}}^{\boldsymbol{\kappa}} \alpha_{\mathbf{i}} \mu_{\mathbf{i}},\}.$$

Ειδικότερα, αν A = ∅ τότε $\sup_{X \in \mathcal{F}_{\boldsymbol{\mu}}} Eh(X) = +\infty$.

Δηλαδή

$$\sup_{X\in\mathcal{F}_\mu} Eh(X) < \infty \;\; \Leftrightarrow \; A \neq \emptyset.$$

(β) Αν Α $\neq \emptyset$ τότε υπάρχει $\boldsymbol{\alpha}^* = \begin{pmatrix} \alpha_o^* \\ \vdots \\ \alpha_\kappa^* \end{pmatrix} \in A$ τέτοιο ώστε

$$\sup_{X\in\mathcal{F}_\mu} Eh(X) = \alpha_o^* + \sum_{i=1}^{\kappa} \alpha_i^* \mu_i.$$

(γ) Αν Α $\neq \emptyset$ και για μια τυχαία μεταβλητή Χ ισχύει η σχέση

$$Eh(X) = \alpha_o^* + \sum_{i=1}^{\kappa} \alpha_i^* \mu_i$$

τότε

$$\mathbb{P}(h(X) = \alpha_o^* + \sum_{i=1}^{\kappa} \alpha_i^* h_i(x)) = 1.$$

Απόδειξη

Θα ακολουθήσουμε τη διαδικασία της απόδειξης του Θεωρήματος Π1, απλώς θα το διασκευάσουμε για τ.μ.

Θα πάρουμε πρώτα $f_0 \equiv 1$.

Η υπόθεση

$$\int_\Omega |f_0| d\mu = \mu(\Omega) < \infty$$

μας εξασφαλίζει ότι το $\mathfrak{B}$ περιέχει <u>ΜΟΝΟ </u>πεπερασμένα μέτρα.

Θα πάρουμε επίσης $M_0 = 1$. Τότε το $\mathfrak{B}_{\mathbf{M}}$ θα περιέχει μόνο μέτρα πιθανότητας.

Τώρα θα θεωρήσουμε μια μετρήσιμη συνάρτηση

$$X: \;(\Omega, \mathcal{A}) \longrightarrow \big(\mathbb{R}, \mathfrak{B}(\mathbb{R})\big)$$

και θα θέσουμε

$$g(\omega) = h\big(X(\omega)\big), \qquad \omega \in \Omega$$

$$f_0(\omega) \equiv 1, \qquad \omega \in \Omega$$

$$f_i(\omega) = h_i\big(X(\omega)\big), \qquad \omega \in \Omega, i = 1, 2, \ldots, \kappa,$$

M_0=1.

$$\mathfrak{B} = \left\{ \begin{array}{c} \mu: \\ (\Omega, \mathcal{A}, \mu) \text{ χώρος μέτρου}, \mu(\Omega) < \infty \\ \text{και} \int_{\Omega} \left|h\big(X(\omega)\big)\right| + \sum_{i=1}^{\kappa} \left|h_i\big(X(\omega)\big)\right| d\mu(\omega) < \infty \end{array} \right\}.$$

Ισοδύναμα, αφού $X : (\Omega, \mathcal{A}, \mu) \longrightarrow (\mathbb{R}, \mathfrak{B}(\mathbb{R}), \mu_X)$ μπορούμε να θεωρήσουμε ως

$$\mathfrak{B} = \{\lambda F(x), \lambda \geq 0, F \text{ συνάρτηση κατανομής με } E|h(X)| < \infty, E|h_i(X)| < \infty,\ i = 1, 2, \dots, \kappa\}.$$

Τώρα θα θέσουμε

$$\Gamma_{\kappa+1} := \left\{ \begin{array}{c} \lambda \cdot \begin{pmatrix} 1 \\ Eh_1(X) \\ \vdots \\ Eh_\kappa(X) \end{pmatrix}, \lambda \geq 0, X \text{ τ. μ. με} E|h(X)| < \infty, E|h_i(X)| < \infty, \\ i = 1, 2, \dots, \kappa \end{array} \right\}.$$

Υπόθεση: Το $\Gamma_{\kappa+1}$ περιέχει εσωτερικά σημεία.

Ερώτημα: Μπορεί να μην περιέχει;

Φυσικά, π.χ. αν κ = 1 και $h_1 \equiv 0$. Τότε $\Gamma_2 = \{(\lambda, 0), \lambda \geq 0\}$.

Ή π.χ. αν $h_1 \equiv 1$ οπότε $\Gamma_2 = \{(\lambda, \lambda), \lambda \geq 0\}$.

Φιξάρουμε λοιπόν ένα εσωτερικό σημείο του $\Gamma_{\kappa+1}$ της μορφής $\mathbf{M} = (1, M_1, \dots M_\kappa)$.

Αυτό σημαίνει δύο πράγματα:

1ον Υπάρχει τ.μ. X με $E|h(X)| < \infty, E|h_i(X)| < \infty, i = 1, 2, \dots, \kappa$

τέτοια ώστε

$$Eh_i(X) = M_i,\ i = 1, 2, \dots, \kappa.$$

2ον Υπάρχει $\delta > 0 : S(\mathbf{M}, \delta) \subseteq \Gamma_{\kappa+1}$.

Αναλυτικά: Μπορώ να βρω $\delta > 0$ ώστε αν $\|\boldsymbol{\beta} - \mathbf{M}\| < \delta$ τότε $\boldsymbol{\beta} \in \Gamma_{\kappa+1}$.

Δηλαδή:

αν

$$(\beta_o - 1)^2 + (\beta_1 - M_1)^2 + \dots + (\beta_\kappa - M_\kappa)^2 < \delta^2$$

τότε

$$\begin{pmatrix} \beta_o \\ \beta_1 \\ \vdots \\ \beta_\kappa \end{pmatrix} = \lambda \cdot \begin{pmatrix} 1 \\ Eh_1(X) \\ \vdots \\ Eh_\kappa(X) \end{pmatrix}$$

για κάποιο λ και για κάποια τ.μ. X με $E|h_i(X)| < \infty, i = 1, 2, \dots, \kappa$.

Όμως το λ δεν είναι «κάποιο λ», αλλά $\lambda = \beta_o$.

Άρα

$$\beta_1 = \beta_o Eh_1(X)$$

$$\vdots$$

$$\beta_\kappa = \beta_o Eh_\kappa(X)$$

όταν

β_o κοντά στο 1

β_1 κοντά στο M_1

$\vdots$

β_κ κοντά στο M_κ.

Δηλαδή το $\begin{pmatrix} Eh_1(X) \\ \vdots \\ Eh_\kappa(X) \end{pmatrix}$ παίρνει όλες τις τιμές $\begin{pmatrix} \beta_1/\beta_o \\ \vdots \\ \beta_\kappa/\beta_o \end{pmatrix}$ σε μια γειτονιά του $(M_1, \dots M_\kappa)$.

<u>Με άλλα λόγια:</u>

Όχι μόνο $Eh_i(X) = M_i$, αλλά ακόμα περισσότερο, υπάρχει $\delta > 0$ τέτοιο ώστε αν

$$\sum_{i=1}^{\kappa} \left(\widetilde{M}_i - M_i\right)^2 < \delta$$

τότε υπάρχει τ.μ. $\tilde{X}$ με

$$E\left|h(\tilde{X})\right| < \infty, \qquad E\left|h_i(\tilde{X})\right| < \infty, \qquad i = 1, 2, \dots, \kappa$$

και

$$Eh_i(\tilde{X}) = \widetilde{M}_i, \qquad i = 1, 2, \dots, \kappa.$$

Αυτές είναι ακριβώς οι υποθέσεις με τις οποίες καταλήγουμε στη σχέση (Π.4) της ενότητας Π2.

Παραδείγματα

1. Το supEX=0, όταν EX = 0, και $EX^2 = 1$.

Τώρα $M_1 = 0$, $M_2 = 1$ και $\mathrm{supEX} = \inf\{\alpha_0 + \alpha_1 \cdot 0 + \alpha_1 \cdot 1\} = \inf\{\alpha_0 + \alpha_2\}$.

$$A_3 = \{\alpha_0, \alpha_1, \alpha_2 : \ x \leq \alpha_0 + \alpha_1 x + \alpha_2 x^2 \text{ για κάθε } x\}$$

$$\alpha_2 x^2 + (\alpha_1 - 1)x + \alpha_0 \geq 0 \ \text{ για κάθε x.}$$

Αν $\alpha_2 = 0$, τότε $\alpha_1 = 1$ και $\alpha_0 \geq 0$.

Αν $\alpha_2 > 0$, τότε για να ισχύει η παραπάνω ανισότητα για κάθε x, θα πρέπει $(\alpha_1 - 1)^2 - 4\alpha_0\alpha_2 < 0 \Longrightarrow \alpha_1 \in \left(1 - 2\sqrt{\alpha_0\alpha_2}, 1 + 2\sqrt{\alpha_0\alpha_2}\right)$.

Συνοψίζοντας τις δύο παραπάνω περιπτώσεις, προκύπτει ότι το A_3 είναι το

$$A_3 = \left\{\alpha_0 \geq 0, \qquad \alpha_2 \geq 0, \qquad 1 - 2\sqrt{\alpha_0\alpha_2} \leq \alpha_1 \leq \ 1 + 2\sqrt{\alpha_0\alpha_2} \ \right\}$$

$$\inf_{A_3}\{\alpha_0 + \alpha_2\} = 0$$

Από το οποίο προκύπτει ότι το supEX = 0.

Τέλος,

$$\alpha_0^* = \alpha_2^* = 0, \ \alpha_1^* = 1, \ \mathbb{P}(X = x) = 1 \quad \forall x.$$

ΚΕΦΑΛΑΙΟ 5

ΕΥΡΕΣΗ ΑΝΩ ΦΡΑΓΜΑΤΩΝ ΓΙΑ ΤΟ ΕΥΡΟΣ n ΔΙΑΤΕΤΑΓΜΕΝΩΝ ΠΑΡΑΤΗΡΗΣΕΩΝ.

Στα προηγούμενα κεφάλαια ασχοληθήκαμε με την εύρεση ενός άνω φράγματος για τη μέγιστη διατεταγμένη παρατήρηση. Στο Κεφάλαιο 4 αναφέρθηκε η έννοια του εύρους δύο διατεταγμένων παρατηρήσεων, καθώς και η σχέση της μέσης τιμής του με τη μέση τιμή της μέγιστης παρατήρησης. Στο Κεφάλαιο αυτό θα προσπαθήσουμε να βρούμε άνω φράγματα για το εύρος n διατεταγμένων παρατηρήσεων.

5.1. Εισαγωγή

Στην ενότητα 5.1 θα αναφέρουμε κάποια φράγματα του range, που προκύπτουν άμεσα από τα φράγματα του Πίνακα στην τελευταία σελίδα του Κεφαλαίου 3. Πριν προχωρήσουμε σε αυτά θα επαναλάβουμε την περίπτωση που έχουμε τυχαίο δείγμα n παρατηρήσεων. Όπως είχαμε αναφέρει στην Ενότητα 1.2, και ειδικότερα στη σχέση (1.3), ένα άνω φράγμα για το εύρος (το οποίο επιτυγχάνεται)[17] δίνεται από την ανισότητα

$$E(R_n) \le \sigma \frac{n\sqrt{2}}{\sqrt{2n-1}}\left\{1-\frac{1}{\binom{2n-2}{n-1}}\right\}^{1/2}.$$

Από το Θεώρημα των Arnold και Groeneveld (1959) προκύπτει ότι για οποιεσδήποτε σταθερές $\lambda_1, \lambda_{2,\ldots,}\lambda_n$,

$$\left|\sum_{i=1}^{n}\lambda_i(E(X_{i:n})-\bar{\mu})\right| \le \left(\sum_{i=1}^{n}(\lambda_i-\bar{\lambda})^2\right)^{1/2}\left(\sum_{i=1}^{n}\left\{(\mu_i-\bar{\mu})^2+\sigma_i^{\,2}\right\}\right)^{1/2}$$

όπου $\bar{\mu}=\frac{1}{n}\sum_{i=1}^{n}\mu_i$ και $\bar{\lambda}=\frac{1}{n}\sum_{i=1}^{n}\lambda_i$.

Αν θεωρήσουμε $\lambda_i=0, i=2,3,\ldots,n-1$, $\lambda_1=-1$ και $\lambda_n=1$, τότε η παραπάνω ανισότητα γίνεται

$$E(X_{i:n})-E(X_{i:n}) \le \sqrt{2}\left(\sum_{i=1}^{n}\left\{(\mu_i-\bar{\mu})^2+\sigma_i^{\,2}\right\}\right)^{1/2}$$

δηλαδή

$$ER_n \le \sqrt{2}\left(\sum_{i=1}^{n}\left\{(\mu_i-\bar{\mu})^2+\sigma_i^{\,2}\right\}\right)^{1/2}.$$

Αντικαθιστώντας στη σχέση (2.15)

$$\left|\sum_{i=1}^{n}\lambda_i(E(X_{i:n})-\bar{\mu})\right| \le \left(\sum_{i=1}^{n}(\lambda_i-\bar{\lambda})^2\right)^{1/2}\left(\sum_{i=1}^{n}\left\{(\mu_i-\bar{\mu})^2+\sigma_i^{\,2}\right\}-n\mathrm{Var}\bar{X}\right)^{1/2}$$

---

[17] Ludwig (1959)

$\lambda_i = 0, i = 2, 3, \dots, n-1$, $\lambda_1 = -1$ και $\lambda_n = 1$, τότε προκύπτει

$$ER_n \le \sqrt{2}\left(\sum_{i=1}^{n}\left\{\left(\mu_i - \bar{\mu}\right)^2 + \sigma_i^{2}\right\} - n\mathrm{Var}\overline{X}\right)^{1/2}.$$

5.2. Ένα άνω φράγμα για το αναμενόμενο εύρος.

Έστω $\mathbf{X} = (X_1, \dots, X_n)$ ένα τυχαίο διάνυσμα με $EX_i = \mu_i$ και $\mathrm{Var}X_i = \sigma_i^2$, όπου $0 < \sigma_i^2 < \infty$ για όλα τα i = 1, 2, …,n. Αν $\boldsymbol{\mu} = (\mu_1, \dots, \mu_n)$, $\boldsymbol{\sigma}^2 = (\sigma_1^2, \dots, \sigma_n^2)$ και $\mathrm{Var}\mathbf{X} = \begin{bmatrix} \sigma_1^2 & \cdots & 0 \\ 0 & & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_n^2 \end{bmatrix}$. Δηλαδή ο διαγώνιος του πίνακα συδιακύμανσης **Σ.** Τα τ.δ. **X** που ικανοποιούν τις παραπάνω προϋποθέσεις, συνοψίζονται στην παρακάτω κλάση

$$\mathcal{F}_n(\boldsymbol{\mu}, \boldsymbol{\sigma}) := \{\mathbf{X}: E\mathbf{X} = \boldsymbol{\mu},\ \mathrm{Var}\mathbf{X} = \boldsymbol{\sigma}^2\}. \tag{5.1}$$

Για παράδειγμα, $X \in \mathcal{F}_1(\mu, \sigma)$ σημαίνει ότι $EX=\mu$ και $\mathrm{Var}X = \sigma^2$.

