ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΕΦΑΡΜΟΣΜΕΝΩΝ ΜΑΘΗΜΑΤΙΚΩΝ ΚΑΙ ΦΥΣΙΚΩΝ ΕΠΙΣΤΗΜΩΝ

ΤΟΜΕΑΣ ΜΑΘΗΜΑΤΙΚΩΝ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Η Μέθοδος Δέλτα με εφαρμογές στην ασυμπτωτική συμπεριφορά των δειγματικών κεντρικών ροπών.

ΔΗΛΕ ΑΝΔΡΙΑΝΗ

<u>Επιβλέπων καθηγητής:</u>

Κουκουβίνος Χρήστος, Καθηγητής Ε.Μ.Π.

Παπαδάτος Νίκος, Αναπληρωτής Καθηγητής Ε.Κ.Π.Α.

ΑΘΗΝΑ, 2014

Πρόλογος

Η παρούσα εργασία πραγματεύεται τη μελέτη της Δέλτα μεθόδου, μέσω της οποίας επιτυγχάνεται η προσέγγιση της συμπεριφοράς των δειγματικών κεντρικών ροπών ασυμπτωτικά.

Έπειτα από την παράθεση βασικών ορισμών και θεωρημάτων στο πρώτο κεφάλαιο, η οποία αποσκοπεί στη βέλτιστη κατανόηση των παρακάτω, στο δεύτερο κεφάλαιο αναφερόμαστε στο Θεώρημα Taylor. Το τελευταίο επιτρέπει την προσέγγιση μιας οποιαδήποτε συνάρτησης μέσω μιας πολυωνυμικής και χρησιμοποιείται στην απόδειξη της Δέλτα μεθόδου. Η μέθοδος αυτή αναλύεται ενδελεχώς στο επόμενο κεφάλαιο και αποτελεί μια γενίκευση του Κεντρικού Οριακού Θεωρήματος. Τέλος, στο τελευταίο κεφάλαιο, εφαρμόζοντας τη Δέλτα μέθοδο, μελετάμε διεξοδικά την οριακή συμπεριφορά των δειγματικών κεντρικών ροπών και πιο συγκεκριμένα, την οριακή τους κατανομή, τη σύγκλιση και την ανεξαρτησία τους από το δειγματικό μέσο.

Θα ήθελα να εκφράσω τις ειλικρινείς ευχαριστίες μου προς τους καθηγητές κ. Χρήστο Κουκουβίνο για τη δυνατότητα εκπόνησης της συγκεκριμένης εργασίας και κ. Νίκο Παπαδάτο για την ανεκτίμητη καθοδήγησή του και τις εποικοδομητικές παρατηρήσεις του, οι οποίες συνέβαλαν στην υπερπήδηση δυσχερειών που ανέκυψαν και υπήρξαν απαραίτητες για την περάτωσή της.

Περιεχόμενα

1 Βασικές Έννοιες Θεωρίας Πιθανοτήτων 5
1.1 Εισαγωγή 5
1.2 Σύγκλιση Τυχαίων Μεταβλητών 6
1.3 Στοιχεία Θεωρίας Μέτρου στη Θεωρία Πιθανοτήτων 10
1.4 Οριακά Θεωρήματα 13
1.5 Χρήσιμοι Ορισμοί-Θεωρήματα 22
1.6 Σχέσεις $O(.)$ και $o(.)$ 23

2 Σειρές Taylor 24
2.1 Ιστορική Αναδρομή 24
2.2 Πολυώνυμο Taylor 25
2.3 Θεώρημα Taylor 28
2.3.1 Μονοδιάστατη Μορφή Θεωρήματος Taylor 30
2.3.2 Πολυδιάστατη Μορφή Θεωρήματος Taylor 31

3 Μέθοδος Δέλτα 36
3.1 Μέθοδος Δέλτα $1^{ης}$ τάξης 37
3.2 Μέθοδος Δέλτα $2^{ης}$ τάξης 40
3.3 Πολυδιάστατη Μέθοδος Δέλτα ανώτερης τάξης 42

4 Οριακή συμπεριφορά δειγματικών κεντρικών ροπών 45
4.1 Συνέπεια δειγματικών κεντρικών ροπών 45
4.2 Οριακή κατανομή δειγματικών κεντρικών ροπών 49
4.3 Σύγκλιση δειγματικών κεντρικών ροπών 61
4.4 Ανεξαρτησία δειγματικών κεντρικών ροπών από το δειγματικό μέσο 71

Βιβλιογραφία 79

Κεφάλαιο 1

Βασικές Έννοιες Θεωρίας Πιθανοτήτων

1.1 Εισαγωγή

Η Θεωρία Πιθανοτήτων αποτελεί τον κλάδο των μαθηματικών που αναφέρεται στον προσδιορισμό της αβεβαιότητας ενός τυχαίου φαινομένου. Η πρώτη συστηματική μελέτη των πιθανοτήτων έχει καταγραφεί κατά τον 16^{o} αιώνα π.Χ. από τον Gerolamo Cardano[1], του οποίου το πάθος για τα τυχερά παιχνίδια γέννησε πλήθος βασικών κανόνων και πρώιμων συμπερασμάτων σχετικά με τις πιθανότητες. Στο βιβλίο του «Liber de Ludo Aleae », με κύριο παράδειγμα τη ρίψη ζαριού, ασχολήθηκε αποκλειστικά με ισοπίθανα γεγονότα και κατέληξε ότι η πιθανότητα να συμβεί ένα γεγονός είναι ο λόγος του αριθμού των ευνοϊκών αποτελεσμάτων προς τον ολικό αριθμό αποτελεσμάτων.