Έστω $X_{1:n} \le \cdots \le X_{n:n}$ το διατεταγμένο δείγμα του **X** και θέτουμε $R_n = X_{n:n} - X_{1:n}$ το εύρος (range) του δείγματος. Στα Κεφάλαια που ακολουθούν μας ενδιαφέρει να βρούμε βέλτιστο άνω φράγμα για το εύρος όταν το **X** βρίσκεται στην κλάση (5.1). Δηλαδή θέλουμε να βρούμε το $\sup_{\mathbf{X} \in \mathcal{F}_n(\boldsymbol{\mu}, \boldsymbol{\sigma})} ER_n$. Θα ξεκινήσουμε με μία ανισότητα για το εύρος, όπως φαίνεται στο παρακάτω Λήμμα

<u>Λήμμα 5.2.1.</u> Για κάθε $\mathbf{X} \in \mathbb{R}^n$, $c \in \mathbb{R}$ και $\lambda > 0$,

$$R_n \le -(n-2)\lambda + \frac{\lambda}{2}\sum_{i=1}^{n}\left\{\left|\frac{X_i - c}{\lambda} - 1\right| + \left|\frac{X_i - c}{\lambda} + 1\right|\right\}. \tag{5.2}$$

Η ισότητα στην (5.2) επιτυγχάνεται αν και μόνο αν,

$$X_{1:n} \le c - \lambda \le X_{2:n} \le \cdots \le X_{n-1:n} \le c + \lambda \le X_{n:n}. \tag{5.3}$$

Απόδειξη

Φιξάρουμε $c \in \mathbb{R}$ και $\lambda > 0$ και θεωρούμε $y_1 = c - \lambda$ και $y_2 = c + \lambda$. Παρατηρούμε ότι $y_1 < y_2$. Οπότε αρκεί να αποδείξουμε ότι

$$R_n \le \frac{1}{2}\sum_{i=1}^{n}\{|X_i - y_1| + |X_i - y_2|\} - \frac{n-2}{2}(y_2 - y_1) \tag{5.4}$$

Επειδή

$$R_n = X_{n:n} - X_{1:n}$$

και

$$\sum_{i=1}^{n}\{|X_i - y_1| + |X_i - y_2|\} = \sum_{i=1}^{n}\{|X_{i:n} - y_1| + |X_{i:n} - y_2|\}$$

έχουμε

$$\frac{1}{2}\sum_{i=1}^{n}\{|X_i - y_1| + |X_i - y_2|\} - \frac{n-2}{2}(y_2 - y_1) - R_n$$

$$= \tfrac{1}{2}\{|X_{1:n} - y_1| + |X_{1:n} - y_2|\} + \tfrac{1}{2}\{|X_{n:n} - y_1| + |X_{n:n} - y_2|\} - (X_{n:n} - X_{1:n})$$

$$+ \tfrac{1}{2}\textstyle\sum_{i=2}^{n-1}\{|X_i - y_1| + |X_i - y_2| - (y_2 - y_1)\}. \qquad (*)$$

Παρατηρούμε τώρα ότι για i = 1, 2,…, n – 1,

$$y_2 - y_1 = |y_2 - y_1| = |y_2 - X_{i:n} + X_{i:n} - y_1|$$

$$\leq |X_{i:n} - y_1| + |X_{i:n} - y_2|.$$

Οπότε αφού το άθροισμα

$$\textstyle\sum_{i=2}^{n-1}\{|X_i - y_1| + |X_i - y_2| - (y_2 - y_1)\}$$

προσθέτει μόνο μη αρνητικούς όρους, είναι μη αρνητικό.

Επίσης, οι όροι εκτός του παραπάνω αθροίσματος στη σχέση (*), είναι οι

$$\frac{1}{2}\{|X_{1:n} - y_1| + |X_{1:n} - y_2| - (X_{n:n} - X_{1:n})\}$$

$$+\frac{1}{2}\{|X_{n:n} - y_1| + |X_{n:n} - y_2| - (X_{n:n} - X_{1:n})\}$$

Όμως

$$X_{n:n} - X_{1:n} = |X_{n:n} - X_{1:n}| = |X_{n:n} - y_1 + y_1 - X_{1:n}|$$

$$\leq |X_{n:n} - y_1| + |X_{1:n} - y_1|.$$

Και ομοίως,

$$X_{n:n} - X_{1:n} \leq |X_{n:n} - y_2| + |X_{1:n} - y_2|.$$

Άρα και οι δύο αυτοί όροι είναι μη αρνητικοί, οπότε ισχύει η ανισότητα (5.4), και κατ’ επέκταση η (5.2).

Ας υποθέσουμε τώρα ότι ισχύει η ισότητα στην (5.4). Τότε θα πρέπει να ισχύουν όλες οι ισότητες (n το πλήθος)

$$X_{n:n} - X_{1:n} = |X_{n:n} - y_1| + |X_{1:n} - y_1| \tag{A}$$

$$X_{n:n} - X_{1:n} = |X_{n:n} - y_2| + |X_{1:n} - y_2| \tag{B}$$

και

$$y_2 - y_1 = |X_{i:n} - y_1| + |X_{i:n} - y_2|, i = 1, 2, \dots, n. \tag{Γ}$$

Όμως

$$|X_{n:n} - y_1| + |X_{1:n} - y_1| = \begin{cases} 2y_1 - X_{1:n} - X_{n:n}, & \text{αν } X_{n:n} < y_1, \\ X_{n:n} - X_{1:n}, & \text{αν } X_{1:n} \leq y_1 \leq X_{n:n}, \\ X_{1:n} + X_{n:n} - 2y_1, & \text{αν } X_{1:n} > y_1. \end{cases}$$

Στην πρώτη περίπτωση έχουμε $X_{n:n} < y_1$ οπότε

$$2y_1 - X_{1:n} - X_{n:n} > 2X_{n:n} - X_{1:n} - X_{n:n} = X_{n:n} - X_{1:n}.$$

Άρα δεν έχω ισότητα στην (A).

Στην τρίτη περίπτωση έχουμε $X_{1:n} > y_1$, οπότε $-2y_1 > -2X_{1:n}$ και

$$X_{1:n} + X_{n:n} - 2y_1 > X_{1:n} + X_{n:n} - 2X_{1:n} = X_{n:n} - X_{1:n}.$$

Οπότε και πάλι δεν έχουμε ισότητα στην (A).

Φυσικά, όταν $X_{1:n} \leq y_1 \leq X_{n:n}$ έχουμε προφανώς ισότητα στην (A).

Το συμπέρασμα είναι ότι

$$X_{n:n} - X_{1:n} = |X_{n:n} - y_1| + |X_{1:n} - y_1| \text{ αν και μόνο αν } X_{1:n} \leq y_1 \leq X_{n:n}.$$

Κάνοντας την ίδια ανάλυση στη (B) προκύπτει ότι

$$X_{n:n} - X_{1:n} = |X_{n:n} - y_2| + |X_{1:n} - y_2| \text{ αν και μόνο αν } X_{1:n} \leq y_2 \leq X_{n:n}.$$

Όμως θέλουμε ταυτόχρονα ισότητα στις (A) και (B). Αυτό μπορεί μόνο αν (αφού $y_1 < y_2$)

$$X_{1:n} \leq y_1 < y_2 \leq X_{n:n}. \tag{Δ}$$

Ας θεωρήσουμε τώρα τις ισότητες στη (Γ). Είναι:

$$|X_{i:n} - y_1| + |X_{i:n} - y_2| = \begin{cases} y_1 + y_2 - 2X_{i:n}, & \text{αν } X_{i:n} < y_1, \\ y_2 - y_1, & \text{αν } y_1 \le X_{i:n} \le y_2, \\ 2X_{i:n} - y_1 - y_2, & \text{αν } X_{i:n} > y_2. \end{cases}$$

Στην πρώτη περίπτωση που $X_{i:n} < y_1$ έχουμε $-2X_{i:n} > -2y_1$. Οπότε

$$y_1 + y_2 - 2X_{i:n} > y_1 + y_2 - 2y_1 = y_2 - y_1.$$

Δηλαδή δεν έχω ισότητα στη (Γ).

Ομοίως, στην Τρίτη περίπτωση που $X_{i:n} > y_2$ έχουμε

$$2X_{i:n} - y_1 - y_2 > 2y_2 - y_1 - y_2 = y_2 - y_1.$$

Οπότε και πάλι δεν έχω ισότητα στη (Γ).

Τελικά έχουμε ισότητα στη (Γ), αν και μόνο αν

$$y_1 \le X_{i:n} \le y_2, \quad i = 2, 3, \dots, n-1 \tag{Ε}$$

Από τις (Δ) και (Ε) έπεται ότι η ισότητα στην (5.4) επιτυγχάνεται αν και μόνο αν

$$X_{1:n} \le y_1 \le X_{2:n} \le \cdots \le X_{n-1:n} \le y_2 \le X_{n:n}.$$

Αντικαθιστώντας τελικά $y_1 = c - \lambda$ και $y_2 = c + \lambda$, προκύπτει η σχέση (5.3).

■

Το παραπάνω συμπέρασμα, δείχνει ότι χρειαζόμαστε δύο μεταβλητές, για το σωστό χειρισμό του R_n. Επίσης, από το προηγούμενο Λήμμα, θα προσπαθήσουμε να βρούμε το $\sup E\{|X-1| + |X+1|\}$, όταν το X είναι τ.μ. με δεδομένη μέση τιμή και διακύμανση.

<u>Λήμμα 5.2.2</u> Έστω $X \in \mathcal{F}(\mu, \sigma^2)$, $\mu \in \mathbb{R}$, $\sigma > 0$. Τότε

$$\sup E\{|X-1| + |X+1|\} = U(\mu, \sigma) \tag{5.5}$$

Όπου

$$U(\mu,\sigma) = \begin{cases} 2\sqrt{\mu^2 + \sigma^2}, & \text{αν } \mu^2 + \sigma^2 \ge 4, \\ 2 + \frac{1}{2}(\mu^2 + \sigma^2), & \text{αν } 2|\mu| \le \mu^2 + \sigma^2 < 4, \\ |\mu| + 1 + \sqrt{(|\mu| - 1)^2 + \sigma^2}, & \text{αν } \mu^2 + \sigma^2 \le 2|\mu| < 4. \end{cases} \tag{5.6}$$

(Όταν σ = 0, $U(\mu, 0) = |\mu - 1| + |\mu + 1| = 2\max\{|\mu|, 1\}$).

Η ισότητα στην

$$E\{|X-1| + |X+1|\} \le U(\mu, \sigma) \tag{5.7}$$

επιτυγχάνεται από μία μοναδική τυχαία μεταβλητή $X^* \in \mathcal{F}(\mu,\sigma^2)$ η οποία παίρνει τρεις συγκεκριμένες τιμές ως εξής:

(α) Αν $\mu^2+\sigma^2 \geq 4$ τότε η μόνη τ.μ. που επιτυγχάνει την ισότητα είναι η δίτιμη με

$$\mathbb{P}\left(X^* = -\sqrt{\mu^2+\sigma^2}\right) = \frac{1}{2}\left(1-\frac{\mu}{\sqrt{\mu^2+\sigma^2}}\right),$$

$$\mathbb{P}\left(X^* = \sqrt{\mu^2+\sigma^2}\right) = \frac{1}{2}\left(1+\frac{\mu}{\sqrt{\mu^2+\sigma^2}}\right).$$

(β) Αν $2|\mu| \leq \mu^2+\sigma^2 < 4$ τότε η μόνη τ.μ. X^* που επιτυγχάνει την ισότητα είναι η τρίτιμη με

$$\mathbb{P}(X^*=-2) = \frac{1}{8}(\mu^2+\sigma^2-2\mu),$$

$$\mathbb{P}(X^*=0) = \frac{1}{4}(4-\mu^2-\sigma^2),$$

$$\mathbb{P}(X^*=2) = \frac{1}{8}(\mu^2+\sigma^2+2\mu).$$

(γ) Αν $\mu^2+\sigma^2 \leq 2\mu < 4$ τότε η μόνη τ.μ. X^* που επιτυγχάνει την ισότητα είναι η δίτιμη

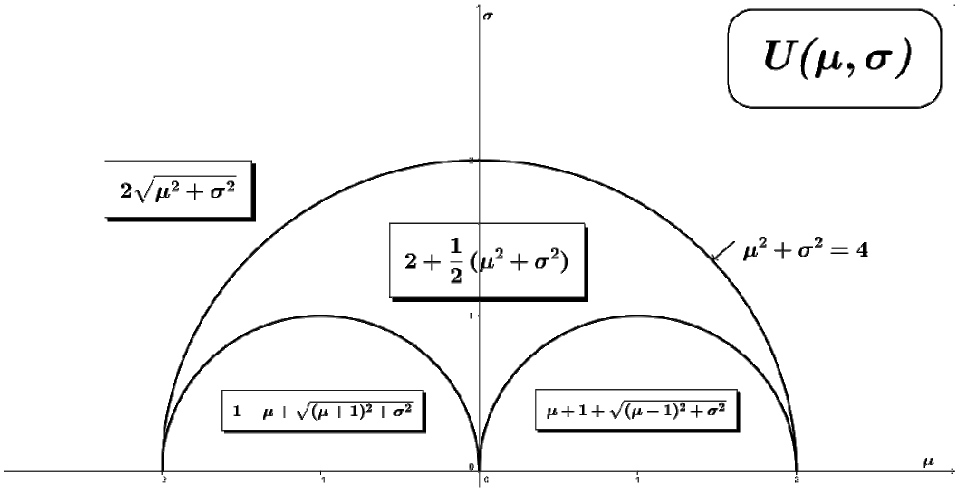
$$\mathbb{P}\left(X^* = 1-\sqrt{(\mu-1)^2+\sigma^2}\right) = \frac{1}{2}\left(1-\frac{\mu-1}{\sqrt{(\mu-1)^2+\sigma^2}}\right),$$

$$\mathbb{P}\left(X^* = 1+\sqrt{(\mu-1)^2+\sigma^2}\right) = \frac{1}{2}\left(1+\frac{\mu-1}{\sqrt{(\mu-1)^2+\sigma^2}}\right).$$

(δ) Αν $\mu^2+\sigma^2 \leq 2|\mu| < 4$ και $\mu < 0$ (δηλαδή $-4 < 2\mu \leq -\mu^2-\sigma^2$) τότε η μόνη τ.μ. X που πετυχαίνει την ισότητα είναι η δίτιμη με

$$\mathbb{P}\left(X^* = -1-\sqrt{(|\mu|-1)^2+\sigma^2}\right) = \frac{1}{2}\left(1+\frac{|\mu|-1}{\sqrt{(|\mu|-1)^2+\sigma^2}}\right),$$

$$\mathbb{P}\left(X^* = -1+\sqrt{(|\mu|-1)^2+\sigma^2}\right) = \frac{1}{2}\left(1-\frac{|\mu|-1}{\sqrt{(|\mu|-1)^2+\sigma^2}}\right).$$



Σχ.5.1 *Το παραπάνω σχήμα, παρουσιάζει τα χωρία στα οποία η U(μ, σ) αλλάζει τύπο.*

Απόδειξη

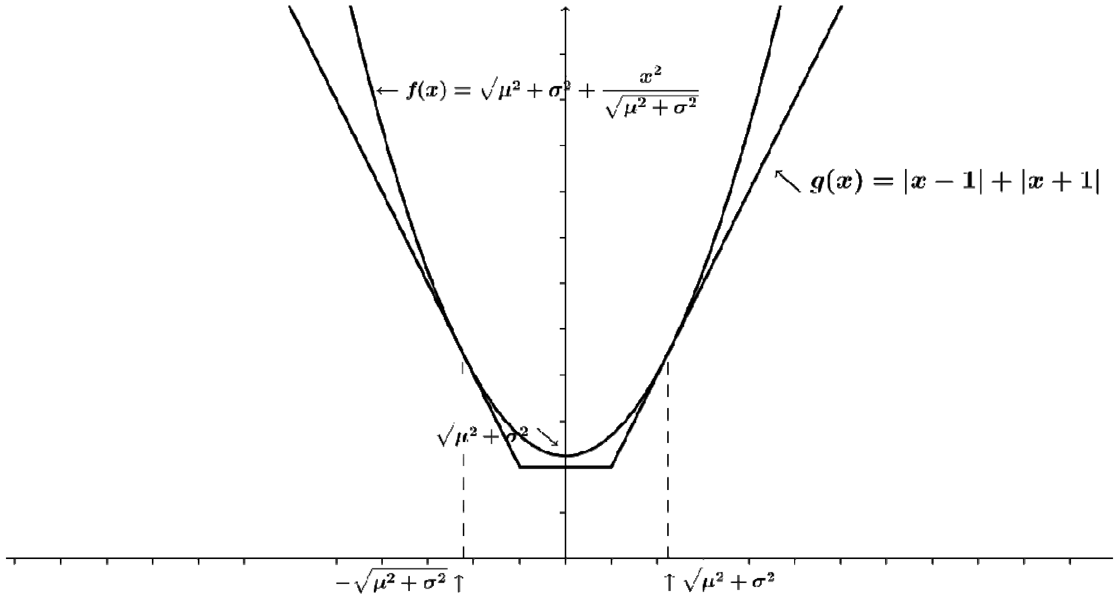
Θέτουμε $g(x) = |x-1| + |x+1| = U(x, 0)$.

(α) Έστω $\mu^2 + \sigma^2 \geq 4$.

Προφανώς

$$g(x) \leq \sqrt{\mu^2+\sigma^2} + \frac{x^2}{\sqrt{\mu^2+\sigma^2}}, \quad x \in \mathbb{R}, \qquad (*)$$

με ισότητα αν και μόνο αν $x \in \left\{-\sqrt{\mu^2+\sigma^2},\ \sqrt{\mu^2+\sigma^2},\ \right\}$.



Σχ.5.2

Η γραφική παράσταση των συναρτήσεων g(x) και $f(x) = \sqrt{\mu^2+\sigma^2} + \frac{x^2}{\sqrt{\mu^2+\sigma^2}}$. *Στο σχήμα φαίνεται ξεκάθαρα ότι η g με την f εφάπτονται στα σημεία με τετμημένες* $-\sqrt{\mu^2+\sigma^2}$ *και* $\sqrt{\mu^2+\sigma^2}$.

Επομένως,

$$\mathrm{Eg(X)} \leq \mathrm{E}\left\{\sqrt{\mu^2+\sigma^2}+\frac{\mathrm{X}^2}{\sqrt{\mu^2+\sigma^2}}\right\} = 2\sqrt{\mu^2+\sigma^2},$$

αφού $\mathrm{EX}^2 = \mu^2+\sigma^2$.

Για ισότητα, πρέπει και αρκεί

$$\mathrm{E}\left\{\sqrt{\mu^2+\sigma^2}+\frac{\mathrm{X}^2}{\sqrt{\mu^2+\sigma^2}}-\mathrm{g(X)}\right\} = 0.$$

Η $\mathrm{h(x)} = \sqrt{\mu^2+\sigma^2}+\frac{\mathrm{x}^2}{\sqrt{\mu^2+\sigma^2}}-\mathrm{g(x)}$ είναι μη αρνητική, οπότε $\mathrm{Eh(x)} = 0$ αν και μόνο αν $\mathbb{P}(\mathrm{h(x)} = 0) = 1$.

Όμως,

$$\{\mathrm{h(x)} = 0\} = \left\{\mathrm{X} = -\sqrt{\mu^2+\sigma^2}\right\} \cup \left\{\mathrm{X} = \sqrt{\mu^2+\sigma^2}\right\}.$$

Άρα για να έχουμε ισότητα πρέπει και αρκεί

$$\mathbb{P}\left(\mathrm{X} = -\sqrt{\mu^2+\sigma^2}\right) = 1-\mathrm{p},$$

$$\mathbb{P}\left(\mathrm{X} = \sqrt{\mu^2+\sigma^2}\right) = \mathrm{p},$$

για κάποιο $\mathrm{p} \in [0,1]$.

Τότε $\mathrm{EX}^2 = \mu^2+\sigma^2$. Επειδή $\mathrm{X} \in \mathcal{F}(\mu,\sigma^2)$, ισχύει ότι ΕΧ=μ. Η σχέση αυτή, θα μας δώσει την τιμή του p.

$$\mathrm{EX} = \mu \Leftrightarrow -(1-\mathrm{p})\sqrt{\mu^2+\sigma^2}+\mathrm{p}\sqrt{\mu^2+\sigma^2} = \mu$$

$$\Leftrightarrow 2\mathrm{p}-1 = \frac{\mu}{\sqrt{\mu^2+\sigma^2}}$$

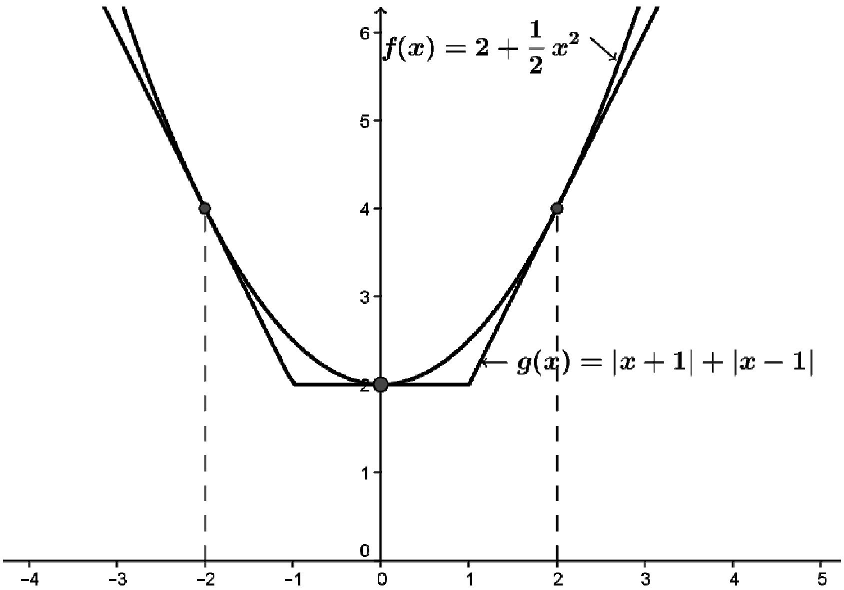
$$\Leftrightarrow \mathrm{p} = \frac{1}{2}\left(1+\frac{\mu}{\sqrt{\mu^2+\sigma^2}}\right).$$

(β) Έστω $2|\mu| \leq \mu^2+\sigma^2 < 4$.

Τώρα χρησιμοποιούμε την ανισότητα

$$g(x) \leq 2 + \frac{1}{2}x^2 \qquad (**)$$

με ισότητα αν και μόνο αν $x \in \{-2,\ 0,\ 2\}$.



Σχ. 5.3

Γραφική παράσταση των συναρτήσεων g(x) και f(x) = $2 + \frac{1}{2}x^2$*. Φαίνεται ξεκάθαρα, ότι ισχύει η ανισότητα* $(**)$*. Η ισότητα ισχύει όταν x = - 2, ή x = 0, ή x = 2.*

Προφανώς, $Eg(X) \leq E(2 + \frac{1}{2}X^2) = 2 + \frac{1}{2}(\mu^2 + \sigma^2)$ αφού $EX^2 = \mu^2 + \sigma^2$. Για ισότητα πρέπει και αρκεί

$$E\left\{2 + \frac{1}{2}X^2 - g(X)\right\} = 0,$$

δηλαδή,

$$Eh(X) = 0 \ \text{ όπου } h(x) = 2 + \tfrac{1}{2}x^2 - g(x) \geq 0.$$

Άρα, πρέπει και αρκεί $\mathbb{P}(h(X) = 0) = 1$, ή ισοδύναμα, $\mathbb{P}(X \in \{-2, 0, 2\}) = 1$.

Έστω $p_1 \in [0,1]$, $p_2 \in [0,1]$ με $p_1 + p_2 \leq 1$. Θέτουμε

$$X = \begin{cases} -2, & \text{με πιθανότητα } p_1, \\ 0, & \text{με πιθανότητα } 1 - p_1 - p_2, \\ 2, & \text{με πιθανότητα } p_2. \end{cases}$$

Η σχέση EX = μ δίνει την εξίσωση

$$p_2 - p_1 = \frac{\mu}{2}. \tag{5.8}$$

Η σχέση $EX^2 = \mu^2 + \sigma^2$ δίνει την εξίσωση

$$p_1 + p_2 = \frac{\mu^2+\sigma^2}{4}. \tag{5.9}$$

Προσθέτοντας και αφαιρώντας τις (5.8) και (5.9) βρίσκουμε

$$2p_1 = \frac{\mu}{2} - \frac{\mu^2+\sigma^2}{4},$$

$$2p_2 = \frac{\mu}{2} + \frac{\mu^2+\sigma^2}{4}.$$

Τελικά

$$p_1 = \mathbb{P}(X = -2) = \frac{1}{8}(\mu^2 + \sigma^2 - 2\mu),$$

$$p_2 = \mathbb{P}(X = 2) = \frac{1}{8}(\mu^2 + \sigma^2 + 2\mu),$$

και

$$1 - p_1 - p_2 = \mathbb{P}(X = 0) = 1 - \frac{1}{4}(\mu^2 + \sigma^2).$$

Τα p_1, p_2, και $1 - p_1 - p_2$ είναι όλα θετικά επειδή $2|\mu| < \mu^2 + \sigma^2 < 4$. Οπότε

$$\mu^2 + \sigma^2 - 2\mu \geq \mu^2 + \sigma^2 - 2|\mu| > 0, \text{ δηλαδή } p_1 > 0,$$

$$\mu^2 + \sigma^2 + 2\mu \geq \mu^2 + \sigma^2 - 2|\mu| > 0, \text{ δηλαδή } p_2 > 0, \text{ και}$$

$$\mu^2 + \sigma^2 < 4 \Longrightarrow 1 - \frac{1}{4}(\mu^2 + \sigma^2) > 0, \ \delta\eta\lambda\alpha\delta\acute{\eta} \ 1 - p_1 - p_2 > 0.$$

Άρα η X που πετυχαίνει η ισότητα είναι γνήσια τρίτιμη.

(γ) Έστω $\mu^2 + \sigma^2 \leq 2\mu < 4$ (άρα $\mu \in (0, 2)$). Θεωρούμε την ανισότητα

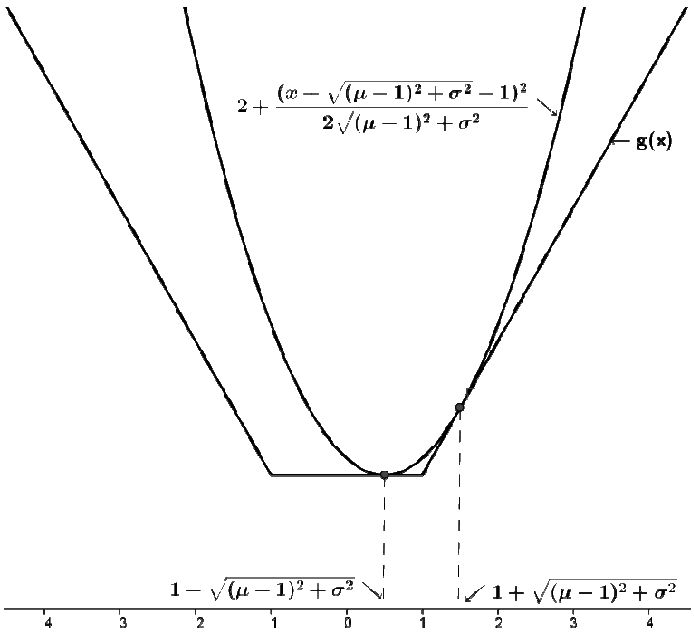
$$g(x) \le 2 + \frac{1}{2\sqrt{(\mu-1)^2+\sigma^2}}\left(x + \sqrt{(\mu-1)^2+\sigma^2} - 1\right)^2 \qquad (***)$$

με ισότητα αν και μόνο αν $x \in \left\{1 - \sqrt{(\mu-1)^2+\sigma^2},\ 1 + \sqrt{(\mu-1)^2+\sigma^2}\right\}$.

Η σχέση $\mu^2 + \sigma^2 \le 2\mu < 4$ δείχνει ότι

$$0 < (\mu-1)^2 + \sigma^2 = \mu^2 + \sigma^2 - 2\mu + 1 = 1 - \left(2\mu - (\mu^2+\sigma^2)\right) \le 1.$$

Άρα τα σημεία $x_1 = 1 - \sqrt{(\mu-1)^2+\sigma^2}$ και $x_1 = 1 + \sqrt{(\mu-1)^2+\sigma^2}$ βρίσκονται στα διαστήματα $[0,1)$ και $(1,2]$ αντίστοιχα.



Σχ. 5.4

Γραφική παράσταση των συναρτήσεων g(x) και f(x) $= 2 + \frac{1}{2\sqrt{(\mu-1)^2+\sigma^2}}\left(x + \sqrt{(\mu-1)^2+\sigma^2} - 1\right)^2$.
Παρατηρούμε ότι ισχύει η ανισότητα $(***)$, *και οι f, g εφάπτονται στα σημεία με τετμημένες*

$$1 - \sqrt{(\mu-1)^2+\sigma^2} \text{ και } 1 + \sqrt{(\mu-1)^2+\sigma^2}.$$

Άρα η ισότητα ισχύει για τα $x \in \left\{1 - \sqrt{(\mu-1)^2+\sigma^2},\ 1 + \sqrt{(\mu-1)^2+\sigma^2}\right\}$.

Μόνο στην περίπτωση που $\mu^2 + \sigma^2 = 2\mu$, δηλαδή

$$(\mu-1)^2 + \sigma^2 = \mu^2 + \sigma^2 - 2\mu + 1 = 1,$$

τα σημεία x_1 και x_2 είναι τα 0, 2 αντίστοιχα και η συνάρτηση γίνεται

$$2+\frac{1}{2\sqrt{(\mu-1)^2+\sigma^2}}\left(x+\sqrt{(\mu-1)^2+\sigma^2}-1\right)^2=2+\frac{1}{2}x^2, \quad \text{όπως στο (β).}$$

Τότε τα σημεία ισότητας είναι τρία, τα $-2, 0, 2$, αλλά το -2 έχει πιθανότητα

$$\mathbb{P}(X=-2)=\frac{1}{8}(\mu^2+\sigma^2-2\mu)=0, \text{ (βλ.(β)).}$$

Γενικά, δουλεύοντας με τον ίδιο τρόπο έχουμε

$$\begin{aligned}
Eg(X) &\leq 2+\frac{1}{2\sqrt{(\mu-1)^2+\sigma^2}}E\left(X+\sqrt{(\mu-1)^2+\sigma^2}-1\right)^2 \\
&= 2+\frac{1}{2\sqrt{(\mu-1)^2+\sigma^2}}\Big\{EX^2+2\left(\sqrt{(\mu-1)^2+\sigma^2}-1\right)\cdot EX+(\mu-1)^2+\sigma^2+1 \\
&\qquad -2\sqrt{(\mu-1)^2+\sigma^2}\Big\} \\
&= 2+\frac{\mu^2+\sigma^2+2\mu\sqrt{(\mu-1)^2+\sigma^2}-2\mu+(\mu-1)^2+\sigma^2+1-2\sqrt{(\mu-1)^2+\sigma^2}}{2\sqrt{(\mu-1)^2+\sigma^2}} \\
&= 2+(\mu-1)+\sqrt{(\mu-1)^2+\sigma^2} \\
&= \mu+1+\sqrt{(\mu-1)^2+\sigma^2}.
\end{aligned}$$

Για ισότητα πρέπει

$$E\left\{2+\frac{\left(X+\sqrt{(\mu-1)^2+\sigma^2}-1\right)^2}{2\sqrt{(\mu-1)^2+\sigma^2}}-g(X)\right\}=0.$$

Τότε έπεται ότι υπάρχει $p\in[0,1]$ ώστε

$$\mathbb{P}\left(X=1+\sqrt{(\mu-1)^2+\sigma^2}\right)=p,$$

και

$$\mathbb{P}\left(X=1-\sqrt{(\mu-1)^2+\sigma^2}\right)=1-p.$$

Η σχέση $EX=\mu$, δίνει

$$\begin{aligned}
\mu &= (1-p)\left[1-\sqrt{(\mu-1)^2+\sigma^2}\right]+p\left[1+\sqrt{(\mu-1)^2+\sigma^2}\right] \\
&= 1+(2p-1)\sqrt{(\mu-1)^2+\sigma^2}
\end{aligned}$$

$$\Longrightarrow 2p - 1 = \frac{\mu-1}{\sqrt{(\mu-1)^2+\sigma^2}}$$

$$\Longrightarrow \qquad p = \frac{1}{2}\left(1 + \frac{\mu-1}{\sqrt{(\mu-1)^2+\sigma^2}}\right), \ p \in (0,1).$$

Τότε

$$\begin{aligned}
EX^2 &= (1-p)\left[1 - \sqrt{(\mu-1)^2+\sigma^2}\right]^2 + p\left[1 + \sqrt{(\mu-1)^2+\sigma^2}\right]^2 \\
&= (1-p)\left\{1 - 2\sqrt{(\mu-1)^2+\sigma^2} + (\mu-1)^2 + \sigma^2\right\} \\
&\qquad + p\left\{1 + 2\sqrt{(\mu-1)^2+\sigma^2} + (\mu-1)^2 + \sigma^2\right\} \\
&= 1 + (\mu-1)^2 + \sigma^2 + 2(2p-1)\sqrt{(\mu-1)^2+\sigma^2} \\
&= 1 + (\mu-1)^2 + \sigma^2 + 2\cdot\frac{\mu-1}{\sqrt{(\mu-1)^2+\sigma^2}}\sqrt{(\mu-1)^2+\sigma^2} \\
&= 1 + (\mu-1)^2 + \sigma^2 + 2(\mu-1) \\
&= \big(1 + (\mu-1)\big)^2 + \sigma^2 \\
&= \mu^2 + \sigma^2.
\end{aligned}$$

Άρα $X \in \mathcal{F}(\mu, \sigma^2)$.