Η ουσιαστική εδραίωση της Θεωρίας Πιθανοτήτων στο μαθηματικό προσκήνιο πραγματοποιήθηκε στα μέσα του $17^{ου}$ αιώνα έπειτα από έκκληση του Γάλλου ευγενή και παίκτη τυχερών παιχνιδιών Chevalier de Méré (1607-1648), προς το μαθηματικό Blaise Pascal[2] με στόχο την επίλυση προβλημάτων τυχερών παιχνιδιών και πιο συγκεκριμένα το «Πρόβλημα των βαθμών».[3] Τα πρόβλημα αυτό υπήρξε εφαλτήριο ενασχόλησης τόσο του τελευταίου όσο και του μαθηματικού Pierre de Fermat[4] με προβλήματα προσδιορισμού αβεβαιότητας, ορίζοντας την

[1]Gerolamo/Geronimo Cardano (1501-1576): Ιταλός μαθηματικός, φυσικός, αστρονόμος γνωστός για το ευρέως φάσματος συγγραφικό του έργο πάνω στην ιατρική, τη φυσική, τη φιλοσοφία, τη θρησκεία, τη μουσική αλλά κυρίως την εξαγωγή κανόνων πιθανότητας.

[2]Blaise Pascal (1623–1662): Γάλλος μαθηματικός, φυσικός, φιλόσοφος. Υπήρξε παιδί θαύμα, αφού στα δεκαέξι του δημοσίευσε μια διατριβή για τις κωνικές τομές, ενώ στα δεκαοκτώ του εφηύρε μια υπολογιστική μηχανή.

[3]Το πρόβλημα καταμερισμού των μεριδίων, όταν το παιχνίδι διακόπτεται πριν τη λήξη του.

[4]Pierre de Fermat (1601-1665): Γάλλος ερασιτέχνης μαθηματικός ο οποίος ασχολήθηκε με τον

πρώτη ουσιαστική μαθηματική προσέγγιση της πιθανότητας.

1.2 Σύγκλιση Τυχαίων Μεταβλητών

Ορισμός 1.2.1. Έστω $(F_n)_{n\in\mathbb{N}}$ ακολουθία συναρτήσεων κατανομής πιθανότητας. Λέγεται ότι η ακολουθία αυτή *συγκλίνει ασθενώς* στη συνάρτηση κατανομής F και συμβολίζεται $F_n \xrightarrow{W} F$ αν και μόνο αν για κάθε $x \in \mathbb{R}$ σημείο ασυνέχειας της F ισχύει $\lim_{n\to+\infty} F_n(x) = F(x)$.

Ορισμός 1.2.2. Έστω $(X_n)_{n\in\mathbb{N}}$ ακολουθία τυχαίων μεταβλητών. Λέγεται ότι η ακολουθία αυτή *συγκλίνει κατά Νόμο (ή συγκλίνει ασθενώς ή συγκλίνει κατά κατανομή)* στην τυχαία μεταβλητή X και συμβολίζεται $X_n \xrightarrow{D} X$ όταν και μόνο όταν $F_{X_n} \xrightarrow{W} F_X$, όπου F_{X_n} η ακολουθία των συναρτήσεων κατανομής πιθανότητας των X_n και F_X η συνάρτηση κατανομής πιθανότητας της X ή ισοδύναμα όταν και μόνο όταν $\lim_{n\to+\infty} P(X_n \le x) = P(X \le x)$ $\forall x$ τ.ω. $P(X = x) = 0$.

Ορισμός 1.2.3. Έστω $(X_n)_{n\in\mathbb{N}}$ ακολουθία τυχαίων μεταβλητών. Θα λέμε ότι ότι η ακολουθία αυτή *συγκλίνει κατά πιθανότητα* στην τυχαία μεταβλητή X και γράφουμε $X_n \xrightarrow{P} X$ αν και μόνο αν για κάθε $\varepsilon > 0$

$$\lim_{n\to+\infty} P(|X_n - X| > \varepsilon) = 0$$

ή ισοδύναμα $\forall\varepsilon > 0$ $\lim_{n\to+\infty} P(|X_n - X| \le \varepsilon) = 1$.

Ορισμός 1.2.4. Έστω $(X_n)_{n\in\mathbb{N}}$ ακολουθία τυχαίων μεταβλητών. Θα λέμε ότι ότι η ακολουθία αυτή *συγκλίνει με πιθανότητα 1 (ή συγκλίνει σχεδόν βεβαίως ή συγκλίνει σχεδόν παντού)* στην τυχαία μεταβλητή X και γράφουμε $X_n \xrightarrow{a.s} X$ αν και μόνο αν

$$P(\lim_{n\to+\infty} |X_n - X| = 0) = 1.$$
$$\Leftrightarrow \lim_{m\to\infty} P[|X_n - X| \ge \varepsilon, \forall n \ge m] = 1, \qquad \forall\varepsilon > 0$$
$$\Leftrightarrow \lim_{n\to\infty} P[\sup_{k\ge n} |X_k - X| \ge \varepsilon] = 0, \quad \forall\varepsilon > 0.$$

απειροστικό λογισμό, τη θεωρία των αριθμών, την αναλυτική γεωμετρία, τη θεωρία πιθανοτήτων και την οπτική. Ειδικότερα, είναι γνωστός για την ανακάλυψη μιας μεθόδου υπολογισμού των ελάχιστων και μέγιστων σημείων σε καμπύλες γραμμές, καθώς και για το τελευταίο θεώρημα του Φερμά.