(δ) Στην περίπτωση $\mu^2 + \sigma^2 \le 2|\mu| < 4$, με $\mu < 0$ (άρα $\mu \in (-2, 0)$) δε χρειάζεται να επαναλάβουμε όλα τα βήματα. Τώρα θεωρούμε τη $-X \in \mathcal{F}(-\mu, \sigma^2)$. Όμως

$$g(x) = g(-x) = |x+1| + |x-1|.$$

Άρα $E(-X) = -\mu = |\mu|$, $\mathrm{Var}(-x) = \sigma^2$, και $(-\mu)^2 + \sigma^2 \le 2(-\mu) < 4$.

Από το (γ) έχουμε

$$\begin{aligned}
E(g(X)) = E\big(g(-X)\big) \le U(-\mu,\sigma) &= (-\mu) + 1 + \sqrt{\big((-\mu)-1\big)^2 + \sigma^2} \\
&= |\mu| + 1 + \sqrt{(|\mu|-1)^2 + \sigma^2}.
\end{aligned}$$

Για ισότητα πρέπει και αρκεί

$$\mathbb{P}\left(-X = 1 - \sqrt{((-\mu)-1)^2+\sigma^2}\right) = \frac{1}{2}\left(1 - \frac{(-\mu)-1}{\sqrt{((-\mu)-1)^2+\sigma^2}}\right),$$

$$\mathbb{P}\left(-X = 1 + \sqrt{\big((-\mu)-1\big)^2+\sigma^2}\right) = \frac{1}{2}\left(1 + \frac{(-\mu)-1}{\sqrt{((-\mu)-1)^2+\sigma^2}}\right).$$

Ισοδύναμα

$$\mathbb{P}\left(X = -1 + \sqrt{((|\mu| - 1)^2 + \sigma^2}\right) = \frac{1}{2}\left(1 - \frac{|\mu|-1}{\sqrt{(|\mu|-1)^2+\sigma^2}}\right),$$

$$\mathbb{P}\left(X = -1 - \sqrt{((|\mu| - 1)^2 + \sigma^2}\right) = \frac{1}{2}\left(1 + \frac{|\mu|-1}{\sqrt{(|\mu|-1)^2+\sigma^2}}\right).$$

■

Παρατήρηση 5.2.3 Ο Isii (1963) παρουσίασε γενικά αποτελέσματα τα οποία περιλαμβάνουν ανισότητες της μορφής του Λήμματος 5.2. Όπως αναφερθήκαμε στο Παράρτημα, το Θεώρημα του Isii στη «γλώσσα» των πιθανοτήτων θα μπορούσε να γραφτεί $\left(\text{π.χ. για } \boldsymbol{\alpha} = (\alpha_0, \alpha_1, \alpha_2)\right)$ ως εξής

Αν $h: \mathbb{R} \longrightarrow \mathbb{R}$ μία Borel συνάρτηση, $\mu \in \mathbb{R}$ και $\sigma > 0$ τότε

$$\sup_{X \in \mathcal{F}_1(\mu,\sigma)} Eh(X)$$

$$= \inf_{\alpha_0,\alpha_1,\alpha_2}\{\alpha_0 + \alpha_1\mu + \alpha_2(\mu^2 + \sigma^2) : \alpha_0 + \alpha_1 x + \alpha_2 x^2 \geq h(x) \text{ για όλα τα } x \in \mathbb{R}\}.$$

Ο Isii έδειξε ότι το παραπάνω infimum επιτυγχάνεται από κάποια $\boldsymbol{\alpha}^* = (\alpha_0^*, \alpha_1^*, \alpha_2^*) \in A$, όπου

$$A = \{(\alpha_0, \alpha_1, \alpha_2) : \alpha_0 + \alpha_1 x + \alpha_2 x^2 \geq h(x) \text{ για όλα τα } x \in \mathbb{R}\} \subseteq \mathbb{R}^3,$$

υπό την προϋπόθεση ότι το infimum είναι πεπερασμένο. Ωστόσο, δεν είναι εύκολο να καθορίσουμε το υποσύνολο A, και τα σημεία $\boldsymbol{\alpha}^*$ που επιτυγχάνεται το infimum. Το προηγούμενο Λήμμα δείχνει ότι αυτό είναι εφικτό για τη συνάρτηση

$$h(x) = |x - 1| + |x + 1|$$

και, το πιο σημαντικό είναι, ότι χαρακτηρίζει την περίπτωση της ισότητας. Τώρα είναι εύκολο να δείξουμε το παρακάτω Πόρισμα.

Πόρισμα 5.2.4 Έστω X μια τυχαία μεταβλητή με μέση τιμή μ και διασπορά σ^2, $0 < \sigma < \infty$. Θεωρούμε $c \in \mathbb{R}$ και $\lambda > 0$. Τότε

$$E\{|X - (c - \lambda)| + |X - (c + \lambda)|\} \leq \lambda U\left(\frac{\mu - c}{\lambda}, \frac{\sigma}{\lambda}\right), \tag{5.10}$$

όπου U(. , .) η συνάρτηση που δίνεται στην (5.6). Η ισότητα στην (5.10) επιτυγχάνεται από μία μοναδική τυχαία μεταβλητή $X \in \mathcal{F}_1(\mu, \sigma)$. Συγκεκριμένα,

(α) Αν $(\mu - c)^2 + \sigma^2 \geq 4\lambda^2$,

$$E\{|X - (c - \lambda)| + |X - (c + \lambda)|\} \leq 2\sqrt{(\mu - c)^2 + \sigma^2} \tag{5.11}$$

Η ισότητα στην (5.11) ισχύει αν και μόνο αν

$$\mathbb{P}(X = \alpha^-) = P^-, \quad \mathbb{P}(X = \alpha^+) = P^+, \tag{5.12}$$

όπου

$$\alpha^- = c - \sqrt{(\mu-c)^2+\sigma^2} \le c - 2\lambda,$$

$$\alpha^+ = c + \sqrt{(\mu-c)^2+\sigma^2} \ge c + 2\lambda,$$

$$P^- = \frac{1}{2}\left(1 - \frac{\mu-c}{\sqrt{(\mu-c)^2+\sigma^2}}\right) > 0,$$

$$P^+ = \frac{1}{2}\left(1 + \frac{\mu-c}{\sqrt{(\mu-c)^2+\sigma^2}}\right) > 0.$$

(β) Αν $2\lambda|\mu-c| < (\mu-c)^2+\sigma^2 < 4\lambda^2$,

$$E\{|X-(c-\lambda)|+|X-(c+\lambda)|\} \le 2\lambda + \frac{1}{2\lambda}[(\mu-c)^2+\sigma^2] \quad (5.13)$$

και η ισότητα στην (5.13) ισχύει αν και μόνο αν

$$\mathbb{P}(X=\beta^-) = p^-,\ \ \mathbb{P}(X=\beta^0) = p^0,\ \ \mathbb{P}(X=\beta^+) = p^+, \quad (5.14)$$

όπου

$$\beta^- = c - 2\lambda, \quad \beta^0 = c, \quad \beta^+ = c + 2\lambda,$$

$$p^- = \frac{1}{8\lambda^2}[(\mu-c)^2+\sigma^2-2\lambda(\mu-c)] > 0,$$

$$p^+ = \frac{1}{8\lambda^2}[(\mu-c)^2+\sigma^2+2\lambda(\mu-c)] > 0,$$

$$p^0 = 1 - \frac{1}{4\lambda^2}[(\mu-c)^2+\sigma^2] > 0.$$

(γ) Αν $(\mu-c)^2+\sigma^2 \le 2\lambda(\mu-c) < 4\lambda^2$,

$$E\{|X-(c-\lambda)|+|X-(c+\lambda)|\} \le \mu - c + \lambda + \sqrt{(\mu-c-\lambda)^2+\sigma^2} \quad (5.15)$$

και η ισότητα στην (5.15) ισχύει αν και μόνο αν

$$\mathbb{P}(X=\gamma^0) = q^0,\ \ \mathbb{P}(X=\gamma^+) = q^+, \quad (5.16)$$

όπου

$$\gamma^0 = c + \lambda - \sqrt{(\mu-c-\lambda)^2+\sigma^2},\ \ \gamma^+ = c + \lambda + \sqrt{(\mu-c-\lambda)^2+\sigma^2}$$

(με $c \le \gamma^0 < c+\lambda < \gamma^+ \le c + 2\lambda$)

$$q^0 = \frac{1}{2}\left(1 - \frac{\mu-c-\lambda}{\sqrt{(\mu-c-\lambda)^2+\sigma^2}}\right) > 0,\ \ q^+ = \frac{1}{2}\left(1 + \frac{\mu-c-\lambda}{\sqrt{(\mu-c-\lambda)^2+\sigma^2}}\right) > 0.$$

(δ) Αν $(\mu-c)^2+\sigma^2 \le -2\lambda(\mu-c) < 4\lambda^2$,

$$E\{|X-(c-\lambda)|+|X-(c+\lambda)|\} \le c - \mu + \lambda + \sqrt{(c-\mu-\lambda)^2+\sigma^2} \quad (5.17)$$

και η ισότητα στην (5.17) ισχύει αν και μόνο αν

$$\mathbb{P}(X = \delta^-) = r^-,\ \mathbb{P}(X = \delta^0) = r^0, \tag{5.18}$$

όπου

$$\delta^- = c - \lambda - \sqrt{(c - \mu - \lambda)^2 + \sigma^2},\ \delta^0 = c - \lambda + \sqrt{(c - \mu - \lambda)^2 + \sigma^2}$$

(με $c - 2\lambda \leq \delta^- < c - \lambda < \delta^0 \leq c$),

$$r^- = \frac{1}{2}\left(1 + \frac{c-\mu-\lambda}{\sqrt{(c-\mu-\lambda)^2+\sigma^2}}\right) > 0,\ r^0 = \frac{1}{2}\left(1 - \frac{c-\mu-\lambda}{\sqrt{(c-\mu-\lambda)^2+\sigma^2}}\right) > 0.$$

Απόδειξη

Έχουμε $|X - (c - \lambda)| + |X - (c + \lambda)| = \lambda\left\{\left|\frac{X-c}{\lambda} - 1\right| + \left|\frac{X-c}{\lambda} + 1\right|\right\}$. Επομένως, αφού $Y = \frac{X-c}{\lambda} \in \mathcal{F}_{\mathbf{1}}\left(\frac{\mu-c}{\lambda}, \frac{\sigma}{\lambda}\right)$, το Λήμμα 5.2.2 οδηγεί στη σχέση (5.10) ως εξής:

$$E\{|X - (c - \lambda)| + |X - (c + \lambda)|\} = \lambda E\{|Y - 1| + |Y + 1|\} \leq \lambda U\left(\frac{\mu-c}{\lambda}, \frac{\sigma}{\lambda}\right).$$

Αφού $\lambda > 0$, η ισότητα επιτυγχάνεται από μία μοναδική τυχαία μεταβλητή $Y^* = \frac{X-c}{\lambda} \in \mathcal{F}_{\mathbf{1}}\left(\frac{\mu-c}{\lambda}, \frac{\sigma}{\lambda}\right)$. Το οποίο σημαίνει ότι η $X^* = c + \lambda Y^*$ είναι η μοναδική τυχαία μεταβλητή που ικανοποιεί την ισότητα στην (5.10). Αν αντικαταστήσουμε τη σ.π.π. της Y^* στις τέσσερεις περιπτώσεις του Λήμματος 5.2.2, προκύπτουν οι σχέσεις (5.11) – (5.18).

■

Θεώρημα 5.2.5 Αν $EX_i = \mu_i$ και $\mathrm{Var}(X_i) = \sigma_i^2 > 0, i = 1, 2, \ldots, n$ τότε

$$ER_n \leq \inf_{\substack{\lambda>0 \\ c\in\mathbb{R}}}\left\{-(n - 2)\lambda + \frac{1}{2}\lambda \sum_{i=1}^{n} U\left(\frac{\mu_i-c}{\lambda}, \frac{\sigma_i^2}{\lambda}\right)\right\} \tag{5.19}$$

όπου η $U(\mu, \sigma)$, $\mu \in \mathbb{R}$, $\sigma > 0$, δίδεται από το Λήμμα 5.2.1.

Ας ονομάσουμε $\varphi_n(c, \lambda) = \varphi_n(c, \lambda;\ \boldsymbol{\mu}, \boldsymbol{\sigma})$ με

$$\varphi_n(c, \lambda) = -(n - 2) + \frac{1}{2}\lambda \sum_{i=1}^{n} U\left(\frac{\mu_i-c}{\lambda}, \frac{\sigma_i^2}{\lambda}\right). \tag{5.20}$$

Τότε το φράγμα είναι

$$ER_n \leq \inf_{\substack{\lambda>0 \\ c\in\mathbb{R}}}\varphi_n(c, \lambda).$$

Απόδειξη

Το Λήμμα 5.2.1 μας δίνει την ανισότητα

$$R_n \leq -(n - 2)\lambda + \frac{\lambda}{2}\sum_{i=1}^{n}\left\{\left|\frac{X_i-c}{\lambda} - 1\right| + \left|\frac{X_i-c}{\lambda} + 1\right|\right\},$$

Από την οποία προκύπτει η ανισότητα

$$ER_n \leq -(n-2)\lambda + \tfrac{\lambda}{2}\textstyle\sum_{i=1}^{n} E\left\{\left|\frac{X_i-c}{\lambda}-1\right| + \left|\frac{X_i-c}{\lambda}+1\right|\right\}.$$

Όμως από το Λήμμα 5.2, $E\left\{\left|\frac{X_i-c}{\lambda}-1\right| + \left|\frac{X_i-c}{\lambda}+1\right|\right\} \leq U\left(\frac{\mu_i-c}{\lambda}, \frac{\sigma_i}{\lambda}\right), i = 1, 2, \ldots, n.$ Οπότε

$$ER_n \leq -(n-2)\lambda + \frac{\lambda}{2}\sum_{i=1}^{n} U\left(\frac{\mu_i - c}{\lambda}, \frac{\sigma_i}{\lambda}\right) = \varphi_n(c, \lambda).$$

Αφού για όλα τα $\lambda > 0$, $c \in \mathbb{R}$ έχω ανισότητα, έπεται ότι

$$ER_n \leq \inf_{\substack{\lambda>0\\ c\in\mathbb{R}}} \varphi_n(c, \lambda).$$

■

<u>Παρατήρηση 5.2.6</u> Σε αυτή τη φάση δεν είναι προφανές ότι το άνω φράγμα της (5.19) είναι βέλτιστο[18], και δεν είναι αυτονόητο πως θα βρούμε $c = c_0$ και $\lambda = \lambda_0$ (αν υπάρ-χουν) ώστε να πραγματοποιείται το infimum στο δεξί μέλος της (5.19). Ωστόσο, όλες οι τιμές $\varphi_n(c, \lambda)$ αποτελούν άνω φράγμα της ER_n. Δηλαδή αν αντικαταστήσουμε κάποια (κατάλληλα) c και λ στη συνάρτηση (5.20), θα δημιουργήσουμε ένα άνω φράγμα για την ER_n. Για παράδειγμα, μία επιλογή που μπορούμε να κάνουμε για το c και το λ, είναι

$$c = \bar{\mu} \text{ και } \lambda = \tfrac{1}{4}AG_n = \tfrac{1}{4}\left\{n\bar{\lambda}\bar{\mu} + \left(\textstyle\sum_{i=1}^{n}(\lambda_i - \bar{\lambda})^2\right)^{1/2} \left(\textstyle\sum_{i=1}^{n}\{(\mu_i - \bar{\mu})^2 + \sigma_i^2\}\right)^{1/2}\right\}.$$

Ένας άλλος τρόπος για να παράγουμε μία κλειστή μορφή για το άνω φράγμα είναι ο εξής:

Για τυχαίο c και $\lambda > 0$, έχουμε:

$$ER_n \leq -(n-2)\lambda + \tfrac{\lambda}{2}\textstyle\sum_{i=1}^{n} E\left\{\left|\frac{X_i-c}{\lambda}-1\right| + \left|\frac{X_i-c}{\lambda}+1\right|\right\}$$

με ισότητα ανν με πιθανότητα 1

$$X_{1:n} \leq c - \lambda \leq X_{2:n} \leq \cdots \leq X_{n-1:n} \leq c + \lambda \leq X_{n:n}.$$

Παρατηρούμε ότι

$$\lambda U\left(\frac{\mu_i-c}{\lambda}, \frac{\sigma_i}{\lambda}\right) \leq 2\lambda + \frac{1}{2\lambda}[(\mu_i - c)^2 + \sigma_i^2],$$

---

[18] Επιτυγχάνεται και έχει κλειστή μορφή όταν n = 2. Την περίπτωση αυτή θα την εξετάσουμε στην πα-ράγραφο 5.3.

επειδή το δεξί μέλος της ανισότητας είναι ένα άνω φράγμα για τη μέση τιμή

$$E\{|X_i-(c-\lambda)|+|X_i-(c+\lambda)|\}$$ [19]

ενώ το αριστερό μέλος είναι το ελάχιστο άνω φράγμα για την ίδια μέση τιμή όταν $X_i \in \mathcal{F}_{\mathbf{1}}(\mu_i,\sigma_i)$. Έπεται ότι

$$\varphi_n(c,\lambda) \le \overline{\varphi_n}(c,\lambda) := 2\lambda + \frac{1}{4\lambda}\sum_{i=1}^{n}[(\mu_i-c)^2+\sigma_i^2].$$

Η ελαχιστοποίηση της $\overline{\varphi_n}(c,\lambda)$ είναι απλή, αρκεί να θέσουμε $c=\bar{\mu}$ και $\lambda^*=\frac{1}{4}AG_n$ όπως πριν. Παρατηρώντας ότι $\sum_{i=1}^{n}\{(\mu_i-\bar{\mu})^2+{\sigma_i}^2\}=\frac{1}{2}{AG_n}^2$ παίρνουμε το άνω φράγμα

$$\begin{aligned} ER_n \le \inf_{\substack{\lambda>0\\ c\in\mathbb{R}}}\varphi_n(c,\lambda) &\le \inf_{\substack{\lambda>0\\ c\in\mathbb{R}}}\overline{\varphi_n}(c,\lambda) \\ &= \overline{\varphi_n}\left(\bar{\mu},\frac{1}{4}AG_n\right) \\ &= \frac{1}{2}AG_n+\frac{1}{2}AG_n \\ &= AG_n. \end{aligned}$$

Για ισότητα πρέπει και αρκεί $\frac{X_i-\bar{\mu}}{\lambda^*}\in\{-2,0,2\}, i=1,2,\dots,n$. Δηλαδή

$$X_i \in \{\bar{\mu}-2\lambda^*,\bar{\mu},\bar{\mu}+2\lambda^*\} \equiv \left\{\bar{\mu}-\frac{AG_n}{2},\ \bar{\mu},\ \bar{\mu}+\frac{AG_n}{2}\right\},\ i=1,2,\dots,n$$

και

$$X_{1:n}=\bar{\mu}-\frac{AG_n}{2}$$

$$X_{2:n}=\cdots=X_{n-1:n}=\bar{\mu}$$

$$X_{n:n}=\bar{\mu}-\frac{AG_n}{2}.$$

---

[19] Αφού $|y-1|+|y+1|\le 2+\frac{1}{2}y^2$ με ισότητα στα $y\in\{-2,\ 0,2\}$,

$$E\{|X_i-(c-\lambda)|+|X_i-(c+\lambda)|\}\le 2\lambda+\frac{1}{2\lambda}(X_i-c)^2.$$

Επίσης ισχύει ότι

$$E\left\{2\lambda+\frac{1}{2\lambda}(X_i-c)^2\right\}=2\lambda+\frac{1}{2\lambda}[(\mu_i-c)^2+\sigma_i^2].$$

Στο επόμενο κεφάλαιο, θα μελετήσουμε πιο αναλυτικά ποια είναι η ικανή και αναγκαία συνθήκη για να έχω ισότητα στο φράγμα των Arnold και Groeneveld για το range n διατεταγμένων παρατηρήσεων.

5.3. Εύρεση ενός άνω φράγματος για το Range χρησιμοποιώντας την κυρτότητα της $\varphi_n(c,\lambda)$.

Όπως αναφέραμε στην προηγούμενη παράγραφο, για το άνω φράγμα της (5.19) δεν έχουμε κλειστή μορφή. Όλες οι τιμές $\varphi_n(c,\lambda)$ αποτελούν άνω φράγμα της ER_n. Αν το φράγμα επιτυγχάνεται τότε θα πρέπει αναγκαστικά να ισχύει το εξής

$$X_{1:n} \le c-\lambda \le X_{2:n} \le \cdots \le X_{n-1:n} \le c+\lambda \le X_{n:n}.$$

με πιθανότητα 1.

Παράλληλα θα πρέπει κάθε X_i, $i=1,\dots,n$ να επιτυγχάνει ισότητα στο φράγμα του Πορίσματος,

$$E\{|X_i-(c-\lambda)|+|X_i-(c+\lambda)|\} \le \lambda U\left(\frac{\mu_i-c}{\lambda},\frac{\sigma_i}{\lambda}\right),\ i=1,\dots,n.$$

Αυτό για να συμβεί θα πρέπει κάθε X_i να παίρνει το πολύ τρείς τιμές. Αν κάποια X_i παίρνουν τρεις τιμές, αυτές θα είναι $c-2\lambda$, c, $c+2\lambda$.

Σίγουρα, η πιο ενδιαφέρουσα περίπτωση είναι εκείνη κατά την οποία ισχύει η (β) του Πορίσματος 5.2.4. Δηλαδή όταν υπάρχουν $\lambda>0$ και $c\in\mathbb{R}$ ώστε

$$2\lambda|\mu_i-c| < (\mu_i-c)^2+{\sigma_i}^2 < 4\lambda^2,\ i=1,\dots,n.$$

Ας υποθέσουμε ότι μας δίνουν τα $\mu_i,\sigma_i>0$ και φιξάρουμε ένα $\lambda>0$. Έχουμε

$$\frac{\partial}{\partial c}\varphi_n(c,\lambda)=\frac{\lambda}{2}\sum_{i=1}^{n}\frac{\partial}{\partial c}U\left(\frac{\mu_i-c}{\lambda},\frac{\sigma_i}{\lambda}\right).$$

Ας θεωρήσουμε τη συνάρτηση $U_y(x)=U(x,y)$ $(y>0)$ με τύπο

- Αν $0<y<1$,

$$U_y(x)=\begin{cases} 2+\frac{1}{2}(x^2+y^2), & \text{αν}\quad 0\le|x|\le 1-\sqrt{1-y^2},\\ |x|+1+\sqrt{(|x|-1)^2+y^2}, & \text{αν}\quad 1-\sqrt{1-y^2}\le|x|\le 1+\sqrt{1-y^2},\\ 2+\frac{1}{2}(x^2+y^2), & \text{αν}\quad 1+\sqrt{1-y^2}\le|x|\le\sqrt{4-y^2},\\ 2\sqrt{x^2+y^2}, & \text{αν}\quad |x|\ge\sqrt{4-y^2}. \end{cases}$$

- Αν $1\le y<2$,

$$U_y(x) = \begin{cases} 2 + \frac{1}{2}(x^2 + y^2), & \text{αν } 0 \le |x| \le \sqrt{4 - y^2}, \\ 2\sqrt{x^2 + y^2}, & \text{αν } \quad |x| \ge \sqrt{4 - y^2}. \end{cases}$$

- Αν $y \ge 2$

$$U_y(x) = 2\sqrt{x^2 + y^2}, \ \ |x| \ge 0.$$

Σε κάθε μία από τις παραπάνω περιπτώσεις, θα ελέγξω την κυρτότητα και τη μονοτονία της U_y, για να ελέγξω πιθανά ελάχιστα.

Στην πρώτη και πιο περίπλοκη περίπτωση $(0 < y < 1)$ παραγωγίζουμε τη U_y και τότε προκύπτει η συνάρτηση

$$U'_y(x) = \begin{cases} x, & \text{αν } \quad 0 \le x \le 1 - \sqrt{1 - y^2}, \\ 1 + \dfrac{1}{2\sqrt{(x-1)^2 + y^2}} 2(x-1), & \text{αν } \ 1 - \sqrt{1 - y^2} \le x \le 1 + \sqrt{1 - y^2}, \\ x, & \text{αν } \ 1 + \sqrt{1 - y^2} \le x \le \sqrt{4 - y^2}, \\ \dfrac{2x}{\sqrt{x^2 + y^2}}, & \text{αν } \quad x \ge \sqrt{4 - y^2}. \end{cases}$$

(Θεωρούμε μόνο την περίπτωση που $x > 0$, γιατί η $Uy(x)$ είναι άρτια).

Βλέπουμε ότι

$$U'_y\left(\left(1 - \sqrt{1 - y^2}\right)_-\right) = 1 - \sqrt{1 - y^2}$$

$$U'_y\left(\left(1 - \sqrt{1 - y^2}\right)_+\right) = 1 + \frac{\left(1-\sqrt{1-y^2}\right)-1}{\sqrt{\left(\left(1-\sqrt{1-y^2}\right)-1\right)^2+y^2}}$$

$$= 1 - \frac{\sqrt{1-y^2}}{\sqrt{1-y^2+y^2}}$$

$$= 1 - \sqrt{1 - y^2}.$$

$$U'_y\left(\left(1 + \sqrt{1 - y^2}\right)_-\right) = 1 + \frac{\left(1+\sqrt{1-y^2}\right)-1}{\sqrt{\left(\left(1+\sqrt{1-y^2}\right)-1\right)^2+y^2}}$$

$$= 1 + \frac{\sqrt{1-y^2}}{\sqrt{1-y^2+y^2}}$$

$$= 1 + \sqrt{1 - y^2}$$

$$= U'_y\left(\left(1 + \sqrt{1 - y^2}\right)_+\right).$$

Η $U'_y(x)$ είνα γνησίως αύξουσα με $U'_y(x) = 0$ αν και μόνο αν $x = 0$. Επομένως η $U_y(x)$ είναι γνήσια κυρτή και άρτια.

Στη δεύτερη περίπτωση που $1 \le y < 2$ παραγωγίζοντας ως προς x τη συνάρτηση προκύπτει

$$\frac{\partial}{\partial x} U_y(x) = x, \qquad \text{αν } |x| < \sqrt{4 - y^2}$$
$$\frac{\partial}{\partial x} U_y(x) = \frac{2x}{\sqrt{x^2 + y^2}}, \quad \text{αν } |x| > \sqrt{4 - y^2}$$

Για $x = \sqrt{4 - y^2}$ έχω

$$U'_y(x_-) = \sqrt{4 - y^2}, \quad U'_y(x_+) = \frac{2\sqrt{4-y^2}}{\sqrt{y^2+4-y^2}} = \sqrt{4 - y^2}.$$

Ομοίως για $x = -\sqrt{4 - y^2}$

$$U'_y(x_-) = U'_y(x_+) = -\sqrt{4 - y^2}.$$

$$\frac{\partial^2}{\partial x^2} U_y(x) = 1, \qquad \text{αν } |x| < \sqrt{4 - y^2}$$
$$\frac{\partial^2}{\partial x^2} U_y(x) = \frac{2y^2}{(x^2 + y^2)^{3/2}}, \quad \text{αν } |x| > \sqrt{4 - y^2}$$

Στο $x = \sqrt{4 - y^2}$ έχω $U''_y(x_-) = 1$ και $U''_y(x_+) = \frac{2y^2}{8} = \frac{1}{4}y^2 < 1.$

Οπότε, όταν $1 \le y < 2$, η $U_y(x)$ δεν είναι δυο φορές παραγωγίσιμη στα σημεία αλλαγής, $x \pm \sqrt{4 - y^2}$. Η $U'_y(x)$ όμως, είναι γνήσια αύξουσα:

$$U_y(x) = \begin{cases} \dfrac{2x}{\sqrt{x^2+y^2}}, & \text{αν } \; x \le -\sqrt{4-y^2}, \\ x, & \text{αν } \; -\sqrt{4-y^2} \le x \le \sqrt{4-y^2}, \\ \dfrac{2x}{\sqrt{x^2+y^2}}, & \text{αν } \; x \ge \sqrt{4-y^2}. \end{cases}$$

΄Αρα η $U_y(x)$ είναι πάλι άρτια, γνήσια κυρτή, με σημείο ελαχίστου το $x = 0$ (και τιμή ίση με 2).

Στην τρίτη, και τελευταία, περίπτωση, που $y \ge 2$, $\frac{\partial}{\partial x} U_y(x) = \frac{2x}{\sqrt{x^2+y^2}}$. Έχει ελάχιστο στο x = 0, και αν ξαναπαραγωγίσουμε, προκύπτει ότι $\frac{\partial^2}{\partial x^2} U_y(x) = \frac{2y^2}{(x^2+y^2)^{3/2}} > 0$. Άρα η $U_y(x)$ είναι γνήσια κυρτή όταν $y \ge 2$, άρτια, με ελάχιστο x = 0.

Οπότε συνοψίζοντας και τις τρεις παραπάνω περιπτώσεις, προκύπτει ότι η $U'_y(x)$ είναι τελικά γνησίως αύξουσα με $U'_y(x) = 0$ αν και μόνον αν x = 0. Τελικά η $U_y(x)$ είναι γνήσια κυρτή και άρτια για κάθε y > 0.