Ορισμός 1.2.5. Έστω $(X_n)_{n\in\mathbb{N}}$ ακολουθία τυχαίων μεταβλητών. Θα λέμε ότι ότι η ακολουθία αυτή *συγκλίνει κατά μέσον p-τάξης (p≥0)* στην τυχαία μεταβλητή X και γράφουμε $X_n \xrightarrow{L^p} X$ όταν $E[|X_n|^p] < \infty$, για κάθε $n \in \mathbb{N}$, $E[|X|^p] < \infty$ και $\lim\limits_{n\to+\infty} E[|X_n - X|^p] = 0$.
Ειδικότερα, για $p = 2$, λέμε ότι η ακολουθία $(X_n)_{n\in\mathbb{N}}$ *συγκλίνει κατά τετραγωνικό μέσο* και συμβολίζουμε $X_n \xrightarrow{L^2} X$ ή $X_n \xrightarrow{s.m} X$.

Θα αποδείξουμε παρακάτω ότι η σύγκλιση κατά πιθανότητα συνεπάγεται την ασθενή σύγκλιση, ενώ η σχεδόν βέβαιη σύγκλιση και η σύγκλιση κατά μέσον συνεπάγονται τη σύγκλιση κατά πιθανότητα.

Θεώρημα 1.2.1. *Έστω* $X_n \xrightarrow{P} X$*. Τότε* $X_n \xrightarrow{D} X$.

Απόδειξη. Έστω $F_n, n \in 1, 2, ...$ οι συναρτήσεις κατανομής πιθανότητας των τυχαίων μεταβλητών X_n και F η συνάρτηση κατανομής πιθανότητας της τυχαίας μεταβλητής X. Θεωρούμε ένα σημείο συνέχειας της F.
Αρκεί ν.δ.ο. $\lim\limits_{n\to+\infty} F_n(x) = F(x)$. Έστω τυχαίο $\varepsilon > 0$, ισχύει

$$\begin{aligned}\{\omega : X(\omega) \leq x - \varepsilon\} &= \{\omega : X(\omega) \leq x - \varepsilon\} \cap (\{\omega : X_n(\omega) \leq x\} \cup \\ &\quad \{\omega : X_n(\omega) > x\}) \\ &= (\{\omega : X(\omega) \leq x - \varepsilon\} \cap \{\omega : X_n(\omega) \leq x\}) \cup \\ &\quad (\{\omega : X(\omega) \leq x - \varepsilon\} \cap \{\omega : X_n(\omega) > x\})\end{aligned}$$

Όμως, $X(\omega) \leq x - \varepsilon$ και $X_n(\omega) > x$ έπεται ότι $X_n(\omega) - X(\omega) > \varepsilon$ και $X(\omega) - X_n(\omega) < -\varepsilon$ άρα $|X_n(\omega) - X(\omega)| > \varepsilon$.
Δηλαδή, $\{\omega : X(\omega) \leq x - \varepsilon\} \cap \{\omega : X_n(\omega) > x\} \subset \{\omega : |X_n(\omega) - X(\omega)| > \varepsilon\}$.
Ακόμα, $\{\omega : X(\omega) \leq x - \varepsilon\} \cap \{\omega : X_n(\omega) \leq x\} \subset \{\omega : |X_n(\omega) \leq x|$. Επομένως, καταλήγουμε ότι $\{X \leq x - \varepsilon\} \subset \{X_n \leq x\} \cup \{|X_n - X| > \varepsilon\}$
και άρα $P(X \leq x - \varepsilon) \leq P(X_n \leq x) + P(|X_n - X| > \varepsilon)$, που σημαίνει ότι

$$F(x - \varepsilon) \leq F_n(x) + P(|X_n - X| > \varepsilon). \tag{1.2.1}$$

Ακόμα, έχουμε

$$\begin{aligned}\{\omega : X(\omega) > x + \varepsilon\} &= \{\omega : X(\omega) > x + \varepsilon\} \cap (\{\omega : X_n(\omega) \leq x\} \cup \\ &\quad \{\omega : X_n(\omega) > x\}) \\ &= (\{\omega : X(\omega) > x + \varepsilon\} \cap \{\omega : X_n(\omega) \leq x\}) \cup \\ &\quad (\{\omega : X(\omega) > x + \varepsilon\} \cap \{\omega : X_n(\omega) > x\}) \\ &\subset \{\omega : |X_n(\omega) - X(\omega)| > \varepsilon\} \cup \{\omega : X_n(\omega) > x\}.\end{aligned}$$

Συνεπώς, $\{X > x + \varepsilon\} \subset \{|X_n - X| > \varepsilon\} \cup \{X_n > x\}$, δηλαδή, $P(X_n > x + \varepsilon) \leq P(|X_n - X| > \varepsilon) + P(X_n > x)$ και άρα

$$1 - F(x + \varepsilon) \leq P(|X_n - X| > \varepsilon) + 1 - F_n(x) \tag{1.2.2}$$