Τώρα βλέπουμε ότι

$$\frac{\partial}{\partial c} U\left(\frac{\mu_i - c}{\lambda}, \frac{\sigma_i}{\lambda}\right) = \frac{\partial}{\partial c} U_{\sigma_i/\lambda}\left(\frac{\mu_i - c}{\lambda}\right)$$
$$= -\frac{1}{\lambda} U'_{\sigma_i/\lambda}\left(\frac{\mu_i - c}{\lambda}\right).$$

Επομένως

$$\frac{\partial}{\partial c} \varphi_n(c,\lambda) = -\frac{1}{2} \sum_{i=1}^{n} U'_{\sigma_i/\lambda}\left(\frac{\mu_i - c}{\lambda}\right).$$

Αν το c είναι μικρότερο από όλα τα μ_i, δηλαδή $c < \mu_{1:n}$, τότε $U'_{\sigma_i/\lambda}\left(\frac{\mu_i - c}{\lambda}\right) > 0$ και $\frac{\partial}{\partial c} \varphi_n(c,\lambda) < 0$.

Ομοίως, αν $c > \mu_{n:n}$, τότε $U'_{\sigma_i/\lambda}\left(\frac{\mu_i - c}{\lambda}\right) < 0$ και $\frac{\partial}{\partial c} \varphi_n(c,\lambda) > 0$.

Η $\varphi_n(c,\lambda)$ γράφεται ως:

$$\varphi_n(c,\lambda) = \frac{\lambda}{2} \sum_{i=1}^{n} \left\{ U_{\sigma_i/\lambda}\left(\frac{\mu_i - c}{\lambda}\right) - \left(1 - \frac{2}{n}\right) \right\},$$

δηλαδή είναι άθροισμα γνήσια κυρτών συναρτήσεων, άρα θα είναι γνήσια κυρτή (ως προς c).

Για δεδομένο λ > 0, η $\varphi_n(c,\lambda)$ ελαχιστοποιείται όταν και μόνο όταν βρίσκουμε

$$c = c(\lambda) : \sum_{i=1}^{n} U'_{\sigma_i/\lambda}\left(\frac{\mu_i - c}{\lambda}\right) = 0 \tag{5.21}$$

Ας εκλέξουμε τώρα $\lambda > 0$ τέτοιο ώστε $\frac{\sigma_i}{\lambda} \geq 2 \Leftrightarrow \lambda \leq \frac{\sigma_i}{2}$, για κάθε $i = 1, 2, \ldots, n$. Τότε

$$U'_{\sigma_i/\lambda}\left(\frac{\mu_i - c}{\lambda}\right) = \frac{2\left(\frac{\mu_i - c}{\lambda}\right)}{\sqrt{\left(\frac{\mu_i - c}{\lambda}\right)^2 + \frac{\sigma_i^2}{\lambda^2}}} = \frac{2(\mu_i - c)}{\sqrt{(\mu_i - c)^2 + \sigma_i^2}}$$

και η (5.21) δίνει

$$c = c(\lambda) \quad : \quad \sum_{i=1}^{n} \frac{c - \mu_i}{\sqrt{(\mu_i - c)^2 + \sigma_i^2}} = 0 \ \ (\text{ανεξάρτητη του } \lambda).$$

Όμως

$$U_{\sigma_i/\lambda}\left(\frac{\mu_i - c}{\lambda}\right) = 2\sqrt{\left(\frac{\mu_i - c}{\lambda}\right)^2 + \frac{\sigma_i^2}{\lambda^2}} = \frac{2}{\lambda}\sqrt{(\mu_i - c)^2 + \sigma_i^2}$$

και

$$\varphi_n(c, \lambda) = -(n-2)\lambda + \sum_{i=1}^{n} \sqrt{(\mu_i - c)^2 + \sigma_i^2}.$$

Αυτή επιτάσσει να βάλουμε μεγάλο λ, $\lambda^* = \frac{1}{2}\min_{1 \leq i \leq n} \sigma_i$.

Οπότε καταλήξαμε σε ένα νέο φράγμα για την ER_n το οποίο δίνεται στη παρακάτω πρόταση:

Πρόταση 5.3.1 Για τη μέση τιμή του εύρους δεδομένων των μ_i και σ_i^2 ισχύει η ανισότητα

$$ER_n \leq -\frac{1}{2}(n-2)\min_{1 \leq i \leq n} \sigma_i + \sum_{i=1}^{n} \sqrt{(\mu_i - c)^2 + \sigma_i^2} \tag{5.23}$$

όπου c η μοναδική λύση της εξίσωσης

$$\sum_{i=1}^{n} \frac{c - \mu_i}{\sqrt{(\mu_i - c)^2 + \sigma_i^2}} = 0. \tag{5.24}$$

■

Το φράγμα πάλι δεν είναι βέλτιστο. Για παράδειγμα, αν θεωρήσουμε $\mathbf{X} \in \mathcal{F}_3(\boldsymbol{\mu}, \boldsymbol{\sigma})$, με $\boldsymbol{\mu} = (4, 0, 7)$ και $\boldsymbol{\sigma} = (1, 3, 1)$, βρίσκουμε ότι $c = 4.136$. Στην περίπτωση αυτή το φράγμα της Πρότασης 5.3.1 είναι ίσο με 8.65. Αν τα συγκεκριμένα **μ, σ,** λ και c τα αντικαταστήσω στην $ER_n \leq -(n-2)\lambda + \frac{\lambda}{2}\sum_{i=1}^{n} U\left(\frac{\mu_i - c}{\lambda}, \frac{\sigma_i}{\lambda}\right)$ βρίσκω ελαφρώς μικρότερο άνω φράγμα.

Παρατήρηση 5.3.2 Ειδική περίπτωση για n = 2.

Έως τώρα δεν είχαμε βρει κάποια κλειστή μορφή για το φράγμα του range και αν τελικά αυτό είναι βέλτιστο. Με αυτά που έχουμε αναφέρει μέχρι αυτό το Κεφάλαιο, η μόνη περίπτωση που μπορούμε να βρούμε βέλτιστο φράγμα είναι η περίπτωση που έχουμε δύο παρατηρήσεις. Εφαρμόζοντας τον (5.24) για n = 2, τελικά βρίσκουμε $c^* = \frac{\sigma_1\mu_2+\sigma_2\mu_1}{\sigma_1+\sigma_2}$. Διαλέγουμε $\lambda^* = \frac{\sigma_1\theta}{2(\sigma_1+\sigma_2)}$ όπου $\theta = \sqrt{(\mu_1-\mu_2)^2+(\sigma_1+\sigma_2)^2}$,

$\sigma_1 \le \sigma_2$. Αντικαθιστώντας στην $\varphi_n(c,\lambda)$, προκύπτει

$$\begin{aligned}
\varphi_2(c^*,\lambda^*) &= \frac{\sigma_1\theta}{4(\sigma_1+\sigma_2)}\left\{U\left(2\frac{|\mu_1-\mu_2|}{\theta}, 2\frac{\sigma_1+\sigma_2}{\theta}\right) + U\left(2\frac{\sigma_2}{\sigma_1}\frac{|\mu_1-\mu_2|}{\theta}, 2\frac{\sigma_2}{\sigma_1}\frac{\sigma_1+\sigma_2}{\theta}\right)\right\} \\
&= \frac{\sigma_1\theta}{4(\sigma_1+\sigma_2)}\left\{2\sqrt{4\frac{(\mu_1-\mu_2)^2}{\theta^2}+4\frac{(\sigma_1+\sigma_2)^2}{\theta^2}} + 2\frac{\sigma_2}{\sigma_1}\sqrt{4\frac{(\mu_1-\mu_2)^2}{\theta^2}+4\frac{(\sigma_1+\sigma_2)^2}{\theta^2}}\right\} \\
&= \frac{\sigma_1\theta}{4(\sigma_1+\sigma_2)}\cdot\frac{4\theta}{\theta}\cdot\left\{1+\frac{\sigma_2}{\sigma_1}\right\} \\
&= \frac{\sigma_1\theta}{\sigma_1+\sigma_2}\cdot\frac{\sigma_1+\sigma_2}{\sigma_1} \\
&= \theta.
\end{aligned}$$

Οπότε αν $c^* = \frac{\sigma_1\mu_2+\sigma_2\mu_1}{\sigma_1+\sigma_2}$, $\lambda^* = \frac{\sigma_1\theta}{2(\sigma_1+\sigma_2)}$ όπου $\theta = \sqrt{(\mu_1-\mu_2)^2+(\sigma_1+\sigma_2)^2}$,

$\sigma_1 \le \sigma_2$, τότε $\varphi_2(c^*,\lambda^*) = \theta = \mathrm{supER}_2$.

ΚΕΦΑΛΑΙΟ 6

ΕΥΡΕΣΗ ΒΕΛΤΙΣΤΩΝ ΑΝΩ ΦΡΑΓΜΑΤΩΝ ΓΙΑ ΤΟ ΕΥΡΟΣ n ΔΙΑΤΕΤΑΓΜΕΝΩΝ ΠΑΡΑΤΗΡΗΣΕΩΝ.

Στο κεφάλαιο 5, έγινε αναφορά για τον τρόπο που μπορούμε να προσδιορίσουμε κάποια άνω φράγματα για τη μέση τιμή του εύρους n διατεταγμένων παρατηρήσεων. Επίσης βρήκαμε και τις αντίστοιχες κατανομές πιθανότητας που επιτυγχάνονται τα φράγματα αυτά. Στο κεφάλαιο αυτό θα ασχοληθούμε με την εύρεση βέλτιστων φραγμάτων για το εύρος, και πιο συγκεκριμένα θα ασχοληθούμε με το πότε το φράγμα των Arnold – Groeneveld επιτυγχάνεται.

6.1. Πότε το φράγμα των Arnold – Groeneveld επιτυγχάνεται;

Όπως παρατηρήσαμε στην ενότητα 5.1, από το Θεώρημα των Arnold και Groeneveld (1979) προκύπτει ότι για οποιεσδήποτε σταθερές $\lambda_1, \lambda_2, \ldots, \lambda_n$,

$$\left|\sum_{i=1}^{n} \lambda_i (E(X_{i:n}) - \bar{\mu})\right| \leq \left(\sum_{i=1}^{n} (\lambda_i - \bar{\lambda})^2\right)^{1/2} \left(\sum_{i=1}^{n} \left\{(\mu_i - \bar{\mu})^2 + \sigma_i^2\right\}\right)^{1/2} \quad (6.1)$$

όπου $\bar{\mu} = \frac{1}{n}\sum_{i=1}^{n} \mu_i$ και $\bar{\lambda} = \frac{1}{n}\sum_{i=1}^{n} \lambda_i$. Εφαρμόζοντας την (6.1) στην περίπτωση του εύρους, προκύπτει η ανισότητα

$$ER_n \leq AG_n := \sqrt{2}\left(\sum_{i=1}^{n} \left\{(\mu_i - \bar{\mu})^2 + \sigma_i^2\right\}\right)^{1/2}. \quad (6.2)$$

Οι Arnold – Groeneveld (1979), ο Rychlik (1993b) και ο Papadatos (2001) έδειξαν ότι αν $\mu_i = \mu$ και $\sigma_i = \sigma$ για όλα τα i, τότε το φράγμα της (6.2), το οποίο γίνεται $\sigma\sqrt{2n}$, επιτυγχάνεται. Τώρα θα παρουσιάσουμε έναν ακριβή χαρακτηρισμό για το πότε επιτυγχάνεται το φράγμα AG_n όταν δίνονται η μέση τιμή και η διασπορά.

Θεώρημα 6.1.1. Έστω $E\mathbf{X} = \boldsymbol{\mu}$ και $\mathrm{Var}\mathbf{X} = \boldsymbol{\sigma}^2$. Τότε η ισότητα στην (6.2) επιτυγχάνεται αν και μόνο αν ικανοποιούνται συγχρόνως οι παρακάτω ανισότητες:

$$\text{(i)} \quad |\mu_i - \bar{\mu}| \leq \frac{\sqrt{2}[(\mu_i - \bar{\mu})^2 + \sigma_i^2]}{\sqrt{\sum_{j=1}^{n}\left\{(\mu_j - \bar{\mu})^2 + \sigma_j^2\right\}}}$$

$$\text{(ii)} \quad \frac{(\mu_i - \bar{\mu})^2 + \sigma_i^2}{\sqrt{\sum_{j=1}^{n}\left\{(\mu_j - \bar{\mu})^2 + \sigma_j^2\right\}}} \leq \frac{1}{2} \qquad i = 1, 2, \ldots, n. \quad (6.3)$$

Θεωρώντας ότι πληρούνται οι (i) και (ii), όλα τα «ακραία» (extremal) τυχαία διανύσματα $\mathbf{X}$[20], έχουν την αναπαράσταση

[20] δηλαδή $\mathbf{X} \in \mathcal{F}_{\mathbf{n}}(\boldsymbol{\mu}, \boldsymbol{\sigma})$ έτσι ώστε $ER_n(\mathbf{X}) = AG_n$

$$\mathbf{X} = \mathrm{g}(\mathrm{X},\mathrm{Y}) := \bar{\boldsymbol{\mu}} + \frac{\mathbf{e}(\mathrm{X})-\mathbf{e}(\mathrm{Y})}{\sqrt{2}}\sqrt{\sum_{j=1}^{n}\left\{\left(\mu_j-\bar{\mu}\right)^2+\sigma_j^2\right\}}, \tag{6.4}$$

όπου $\mathbf{e}(\mathrm{i}) = (0,\dots,1,\dots,0)^t \in \mathbb{R}^n$ είναι το i-οστό μοναδιαίο διάνυσμα, $\bar{\boldsymbol{\mu}} = (\bar{\mu},\dots,\bar{\mu})^t \in \mathbb{R}^n$, και (X,Y) είναι ένα διακριτό τυχαίο ζεύγος με περιθώριες κατανομές (marginals)

$$\mathbb{P}[\mathrm{X}=\mathrm{i}] = p_i^{+} = \frac{(\mu_i-\bar{\mu})^2+\sigma_i^2+\frac{1}{2}(\mu_i-\bar{\mu})\mathrm{AG}_n}{\sum_{j=1}^{n}\left\{\left(\mu_j-\bar{\mu}\right)^2+\sigma_j^2\right\}},$$

$$\mathrm{i} = 1, 2, \dots, n, \tag{6.5}$$

$$\mathbb{P}[\mathrm{Y}=\mathrm{i}] = p_i^{-} = \frac{(\mu_i-\bar{\mu})^2+\sigma_i^2-\frac{1}{2}(\mu_i-\bar{\mu})\mathrm{AG}_n}{\sum_{j=1}^{n}\left\{\left(\mu_j-\bar{\mu}\right)^2+\sigma_j^2\right\}},$$

έτσι ώστε

$$\mathbb{P}[\mathrm{X}=\mathrm{Y}] = 0. \tag{6.6}$$

Επιπρόσθετα, αν οι ανισότητες στην (6.3) είναι γνήσιες για όλα τα i, μπορούμε να βρούμε απείρως πολλά «ακραία» τυχαία διανύσματα $\mathbf{X}$, και αν η (6.3) ικανοποιείται και για κάποια i έχουμε ισότητα στη (ii), τότε το ακραίο τυχαίο διάνυσμα είναι μοναδικό.

Παρατήρηση 6.1.2. Έστω $(\mu_1,\mu_2,\mu_3) = (-1,0,1)$, $(\sigma_1^2,\ \sigma_2^2,\sigma_3^2) = (1,3,2)$, τότε η (6.3) ισχύει. Ωστόσο, η (6.3) (ii) ικανοποιείται με γνήσιες ανισότητες, κάτι το οποίο δε συμβαίνει στην (6.3)(i). Βρίσκουμε AG_3=4 και $\boldsymbol{p}^{+} = \left(0,\frac{3}{8},\frac{5}{8}\right)$, $\boldsymbol{p}^{-} = \left(\frac{1}{2},\frac{3}{8},\frac{1}{8}\right)$. Είναι εύκολο να δούμε ότι η κατανομή του $(\mathrm{X}_1,\ \mathrm{X}_2,\ \mathrm{X}_3)$ (που φαίνεται στην (6.4)) ορίζεται μοναδικά: αντιστοιχούμε τις πιθανότητες $\frac{2}{8},\frac{2}{8},\frac{3}{8},\frac{1}{8}$, στα σημεία $(-2,2,0)$, $(-2,0,2)$, $(0,-2,2)$, $(0,2,-2)$. Έπεται ότι το τυχαίο διάνυσμα που επιτυγχάνει το φράγμα AG_n, μπορεί να είναι μοναδικό ακόμη και αν η (6.3)(ii) ικανοποιείται με γνήσιες ανισότητες για όλα τα i.

Για την απόδειξη του Θεωρήματος 6.1.1., θα πρέπει να αναφέρουμε τα παρακάτω:

Έστω (X,Y) το διδιάστατο τυχαίο ζευγάρι μεταβλητών με στήριγμα (support) ένα υποσύνολο του $\{1,2,\dots,n\}^2$. Το $(\mathrm{X},\mathrm{Y})\sim\mathrm{A}$, όπου $\mathrm{A} = \left(\alpha_{ij}\right)$, αν $\mathbb{P}[\mathrm{X}=\mathrm{i},\mathrm{Y}=\mathrm{j}] = \alpha_{ij}$ για όλα τα i, j. Πιο απλά, για την τυχαία μεταβλητή X, με στήριγμα ένα υποσύνολο του $\{1,2,\dots,n\}$, λέμε ότι $X\sim\boldsymbol{\alpha}$, όπου $\boldsymbol{\alpha} = (\alpha_1,\dots,\alpha_n)$, για να περιγράψουμε ότι η κατανομή πιθανότητάς του είναι $\mathbb{P}[X = i] = \alpha_i$ για κάθε i.

Λήμμα 6.1.3. Έστω $\boldsymbol{p} = (p_1,\dots,p_n)$ και $\boldsymbol{q} = (q_1,\dots,q_n)$ δυο διανύσματα πιθανότητας, δηλαδή $p_i \geq 0$, $q_i \geq 0$ για όλα τα i, και $\sum_{i=1}^{n} p_i = \sum_{i=1}^{n} q_i = 1$. Ικανή και αναγκαία συνθήκη για την ύπαρξη του (X, Y) με

$$\mathbb{P}[X = Y] = 0,\ X\sim\boldsymbol{p},\ Y\sim\boldsymbol{q} \tag{6.7}$$

είναι η παρακάτω:

$$\max_{1\leq i\leq n}\{p_i + q_i\} \leq 1. \tag{6.8}$$

Αν η ισότητα ισχύει στην (6.8) τότε το τυχαίο ζευγάρι ορίζεται μοναδικά. Αν στην (6.8) ισχύει γνήσια ανισότητα και $\min_i\{p_i q_i\} > 0$ τότε υπάρχουν απείρως πολλά ζευγάρια διανυσμάτων που ικανοποιούν την (6.7).

Απόδειξη Λήμματος

Για n = 1 οι (6.7) και (6.8) είναι άκυρες, έτσι δεν έχουμε κάτι να αποδείξουμε. Για n = 2 το αποτέλεσμα είναι τετριμμένο (έχουμε μοναδικότητα αν η (6.7) ικανοποιείται, και έχουμε ισότητα στην (6.8) οποτεδήποτε πληρείται).

Υποθέτουμε ότι $n \geq 3$ και θεωρούμε το σύνολο από όλους τους πίνακες πιθανότητας με δεδομένες τις περιθώριες πιθανότητες (marginals),

$$\mathcal{M}(\boldsymbol{p},\boldsymbol{q}) := \left\{Q = \left(q_{ij}\right) \in \mathbb{R}^{nxn}: q_{ij} \geq 0, \sum_{i=1}^{n} q_{ij} = q_j, \sum_{j=1}^{n} q_{ij} = p_i \text{ για όλα τα } i, j \right\}.$$

Το σύνολο $\mathcal{M}(\boldsymbol{p},\boldsymbol{q})$ είναι μη κενό, αφού περιέχει τον πίνακα $Q = \left(p_i q_j\right)$. Επίσης η συνάρτηση $f(Q) :=$ ίχνος$(Q) = \sum_{i=1}^{n} q_{ii}$ είναι συνεχής στο μετρικό χώρο $\langle\mathbb{R}^{nxn}, d\rangle$, όπου d είναι η απόσταση ολικής κύμανσης (total variation distance), $d\left(Q,\widetilde{Q}\right) = \sum_{i,j}\left|q_{ij} - \widetilde{q_{ij}}\right|$ (ή οποιοδήποτε ισοδύναμη μέτρική στον $\mathbb{R}^{nxn}$). Επιπρόσθετα, το $\mathcal{M}(\boldsymbol{p},\boldsymbol{q})$ είναι συμπαγές υποσύνολο του $\mathbb{R}^{nxn}$, αφού προφανώς είναι κλειστό, και περιέχεται σε μία μπάλα με κέντρο τον πίνακα 0_{nxn} και ακτίνα 1. Αφού το πεδίο ορισμού της f είναι το κλειστό και φραγμένο σύνολο $\mathcal{M}(\boldsymbol{p},\boldsymbol{q})$, έπεται ότι η $f(Q)$ αποκτάει ελάχιστο για κάποια $Q^* \in \mathcal{M}(\boldsymbol{p},\boldsymbol{q})$.

Έστω $(X, Y)\sim Q^*$ όπου $Q^* \in \mathcal{M}(\boldsymbol{p},\boldsymbol{q})$ είναι ένας πίνακας ελαχιστοποίησης. Τότε $X\sim\boldsymbol{p}$, $Y\sim\boldsymbol{q}$ και $f(Q^*) = \mathbb{P}[X = Y]$. Η κύρια διαγώνιος ενός οποιουδήποτε πίνακα ελαχιστοποίησης Q^*, μπορεί να περιέχει το πολύ ένα μη μηδενικό στοιχείο. Πράγματι, αν ${q_{ii}}^* > 0\ \textit{και}\ {q_{jj}}^* > 0$ με $i \neq j$, τότε θέτουμε $\gamma = \min\left\{{q_{ii}}^*, {q_{jj}}^*\right\} > 0$. Μετά, θεωρούμε τον πίνακα $\widetilde{Q} = \left(\tilde{q}_{ij}\right)$, ο οποίος διαφέρει από τον Q^* στα τέσσερα επόμενα στοιχεία $\tilde{q}_{ii} = q_{ii}^* - \gamma$, $\tilde{q}_{jj} = q_{jj}^* - \gamma$, $\tilde{q}_{ij} = q_{ij}^* + \gamma$, $\tilde{q}_{ji} = q_{ji}^* + \gamma$. Αφού τα αθροίσματα των γραμμών και των στηλών παραμένουν αμετάβλητα και τα στοιχεία του $\widetilde{Q}$ είναι μη αρνητικά, είναι προφανές ότι $\widetilde{Q} \in \mathcal{M}(\boldsymbol{p},\boldsymbol{q})$ και οδηγούμαστε στην αντίφαση

$f(\widetilde{Q}) = f(Q^*) - 2\gamma < f(Q^*)$. Οπότε τα διαγώνια στοιχεία ενός πίνακα ελαχιστοποίησης Q^* είναι μηδενικά, με πιθανή εξαίρεση ενός το πολύ εξ αυτών.

Απόδειξη ικανότητας:

Υποθέτουμε ότι η (6.8) ικανοποιείται, και υποθέτουμε ότι $\min_Q f(Q) = f(Q^*) = \theta > 0$. Έστω $q^*_{\kappa\kappa} = \theta$, έτσι $q^*_{ii} = 0$ για όλα τα $i \neq \kappa$. Τότε,

$$\mathbb{P}[\{X = \kappa\} \cup \{Y = \kappa\}] = \mathbb{P}[X = \kappa] + \mathbb{P}[Y = \kappa] - \mathbb{P}[X = \kappa, Y = \kappa] = p_\kappa + q_\kappa - \theta.$$

Αφού $1 - p_\kappa - q_\kappa \geq 0$ (από την (6.8)) προκύπτει ότι

$$\mathbb{P}[X \neq \kappa, Y \neq \kappa] = 1 - \mathbb{P}[\{X = \kappa\} \cup \{Y = \kappa\}] = \theta + (1 - p_\kappa - q_\kappa) \geq \theta > 0.$$

Από την άλλη πλευρά, αφού $q^*_{ii} = 0$ για όλα τα $i \neq \kappa$, έχουμε

$$\mathbb{P}[X \neq \kappa, Y \neq \kappa] = \sum_{i \neq \kappa,\ j \neq \kappa,\ i \neq j} q^*_{ij}.$$

Η παραπάνω πιθανότητα είναι τουλάχιστον θ, και έτσι, είναι γνήσια θετική. Έτσι οδηγούμαστε στο ότι το άθροισμα (το οποίο είναι άθροισμα μη αρνητικών όρων) περιλαμβάνει τουλάχιστον ένα θετικό όρο. Έτσι μπορούμε να φιξάρουμε δυο δείκτες r, s με $r \neq s$, $s \neq \kappa$, $r \neq s$, έτσι ώστε $q^*_{rs} > 0$. Θέτουμε $\delta = \min\{\theta, q^*_{rs}\} > 0$ και θεωρούμε τον πίνακα $\widetilde{Q} = (\tilde{q}_{ij})$, ο οποίος διαφέρει από τον Q^* μόνο στα τέσσερα επόμενα στοιχεία

$$\tilde{q}_{\kappa\kappa} = q^*_{\kappa\kappa} - \delta = \theta - \delta,\ \tilde{q}_{rs} = q^*_{rs} - \delta,\ \tilde{q}_{r\kappa} = q^*_{r\kappa} + \delta,\ \tilde{q}_{\kappa s} = q^*_{\kappa s} + \delta.$$

Αφού τα αθροίσματα των γραμμών και των στηλών παραμένουν αμετάβλητα και τα στοιχεία του $\widetilde{Q}$ είναι μη αρνητικά, είναι προφανές ότι $\widetilde{Q} \in \mathcal{M}(\boldsymbol{p}, \boldsymbol{q})$ και οδηγούμαστε στην αντίφαση $f(\widetilde{Q}) = \theta - \delta < \theta$. Οπότε $f(Q^*) = \mathbb{P}[X = Y] = 0$. Αυτό αποδεικνύει την ύπαρξη των τυχαίων διανυσμάτων που ικανοποιούν την (6.7).

Απόδειξη της αναγκαιότητας:

Αυτό είναι απλό. Αν ένα τυχαίο διάνυσμα (X, Y) ικανοποιεί την (6.7), τότε $(X, Y) \sim Q$ για κάποια $Q \in \mathcal{M}(\boldsymbol{p}, \boldsymbol{q})$ με $q_{ii} = 0$ για όλα τα i. Έτσι για κάθε i,

$$p_i + q_i = p_i + q_i - q_{ii} = \mathbb{P}[X = i] + \mathbb{P}[Y = i] - q_{ii} = \mathbb{P}[\{X = i\} \cup \{Y = i\}] \leq 1.$$

Περίπτωση μοναδικότητας:

Υποθέτουμε, ότι $\max_i\{p_i + q_i\} = 1$ και διαλέγουμε κ με $p_\kappa + q_\kappa = 1$. Αν $(X, Y) \sim Q$ ικανοποιεί την (6.7) έχουμε $\mathbb{P}[\{X = \kappa\} \cup \{Y = \kappa\}] = p_\kappa + q_\kappa - q_{\kappa\kappa} \geq p_\kappa + q_\kappa - \mathbb{P}[X = Y] = p_\kappa + q_\kappa = 1$. Αυτό οδηγεί στο ότι ο Q μπορεί να έχει μη μηδενικά στοιχεία μόνο στην κ – οστή του γραμμή και στην κ – οστή του στήλη. Έτσι $q_{i\kappa} = p_i$ για όλα τα $i \neq \kappa$, $q_{\kappa j} = q_j$ για όλα τα $j \neq \kappa$, και $q_{ij} = 0$ διαφορετικά. Έτσι, ο Q καθορίζεται μοναδικά από τα $\boldsymbol{p}, \boldsymbol{q}$. Σημειώνουμε ότι το κ, δεν είναι απαραίτητα μοναδικό, αλλά ο Q είναι πάντα μοναδικός. Για παράδειγμα, αν $\boldsymbol{p} = (1-\mathrm{p}, \mathrm{p}, 0, \ldots, 0)$ και $\boldsymbol{q} = (\mathrm{p}, 1-\mathrm{p}, 0, \ldots, 0)$ με $0 \le \mathrm{p} \le 1$, αποκτούμε τη μοναδική λύση στην (6.7) $\mathbb{P}[\mathrm{X}=2, \mathrm{Y}=1] = \mathrm{p} = 1 - \mathbb{P}[\mathrm{X}=1, \mathrm{Y}=2]$. Στην πραγματικότητα, κάποιος μπορεί εύκολα να επιβεβαιώσει ότι το παράδειγμα περιγράφει την πιο γενική μορφή όπου η σχέση $\mathrm{p}_\kappa + \mathrm{q}_\kappa = 1$ ισχύει για περισσότερους από ένα δείκτες κ.

Περίπτωση μη μοναδικότητας:

Υποθέτουμε ότι όλα τα $\mathrm{p_i}$ και $\mathrm{q_i}$ είναι θετικά και ότι η (6.8) είναι γνήσια ανισότητα, δηλαδή $\mathrm{p_i} + \mathrm{q_i} < 1$ για όλα τα i. [η τελευταία υπόθεση είναι πιθανή, αν και μόνο αν n≥3.] Θέτουμε

$$\beta = \frac{1}{\mathrm{n}^2}[1 - \max_{\mathrm{i}}\{\mathrm{p_i} + \mathrm{q_i}\}] > 0, \ \delta = \frac{1}{2}\min_{\mathrm{i,j}}\{\mathrm{p_i q_j}\} > 0 \text{ και } \varepsilon = \min\{\beta, \delta\}.$$

Ορίζουμε

$$\mathcal{M}_\varepsilon(\boldsymbol{p}, \boldsymbol{q}) := \{\mathrm{Q} \in \mathcal{M}(\boldsymbol{p}, \boldsymbol{q}) : \mathrm{q_{ij}} \ge \varepsilon, \text{ για όλα τα } \mathrm{i, j} \text{ με } \mathrm{i} \ne \mathrm{j}\}.$$

Παρατηρούμε ότι το $\mathcal{M}_\varepsilon(\boldsymbol{p}, \boldsymbol{q})$ είναι μη κενό (αφού περιέχει τον $\mathrm{Q} = (\mathrm{p_i q_j})$), συμπαγές υποσύνολο του $\mathbb{R}^{\mathrm{nxn}}$. Εφαρμόζοντας τα ίδια επιχειρήματα όπως στην αρχή, παρατηρούμε ότι η συνεχής συνάρτηση $f(\mathrm{Q}) := \text{ίχνος}(\mathrm{Q})$ επιτυγχάνει το ελάχιστο στον πίνακα $\mathrm{Q}_\varepsilon{}^* = (\mathrm{q_{ij}}^*) \in \mathcal{M}_\varepsilon(\boldsymbol{p}, \boldsymbol{q})$. O $\mathrm{Q}_\varepsilon{}^*$ έχει το πολύ ένα μη μηδενικό διαγώνιο στοιχείο ενώ, από τον ορισμό του $\mathcal{M}_\varepsilon(\boldsymbol{p}, \boldsymbol{q})$, όλα τα μη διαγώνια στοιχεία είναι τουλάχιστον ε. Έστω (X, Y) $\sim\mathrm{Q}_\varepsilon{}^*$. Υποθέτοντας ότι $\mathbb{P}[\mathrm{X}=\mathrm{Y}] = \theta > 0$, μπορούμε να βρούμε ένα μοναδικό δείκτη κ τέτοιον ώστε $\mathrm{q}^*_{\kappa\kappa} = \theta$. Υπολογίζουμε