Από τις σχέσεις (1.2.1), (1.2.2) συμπεραίνουμε ότι
$F(x - \varepsilon) - P(|X_n - X| > \varepsilon) \leq F_n(x) \leq F(x + \varepsilon) + P(|X_n - X| > \varepsilon)$.
Και επειδή από υπόθεση $X_n \xrightarrow{P} X$ συνεπάγεται ότι $\lim_{n\to+\infty} P(|X_n - X| > \varepsilon) = 0$.
Άρα $F(x-\varepsilon) \leq \lim_{n\to+\infty} infF_n(x) \leq \lim_{n\to+\infty} supF_n(x) \leq F(x+\varepsilon)$ για τυχαίο ε>0.
Άρα καταλήγουμε ότι $\lim_{n\to+\infty} F_n(x) = F(x), \forall\varepsilon > 0$.

□

Για την απόδειξη του παρακάτω Θεωρήματος θα χρειαστούμε το εξής Λήμμα:

Λήμμα 1.2.1. *Θεωρούμε $(A_n)_{n\in\mathbb{N}}$ ακολουθία ενδεχομένων.Αν $(A_n)_{n\in\mathbb{N}}$ είναι αύξουσα, δηλαδή $A_n \subset A_{n+1}, \forall n \in \mathbb{N}$, τότε $\lim_{n\to+\infty} P[A_n] = P[\bigcup_{n=1}^{\infty} A_n]$.*

Απόδειξη. Έχουμε $A_n \subset A_{n+1}, \forall n \in \mathbb{N} \Rightarrow P(A_n) \leq P(A_{n+1}), \forall n \in \mathbb{N}$, άρα το όριο $\lim_{n\to+\infty} P[A_n]$ υπάρχει στο $\mathbb{R}$. Θέτοντας $B_1 = A_1, B_n = A_n - A_{n-1}$
για $n \in \{2, 3, ...\}$, παρατηρούμε ότι
$\bigcup_{i=1}^{n} B_i = A_1 \cup A_2 - A_1 \cup A_3 - A_2 \cup ... \cup A_n - A_{n-1} = A_n$, γιατί $A_n \subset A_{n+1}$
$\bigcup_{i=1}^{\infty} B_i = A_1 \cup A_2 - A_1 \cup A_3 - A_2 \cup ... \cup A_n - A_{n-1} \cup ... = A_1 \cup A_2 ... = \bigcup_{i=1}^{\infty} A_i$
$B_i \cap B_j = \emptyset, \forall i \neq j$. Συμπεραίνουμε ότι
$P[\bigcup_{i=1}^{\infty} A_i] = P[\bigcup_{i=1}^{\infty} B_i] = \sum_{i=1}^{\infty} P[B_i] = \lim_{n\to\infty} \sum_{i=1}^{n} P[B_i] = \lim_{n\to\infty} P[\bigcup_{i=1}^{n} B_i] = \lim_{n\to\infty} P[A_n]$.

□

Θεώρημα 1.2.2. *Έστω $X_n \xrightarrow{a.s} X$. Τότε $X_n \xrightarrow{P} X$.*

Απόδειξη. Από υπόθεση $X_n \xrightarrow{a.s} X$ άρα από ορισμό σχεδόν βέβαιης σύγκλισης συνεπάγεται ότι $P(\lim_{n\to+\infty} |X_n - X| = 0) = 1$, άρα
για κάθε $\nu \in \mathbb{N}, \exists n_0 \in \mathbb{N}$ τ.ω. $\forall n \geq n_0$ ισχύει $|X_n - X| < \frac{1}{\nu}$.
Δηλαδή, το ενδεχόμενο $\{\lim_{n\to+\infty} X_n = X\}$ ισοδυναμεί με το ενδεχόμενο
$\bigcap_{\nu} \bigcup_{n_0} \bigcap_{n\geq n_0}^{\infty} \{|X_n - X| < \frac{1}{\nu}\}$.

Επομένως, $X_n \xrightarrow{a.s} X \Leftrightarrow P[\bigcap\limits_{\nu}\bigcup\limits_{n_0}\bigcap\limits_{n\geq n_0}^{\infty}\{|X_n - X| < \frac{1}{\nu}\}] = 1.$

Όμως, $\forall \nu \in \mathbb{N}, \quad \bigcap\limits_{\nu}\bigcup\limits_{n_0}\bigcap\limits_{n\geq n_0}^{\infty}\{|X_n - X| < \frac{1}{\nu}\} \subset \bigcup\limits_{n_0}\bigcap\limits_{n\geq n_0}^{\infty}\{|X_n - X| < \frac{1}{\nu}\}.$

Άρα ισχύει $P[\bigcup\limits_{n_0}\bigcap\limits_{n\geq n_0}^{\infty}\{|X_n - X| < \frac{1}{\nu}\}] = 1, \forall \nu \in \mathbb{N}$. Από Λήμμα 1.2.1. και αφού η ακολουθία ενδεχομένων $\bigcap\limits_{n\geq n_0}^{\infty}\{|X_n - X| < \frac{1}{\nu}\}$ είναι αύξουσα καταλήγουμε ότι $\lim\limits_{n_0\to\infty} P[\bigcap\limits_{n\geq n_0}^{\infty}\{|X_n - X| < \dfrac{1}{\nu}\}] = 1.$

Ακόμα, $\bigcap\limits_{n\geq n_0}^{\infty}\{|X_n - X| < \frac{1}{\nu}\} \subset \{|X_{n_0} - X| < \frac{1}{\nu}\}$

άρα $\lim\limits_{n_0\to\infty} P[|X_{n_0} - X|] = 1, \forall \nu \in \mathbb{N} \Rightarrow X_n \xrightarrow{P} X.$ □

Θεώρημα 1.2.3. *Έστω* $X_n \xrightarrow{L^p} X$. *Τότε* $X_n \xrightarrow{P} X$.

Απόδειξη. Για την απόδειξη θα κάνουμε χρήση της Ανισότητας Chebychev[5], την οποία παραθέτουμε:

Λήμμα 1.2.2. ***Ανισότητα Chebychev:*** *Για κάθε τυχαία μεταβλητή* X *και για τυχαίες σταθερές* $\alpha > 0$ *και* c, *ισχύει*

$$E[(X-c)^2] \geq \alpha^2 P[|X-c| \geq \alpha]. \tag{1.2.3}$$

Απόδειξη. Θα αποδείξουμε τη συνεχή περίπτωση. Υποθέτουμε ότι X συνεχής τ.μ με σ.π.π. f. Τότε, έχουμε:

$$\begin{aligned} E[(X-c)^2] &= \int_{-\infty}^{+\infty}(t-c)^2 f(t)dt \geq \int_{-\infty}^{c-\alpha}(t-c)^2 f(t)dt + \int_{c+\alpha}^{+\infty}(t-c)^2 f(t)dt \\ &\geq \int_{-\infty}^{c-\alpha}\alpha^2 f(t)dt + \int_{c+\alpha}^{+\infty}\alpha^2 f(t)dt, \\ &\quad \text{διότι } t \leq c-\alpha \Rightarrow \alpha \leq |t-c| \Rightarrow \alpha^2 \leq |t-c|^2. \\ &= \alpha^2[\int_{-\infty}^{c-\alpha} f(t)dt + \int_{c+\alpha}^{+\infty} f(t)dt] \end{aligned}$$

[5]Μια πρώτη μορφή της ανισότητας και η απόδειξή της εμφανίστηκε από τον Γάλλο στατιστικό Irénée-Jules Bienaymé (1796-1878), όμως φέρει το όνομα του Ρώσου Μαθηματικού Pafnuty Chebyshev (1821- 1894), ο οποίος επίσης την απέδειξε.

$$= \alpha^2 P[X \leq c - \alpha \text{ ή } X \geq c + \alpha]$$
$$= \alpha^2 P[|X - c| \geq \alpha]. \qquad \square$$

Συνέχεια Απόδειξης Θεωρήματος.

Από την Ανισότητα Chebychev $E[(X_n - X)^p] \geq \alpha^p P[|X_n - X| > \alpha]$ όμως αφού $X_n \xrightarrow{L^p} X$ εξ' ορισμού $\lim\limits_{n\to+\infty} E[|X_n - X|^p] = 0$.Επομένως, από Κριτήριο Παρεμβολής προκύπτει $\lim\limits_{n\to+\infty} P(|X_n - X| > \alpha) = 0$ και άρα $X_n \xrightarrow{P} X$. $\square$

1.3 Στοιχεία Θεωρίας Μέτρου στη Θεωρία Πιθανοτήτων

Ορισμός 1.3.1. Θεωρούμε ένα σύνολο $\Omega \neq \emptyset$. Μια κλάση F υποσυνόλων του Ω λέγεται *σ-άλγεβρα* εφόσον ικανοποιεί τις παρακάτω συνθήκες:

- $\Omega \in F$
- αν $A \in F \Rightarrow A^c \in F$
- αν $\{A_n : n = 1, 2, 3\} \subset F \Rightarrow \bigcup\limits_{n=1}^{\infty} A_n \in F$

Πρόταση 1.3.1. *Αν $(F_i)_{i\in I}$ μια οικογένεια σ-αλγεβρών υποσυνόλων του Β, τότε κλάση $F = \bigcap\limits_{i\in I} F_i$ είναι σ-άλγεβρα υποσυνόλων του Ω.*

Απόδειξη. Θα αποδείξουμε ότι για την κλάση F ικανοποιούνται οι τρεις προϋποθέσεις της σ-άλγεβρας από τον ορισμό. Εφόσον F_i σ-άλγεβρες από υπόθεση άρα $\Omega \in F_i, \forall i \in I \Rightarrow \Omega \in \bigcap\limits_{i\in I} \Rightarrow \Omega \in F$. Αν $A \in F \Rightarrow A \in \bigcap\limits_{i\in I} F_i \Rightarrow A \in F_i, \forall i \in I$ και από υπόθεση F_i σ-άλγεβρες οπότε$A^c \in F_i, \forall i \in I \Rightarrow A^c \in \bigcap\limits_{i\in I} F_i \Rightarrow A^c \in F$. Τέλος, αν $\{A_n\}_{n=1,2,..} \subset F \Rightarrow \{A_n\}_{n=1,2,..} \subset F_i, \forall i \in I \Rightarrow \bigcup\limits_{n=1}^{\infty} A_n \in F_i, \forall i \in I \Rightarrow \bigcup\limits_{n=1}^{\infty} A_n \in F$. $\square$

Ορισμός 1.3.2. Έστω $C \neq \emptyset$ μια κλάση υποσυνόλων του Ω και F η τομή όλων των σ-αλγεβρών που περιέχουν την C. Η F είναι όπως αποδείξαμε στην προηγούμενη πρόταση σ-άλγεβρα ως τομή σ-αλγεβρών, ονομάζεται *σ-άλγεβρα παραγόμενη από την C* και συμβολίζεται $F = \sigma(C)$.