$\mathbb{P}[\{\mathrm{X}=\kappa\} \cup \{\mathrm{Y}=\kappa\}] = \mathrm{p}_\kappa + \mathrm{q}_\kappa - \theta$. Αφού $\mathrm{q}^*_{\mathrm{ii}} = 0$ για τα $\mathrm{i} \ne \kappa$, έχουμε

$$\sum_{\mathrm{i}\ne\kappa,\ \mathrm{j}\ne\kappa,\ \mathrm{i}\ne\mathrm{j}} \mathrm{q}^*_{\mathrm{ij}} = \mathbb{P}[\mathrm{X} \ne \kappa, \mathrm{Y} \ne \kappa] = \theta + (1 - \mathrm{p}_\kappa - \mathrm{q}_\kappa) \ge \theta + \left[1 - \max_{\mathrm{i}}\{\mathrm{p_i} + \mathrm{q_i}\}\right]$$

$$= \theta + \mathrm{n}^2\beta \ge \theta + \mathrm{n}^2\varepsilon > \mathrm{n}^2\varepsilon.$$

Το παραπάνω άθροισμα περιλαμβάνει $(\mathrm{n}-1)(\mathrm{n}-2) < \mathrm{n}^2$ μη αρνητικούς όρους και η ανισότητα δείχνει ότι τουλάχιστον ένας από αυτούς είναι μεγαλύτερος από ε. Έτσι μπορούμε να βρούμε δυο δείκτες r, s με $\mathrm{r} \ne \mathrm{s}, \mathrm{s} \ne \kappa, \mathrm{r} \ne \mathrm{s}$, έτσι ώστε $\mathrm{q}^*_{\mathrm{rs}} > \varepsilon$. Μπορούμε να πούμε ότι ο $\mathrm{q}^*_{\mathrm{rs}}$ γράφεται $\mathrm{q}^*_{\mathrm{rs}} = \varepsilon + \gamma$ με $\gamma > 0$. Θέτω $\lambda = \min\{\theta, \gamma\} > 0$ και θεωρούμε τον πίνακα $\widetilde{\mathrm{Q}}_\varepsilon = (\tilde{\mathrm{q}}_{\mathrm{ij}})$, ο οποίος διαφέρει από τον $\mathrm{Q}^*{}_\varepsilon$ μόνο στα τέσσερα επόμενα στοιχεία

$\tilde{\mathrm{q}}_{\kappa\kappa} = \mathrm{q}^*_{\kappa\kappa} - \lambda = \theta - \lambda \ge 0$, $\tilde{\mathrm{q}}_{\mathrm{rs}} = \mathrm{q}^*_{\mathrm{rs}} - \lambda = \varepsilon + (\gamma - \lambda) \ge \varepsilon$, $\tilde{\mathrm{q}}_{\mathrm{r}\kappa} = \mathrm{q}^*_{\mathrm{r}\kappa} + \lambda$, $\tilde{\mathrm{q}}_{\kappa\mathrm{s}} = \mathrm{q}^*_{\kappa\mathrm{s}} + \lambda$. . Είναι προφανές ότι $\widetilde{\mathrm{Q}}_\varepsilon \in \mathcal{M}_\varepsilon(\boldsymbol{p}, \boldsymbol{q})$ και οδηγούμαστε στην αντίφαση $f(\widetilde{\mathrm{Q}}_\varepsilon) = \theta - \lambda < \theta$. Οπότε $f(\mathrm{Q}^*{}_\varepsilon) = \mathbb{P}[\mathrm{X}=\mathrm{Y}] = 0$. Αυτό δείχνει ότι η ύπαρξη του τυχαίου διανύσματος (X, Y) ικανοποιεί την (6.7) με την πρόσθετη ιδιότητα

$\mathbb{P}[X = i, Y = j] \geq \varepsilon > 0$ για όλα τα $i \neq j$, θεωρώντας το $\varepsilon > 0$ αρκετά μικρό. Δεδομένου ενός πίνακα πιθανότητας $Q_\varepsilon{}^* = (q_{ij}{}^*)$, είναι δυνατόν να κατασκευάσουμε μια διαφορετική λύση, $Q = (q_{ij})$, στην (6.7). Για παράδειγμα, θέτω $q_{12} = q_{12}{}^* - {}^{\varepsilon}/_{2}$, $q_{13} = q_{13}{}^* + {}^{\varepsilon}/_{2}$, $q_{21} = q_{21}{}^* + {}^{\varepsilon}/_{2}$, $q_{23} = q_{23}{}^* - {}^{\varepsilon}/_{2}$, $q_{31} = q_{31}{}^* - {}^{\varepsilon}/_{2}$, $q_{32} = q_{32}{}^* + {}^{\varepsilon}/_{2}$, και αφήνουμε τους υπόλοιπους όρους αμετάβλητους. Τελικά, είναι εύκολο να δούμε ότι αν οι Q_1 και Q_2 λύνουν την (6.7) τότε το ίδιο είναι αληθές για την $Q(t) = tQ_1 + (1-t)Q_2$, $0 \leq t \leq 1$, και η απόδειξη ολοκληρώθηκε.

■

Απόδειξη Θεωρήματος 6.1.1.

Υποθέτουμε ότι $ER_n = AG_n$ για κάποια τυχαία διανύσματα $\mathbf{X}$ με $E\mathbf{X} = \boldsymbol{\mu}$ και $\mathrm{Var}\mathbf{X} = \boldsymbol{\sigma}^2$. Θέτω $c = \bar{\mu}$, $\lambda = \frac{1}{4}AG_n > 0$ και παίρνουμε στη σχέση (5.2) του προηγούμενου κεφαλαίου μέσες τιμές, για να πάρουμε την ανισότητα (παρατήρηση 5.2.6)

$$AG_n = ER_n \leq \frac{-(n-2)AG_n}{4} + \frac{AG_n}{8}\sum_{i=1}^n E\left\{\left|\frac{X_i-\bar{\mu}}{AG_n/_4} - 1\right| + \left|\frac{X_i-\bar{\mu}}{AG_n/_4} + 1\right|\right\}.$$

Στη συνέχεια, από την $|y-1| + |y+1| \leq 2 + \frac{1}{2}y^2$ με ισότητα αν και μόνο αν $y \in \{-2,0,2\}$ παίρνουμε

$$\sum_{i=1}^n E\left\{\left|\frac{X_i-\bar{\mu}}{AG_n/_4} - 1\right| + \left|\frac{X_i-\bar{\mu}}{AG_n/_4} + 1\right|\right\} \leq 2n + \frac{1}{2}\sum_{i=1}^n E\left\{\left(\frac{X_i-\bar{\mu}}{AG_n/_4}\right)^2\right\} = 2n + 4.$$

Αφού $\frac{-(n-2)AG_n}{4} + \frac{AG_n}{8}(2n+4) = AG_n$, οι προηγούμενες ανισότητες είναι, στην πραγματικότητα, ισότητες. Επομένως η $AG_n = ER_n$ είναι ισοδύναμη με την σχέση (5.3) του προηγούμενου κεφαλαίου

$$X_{1:n} \leq c - \lambda \leq X_{2:n} \leq \cdots \leq X_{n-1:n} \leq c + \lambda \leq X_{n:n},$$

(με $c = \bar{\mu}$, $\lambda = \frac{1}{4}AG_n$) και $\frac{X_i-\bar{\mu}}{AG_n/_4} \in \{-2,0,2\}$, $i = 1,\ldots,n$ (φυσικά, αρκεί να ισχύει με πιθανότητα 1). Επομένως, $AG_n = ER_n$ αν και μόνο αν

(α) $X_{1:n} \leq \bar{\mu} - \frac{1}{4}AG_n \leq X_{2:n} \leq \cdots \leq X_{n-1:n} \leq \bar{\mu} + \frac{1}{4}AG_n \leq X_{n:n}$

και (6.9)

(β) $X_i \in \left\{\bar{\mu} - \frac{1}{2}AG_n, \bar{\mu}, \bar{\mu} + \frac{1}{2}AG_n\right\}, i = 1, \ldots, n,$

με πιθανότητα 1. Ως εκ τούτου, το στήριγμα (support) οποιουδήποτε ακραίου (extremal) τυχαίου διανύσματος, είναι ένα υποσύνολο του

$$S := \left\{\left(\bar{\mu}, \bar{\mu}, \dots, \bar{\mu}, \bar{\mu} + \tfrac{1}{2}AG_n, \bar{\mu}, \dots, \bar{\mu}, \bar{\mu} - \tfrac{1}{2}AG_n, \bar{\mu}, \dots, \bar{\mu}\right)\right\},$$

Όπου τα πρόσημα «+» και « - » παρουσιάζονται σε δύο διαφορετικά σημεία. Το S, έχει n(n – 1) στοιχεία και μπορεί να γραφτεί

$$S = \left\{\bar{\boldsymbol{\mu}} + \frac{e(i)-e(j)}{2}AG_n : (i,j) \in \{1,2,\dots,n\}^2, i \neq j\right\}.$$

Έστω $S' := \{(i,j) \in \{1,2,\dots,n\}^2, i \neq j\}$ και παρατηρούμε ότι η συνάρτηση $g\colon S' \longrightarrow S$, η οποία στέλνει το (i,j) στο $g(i,j) = \bar{\boldsymbol{\mu}} + \frac{e(i)-e(j)}{2}AG_n$, είναι 1 – 1 και επί. Αυτό σημαίνει ότι το $(X,Y) := g^{-1}(\mathbf{X})$ είναι ένα τυχαίο ζευγάρι με τιμές σε ένα υποσύνολο του S', και $\mathbf{X} = g(X,Y)$. Αυτό επιβεβαιώνει την αναπαράσταση στη σχέση (6.4). Για $i \in \{1,2,\dots n\}$ θέτουμε

$$p_i^+ := \mathbb{P}\left[X_i = \bar{\mu} + \tfrac{1}{2}AG_n\right] = \mathbb{P}[X = i], \quad p_i^- := \mathbb{P}\left[X_i = \bar{\mu} - \tfrac{1}{2}AG_n\right] = \mathbb{P}[Y = i],$$

ούτως ώστε, $\mathbb{P}[X_i = \bar{\mu}] = 1 - p_i^+ - p_i^-$. Από την $EX_i = \mu_i$ παίρνουμε $p_i^+ + p_i^- = \frac{2(\mu_i - \bar{\mu})}{2AG_n}$ και από την $E\{(X_i - \bar{\mu})^2\} = (\mu_i - \bar{\mu})^2 + \sigma_i^2$ παίρνουμε τη σχέση $p_i^+ + p_i^- = \frac{4[(\mu_i - \bar{\mu})^2 + \sigma_i^2]}{AG_n^2}$. Επομένως

$$p_i^+ = \frac{2[(\mu_i-\bar{\mu})^2+\sigma_i^2]+(\mu_i-\bar{\mu})AG_n}{AG_n^2}, \; p_i^- = \frac{2[(\mu_i-\bar{\mu})^2+\sigma_i^2]-(\mu_i-\bar{\mu})AG_n}{AG_n^2},$$

οπότε προκύπτει η (6.5). Επομένως, μπορούμε να βρούμε ένα τυχαίο διάνυσμα $\mathbf{X}$ με $E\mathbf{X} = \boldsymbol{\mu}$ και $\mathrm{Var}\mathbf{X} = \boldsymbol{\sigma}^2$ και $ER_n = AG_n$ αν και μόνο αν η παραπάνω κατασκευή του τυχαίου ζεύγους (X, Y) με $\mathbb{P}[X = Y] = 0$, είναι εφικτή. Σύμφωνα με το Λήμμα 6.1.3, αυτό είναι ισοδύναμο με το $\max_i\{p_i^+ + p_i^-\} \leq 1$, που δίνει την (6.3)(ii) (επίσης εγγυάται ότι $\mathbb{P}[X_i = \bar{\mu}] = 1 - p_i^+ - p_i^- \geq 0$), ενώ η (6.3)(i) προκύπτει από το ότι $p_i^+ \geq 0$ και $p_i^- \geq 0$.

Τέλος, οι ανισότητες στην (6.3) είναι γνήσιες για όλα τα i αν και μόνο αν $p_i^+ + p_i^- < 1$ και $p_i^+ p_i^- > 0$ για όλα τα i. Το Λήμμα 6.1.3 δείχνει ότι υπάρχουν απείρως πολλά διανύσματα με τις παραπάνω ιδιότητες σε αυτήν την περίπτωση. Επίσης, αν η (6.3) ικανοποιείται και έχουμε ισότητα στην (6.3)(ii) για κάποια i, η μοναδικότητα προκύπτει ξανά από το Λήμμα 6.1.3.

Παρατήρηση – Παραπομπή 6.1.5. Στο Κεφάλαιο αυτό βρήκαμε πότε το φράγμα των Arnold και Groeneveld είναι βέλτιστο. Μία γενική περίπτωση για ένα βέλτιστο φράγμα του εύρους βρίσκεται στο άρθρο Papadatos, N. (2014). *Maximizing the expected range from dependent observations under mean – variance information.*

Βιβλιογραφία

[1] Arnold, B.C., Groeneveld, R.A.,(1979). *Bounds on expectations of linear systematic statistics based on depended samples.* Ann. Statist. **7**, 220 – 223; Correction 8 1401.

[2] Bertsimas, D., Natarajan, K., Teo, C – P., (2006). *Tight Bounds on expected order statistics.* Prob. Eng. Inf. Sc. **20**, 667 – 686.

[3] Gumbel, E.J., (1954). *The maxima of the mean largest value and of the range.* Ann. Math. Statist. **25**, 76 – 84.

[4] Galambos, J., (1978). *The Asymptotic Theory of extreme order statistics.* U.S.A, Wiley.

[5] Hartley, H.O. and David, H.A., (1954). *Universal boundsfor mean range and extreme observation.* Ann. Math. Statist. **25**, 85 – 99.

[6] Hawkins, D. M., (1971). *On the bounds of the range of order statistics.* J. Amer. Statist. Assoc. **66**, 644 – 645.

[7] Issi, K., (1962). *On sharpness of Tchebycheff – Type inequalities.* Ann. Inst. Statist. Math, **14** iss. 1, 185 – 197.

[8] Ludwig, O. (1959). *Ungleichungen für Extremwerte und andere Ranggrößen in Anwendung auf biometrixche Probleme.* Biom. Zeit. **1**, 203 – 209.

[9] Pearson, E. S. and Chandra Secar, C. (1936). *The efficiency of statistical tools and a criterion for rejection for outlying observations.* Biometrica **28,** 308 – 320.

[10] Papadatos, N., (2001). *Expectation bounds on linear estimators from dependent samples.* J. Statist. Plann. **93**, 17 – 27.

[11] Papadatos, N., (2014). *Maximizing the expected range from dependent observations under mean – variance information.*

[12] Rychlik, T. (1993b). *Sharp bounds on L – estimates and their expectations for depended samples.* Commun. Statist. Theory Meth **22,** 1053 – 1068. Erratum in Commun. Statist. Theory Meth. **23,** 305 – 306.

[13] Αντωνίου, Α. Δ., (2004). *Φράγματα αναμενόμενων τιμών γενικών γραμμικών εκτιμητριών για ισόνομες παρατηρήσεις.*

[14] Νεγρεπόντης, Σ., Γιωτόπουλος, Σ., Γιαννακούλιας, Ε., (1999). *Απειροστικός Λογισμός (Τόμος Ι, ΙΙα).* Εκδόσεις Συμμετρία.

[15] Παπαδάτος, Ν., (2006). *Θεωρία Πιθανοτήτων.* Ε.Κ.Π.Α.

[16] Σταθακόπουλος Κ. Α., (1999). *Πραγματική Ανάλυση.* Αίθρα.

[17] Τσίτσας, Λ., (2002). *Εφαρμοσμένος διανυσματικός Απειροστικός Λογισμός.* Εκδόσεις Συμμετρία.

[18] Χαραλαμπίδης, Χ. Α., (1999). *Θεωρία Πιθανοτήτων και εφαρμογές.* Εκδόσεις Συμμερία.