Παρατήρηση 1.3.1. Παρατηρούμε ότι η σ-άλγεβρα παραγόμενη από την C είναι η ελάχιστη σ-άλγεβρα που περιέχει την C.

Ορισμός 1.3.3. Ονομάζουμε *Borel σ-άλγεβρα*, τη σ-άλγεβρα που παράγεται από τα ανοικτά υποσύνολα του $\mathbb{R}$ και τη συμβολίζουμε $\mathfrak{B}$,ενώ τα στοιχεία της λέγονται *σύνολα Borel*.

Πρόταση 1.3.2. Η Borel σ-άλγεβρα $\mathfrak{B}$ παράγεται από:

- Τα ανοικτά διαστήματα $\mathfrak{E}_1 = \{(a,b) : a < b\}$
- Τα κλειστά διαστήματα $\mathfrak{E}_2 = \{[a,b] : a < b\}$
- Τα διαστήματα $\mathfrak{E}_{3,1} = \{(a,b] : a < b\}$ ή $\mathfrak{E}_{3,2} = \{[a,b) : a < b\}$
- Τα διαστήματα $\mathfrak{E}_{4,1} = \{(a,\infty) : a \in \mathbb{R}\}$ ή $\mathfrak{E}_{4,2} = \{(-\infty,b) : b \in \mathbb{R}\}$
- Τα διαστήματα $\mathfrak{E}_{5,1} = \{[a,\infty) : a \in \mathbb{R}\}$ ή $\mathfrak{E}_{5,2} = \{(-\infty,b] : b \in \mathbb{R}\}$

Ορισμός 1.3.4. Ονομάζουμε χώρο πιθανότητας την τριάδα (Ω,F,P), όπου $\Omega \neq \emptyset$, F σ-άλγεβρα υποσυνόλων του Ω, P:$F \to \mathbb{R}$ μέτρο πιθανότητας τέτοιο ώστε να ικανοποιεί τα παρακάτω:

- $P(A) \geq 0, \forall A \in F$
- P(Ω)=1
- Αν $\{A_n : n = 1,2,...\} \subset F$ με $A_n \cap A_m = \emptyset \forall n \neq m$ τότε ισχύει $P[\bigcup_{n=1}^{\infty} A_n] = \sum_{n=1}^{\infty} P[A_n]$

Πρόταση 1.3.3. *Έστω (Ω,F,P) χώρος πιθανότητας και* $\{A_n : n = 1,2,...\} \subset F$.
i) Αν $A_n \subset A_{n+1}, \forall n \in \mathbb{N}$ τότε $P[\bigcup_{n=1}^{\infty} A_n] = \lim_{n\to\infty} P[A_n]$.
ii)Αν Αν $A_n \supset A_{n+1}, \forall n \in \mathbb{N}$ τότε $P[\bigcap_{n=1}^{\infty} A_n] = \lim_{n\to\infty} P[A_n]$.

Απόδειξη. i)Αφού $A_n \subset A_{n+1}, \forall n \in \mathbb{N}$, τότε $P[A_n] \leq P[A_{n+1}] \leq P[\Omega] = 1$ άρα το όριο $\lim_{n\to\infty} P[A_n]$ υπάρχει στο $\mathbb{R}, \forall n \in \mathbb{N}$.
Θέτουμε $B_1 = A_1, B_n = A_n - A_{n-1}$,για $n = 2,3,...$. Τότε,προκύπτει ότι:
$\bigcup_{i=1}^{n} B_i = B_1 \cup B_2 \cup ... \cup B_n = A_1 \cup \{A_2 - A_1\}, ... \{A_n - A_{n-1}\}$
$\overset{A_n \subset A_{n+1}}{=} A_1 \cup A_2 \cup ... \cup A_n = A_n$.
$\bigcup_{n=1}^{\infty} B_n = B_1 \cup B_2 \cup ... \cup B_n \cup ... = A_1 \cup \{A_2 - A_1\}, ... \{A_n - A_{n-1}\} \cup ...$

$$\overset{A_n \subset A_{n+1}}{=} A_1 \cup A_2 \cup ... \cup A_n \cup A_{n+1} \cup ... = \bigcup_{n=1}^{\infty} A_n.$$

$\forall i \neq j, B_i \cap B_j = \{A_i - A_{i-1}\} \cap \{A_j - A_{j-1}\} \overset{i<j \Rightarrow A_i \subseteq A_{j-1}}{=} \emptyset$[6]

Θα δείξουμε ότι $P[\bigcup_{n=1}^{\infty} A_n] = \lim_{n\to\infty} P[A_n]$.

$P[\bigcup_{n=1}^{\infty} A_n] = P[\bigcup_{n=1}^{\infty} B_n] \overset{Op.1.3.4.}{=} \sum_{n=1}^{\infty} P[B_n] = \lim_{k\to\infty} \sum_{n=1}^{k} P[B_n] = \lim_{k\to\infty} P[A_k]$.

ii)Αντιστοίχως, θέτουμε $C_n = A_n^c$, για $n \in \mathbb{N}$.

Τότε, $\forall n \in \mathbb{N}, A_n \supset A_{n+1} \Rightarrow A_n^c \subset A_{n+1}^c \Rightarrow C_n \subset C_{n+1}$. Οπότε από ερώτημα (i) ισχύει ότι $P[\bigcup_{n=1}^{\infty} C_n] = \lim_{n\to\infty} P[C_n] \Longrightarrow 1 - P[\{\bigcup_{n=1}^{\infty} C_n\}^c] = \lim_{n\to\infty}(1 - P[C_n^c])$

$$\Longrightarrow 1 - P[\bigcap_{n=1}^{\infty} A_n] = \lim_{n\to\infty}(1 - P[A_n])$$

$$\Longrightarrow P[\bigcap_{n=1}^{\infty} A_n] = \lim_{n\to\infty} P[A_n],$$

διότι ορίσαμε $C_n = A_n^c$ και άρα $[\bigcup_{n=1}^{\infty} C_n]^c = \bigcap_{n=1}^{\infty} C_n^c = \bigcap_{n=1}^{\infty} A_n$.

□

Πόρισμα 1.3.1. (Ανισότητα Boole[7]) *Έστω* $\{A_n : n = 1, 2, ...\} \subset F$. *Τότε, ισχύει* $P[\bigcup_{n=1}^{\infty} A_n] \leq \sum_{n=1}^{\infty} P[A_n]$.

Απόδειξη. [8] Θέτουμε $B_k = \bigcup_{n=1}^{k} A_n$. Τότε, παρατηρούμε ότι $B_k \subset B_{k+1}$, διότι $A_1 \subset A_1 \cup A_2 \subset ... \subset \bigcup_{n=1}^{k} A_n \subset \bigcup_{n=1}^{k+1} A_n$. Άρα έχουμε,

$$P[\bigcup_{n=1}^{\infty} A_n] = P[\bigcup_{k=1}^{\infty}\{\bigcup_{n=1}^{k} A_n\}] = P[\bigcup_{k=1}^{\infty} B_k] \overset{\Pi\rho o\tau\alpha\sigma\eta 1.3.3(i)}{=} \lim_{k\to\infty} P[B_k] = \lim_{k\to\infty} P[\bigcup_{n=1}^{k} A_n]$$

$$\overset{P[\bigcup_{i=1}^{n} A_i] \leq \sum_{i=1}^{n} P[A_i]}{\leq} \lim_{k\to\infty} \sum_{n=1}^{k} P[A_n] = \sum_{n=1}^{\infty} P[A_n].$$

□

[6] αντίστοιχα αν $j < i$.

[7] George Boole (1815–1864): Άγγλος μαθηματικός, φιλόσοφος του οποίου η ενεργητικότητα διαφαίνεται σε διάφορους κλάδους των μαθηματικών, όπως οι διαφορικές εξισώσεις, η θεωρία πιθανοτήτων, η αλγεβρική λογική ενώ θεωρείται πατέρας της πληροφορικής.

[8] Η ανισότητα μπορεί να αποδειχθεί και με χρήση της μαθηματικής επαγωγής.

Ορισμός 1.3.5. Το ανώτερο και το κατώτερο όριο μιας ακολουθίας ενδεχομένων $\{A_n : n = 1, 2, ...\}$ ορίζονται ως εξής:

$$\lim_{n\to\infty} supA_n := \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} A_m$$

$$\lim_{n\to\infty} infA_n := \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m$$

Πρόταση 1.3.4. (Λήμμα Borel[9]-Cantelli[10]) *Έστω ο χώρος πιθανότητας (Ω,F,P) και* $\{A_n : n = 1, 2, ...\} \subset F$. *Εάν* $\sum\limits_{n=1}^{\infty} P[A_n] < \infty \Longrightarrow P[\lim\limits_{n\to\infty} \sup A_n] = 0.$

Απόδειξη. Έχουμε, $\bigcup\limits_{n=1}^{\infty} \bigcap\limits_{m=n}^{\infty} A_m \subset \bigcap\limits_{m=n}^{\infty} A_m$. (1) Άρα,

$$P[\lim_{n\to\infty} \sup A_n] = P[\bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} A_m] \overset{(1)}{\leq} P[\bigcup_{m=n}^{\infty} A_m] \overset{\Pi o\rho\iota\sigma\mu\alpha 1.3.1}{\leq} \sum_{m=n}^{\infty} P[A_m]. \quad (2)$$

Από υπόθεση $\sum\limits_{n=1}^{\infty} P[A_n] < \infty$, άρα ορίζεται στο $\mathbb{R}$ το όριο της σειράς $\sum_{m=n}^{\infty} P[A_m]$ και παρατηρούμε ότι $\lim\limits_{n\to\infty} \sum_{m=n}^{\infty} P[A_m] = 0$. Οπότε, χρησιμοποιώντας την (2) καταλήγουμε ότι $P[\lim\limits_{n\to\infty} \sup A_n] = 0$.

□

Ορισμός 1.3.5. Μια συνάρτηση $g : \mathbb{R}^m \to \mathbb{R}^n$ λέγεται *συνάρτηση Borel* αν $\forall B \in \mathfrak{B}^n$ ισχύει $g^{-1}(B) \in \mathfrak{B}^m$.

1.4 Οριακά Θεωρήματα

Τα Οριακά Θεωρήματα προσδιορίζουν τη συμπεριφορά Borel συναρτήσεων, ενός συνόλου τυχαίων μεταβλητών $(X_1, X_2, ... X_n)$ που ανήκουν στον ίδιο χώρο πιθανότητας (Ω,F,P) όταν το n (πλήθος τυχαίων μεταβλητών X_i) τείνει στο άπειρο. Ανάλογα με το είδος της σύγκλισης χωρίζονται σε τρεις κατηγορίες: (Α.Ν.Μ.Α), (Ι.Ν.Μ.Α), (Κ.Ο.Θ).

Το πρώτο Οριακό Θεώρημα διατυπώθηκε από τον Johanne Bernoulli[11] το 1713. Ο Bernoulli το ονόμασε «Golden Theorem» (Χρυσό Θεώρημα) αποσκοπώντας

[9] Félix Édouard Justin Émile Borel (1871-1956): Γάλλος μαθηματικός και πολιτικός. Υπήρξε ένας από τους πρωτοπόρους που ασχολήθηκαν ενδελεχώς με τη θεωρία μέτρου και την εφαρμογή της στη θεωρία πιθανοτήτων.

[10] Francesco Paolo Cantelli (1875-1966): Ιταλός μαθηματικός, το ενδιαφέρον του οποίου προσανατολίστηκε στην αστρονομία, την ουράνια μηχανική και την εφαρμογή των πιθανοτήτων στα οικονομικά.

[11] βλ. Κεφ. 2 υποσημείωση 3.

στο να τονίσει τη σημαντικότητά του. Η ονομασία Ισχυρός Νόμος Μεγάλων Αριθμών δόθηκε μεταγενέστερα, από τον μαθηματικό Siméon Denis Poisson [12].

Ασθενείς Νόμοι Μεγάλων Αριθμών (Α.Ν.Μ.Α.)

Θεώρημα 1.4.1. ***(Bernoulli)*** *Θεωρούμε ακολουθία ανεξάρτητων τυχαίων μεταβλητών* $\{X_n : n = 1, 2, ...\}$*, οι οποίες ακολουθούν κατανομή Bernoulli με* $P[X_n = 1] = p, P[X_n = 0] = 1 - p, n \in \mathbb{N}$. *Έστω* $S_n = X_1 + X_2 + \dots X_n$ *και άρα* $\frac{S_n}{n}$ *η σχετική συχνότητα επιτυχιών. Τότε* $\frac{S_n}{n} \xrightarrow{P} p$.

Απόδειξη. Για την απόδειξη θα χρησιμοποιήσουμε το επόμενο Λήμμα, το οποίο αποτελεί Θεώρημα Α.Ν.Μ.Α.

Λήμμα 1.4.1.(Chebyshev[13]) Θεωρούμε ακολουθία ασυσχέτιστων τυχαίων μεταβλητών $\{X_n : n = 1, 2, ...\}$ με πεπερασμένες διασπορές $\sigma_i^2 = Var[X_i] < \infty$ και $\lim\limits_{n\to\infty} \frac{1}{n^2} \sum\limits_{i=1}^{n} \sigma_i^2 = 0$.

Έστω $S_n = X_1 + X_2 + \dots X_n, n \in \mathbb{N}$. Τότε, $\frac{S_n - E[S_n]}{n} \xrightarrow{P} 0$.

Απόδειξη. (Λήμματος) Από υπόθεση οι $X_i, i = 1, 2, \dots$ είναι ασυσχέτιστες, δηλαδή $Cov[X_i, X_j] = 0, \forall i \neq j$ και από τη γνωστή ιδιότητα διασποράς $Var[X + Y] = Var[X] + Var[Y] + 2Cov[X, Y]$ έχουμε:

$$Var[\tfrac{S_n}{n}] = Var[\tfrac{1}{n} \sum_{i=1}^{n} X_i] = \tfrac{1}{n^2} Var[\sum_{i=1}^{n} X_i] = \tfrac{1}{n^2} \sum_{i=1}^{n} Var[X_i] = \tfrac{1}{n^2} \sum_{i=1}^{n} \sigma_i^2 .(1)$$

Από Ανισότητα Chebyshev (βλ. Λήμμα 1.2.2.) ισχύει ότι:
$E[(X - c)^2] \geq \varepsilon^2 P[|X - c| \geq \varepsilon]$. Άρα έχουμε $\forall \varepsilon > 0$:

$$P[|\tfrac{S_n - E[S_n]}{n}| > \varepsilon] = P[|\tfrac{S_n}{n} - \tfrac{E[S_n]}{n}| > \varepsilon] = P[|\tfrac{S_n}{n} - E[\tfrac{S_n}{n}]| > \varepsilon]$$

$$\leq \frac{E[(\frac{S_n}{n} - E[\frac{S_n}{n}])^2]}{\varepsilon^2} = \frac{Var[\frac{S_n}{n}]}{\varepsilon^2} \overset{(1)}{=} \frac{\frac{1}{n^2} \sum_{i=1}^{n} \sigma_i^2}{\varepsilon^2} \longrightarrow 0,$$ καθώς

$n \to \infty$ διότι από υπόθεση $\lim\limits_{n\to\infty} \frac{1}{n^2} \sum\limits_{i=1}^{n} \sigma_i^2 = 0$. Άρα από ορισμό της σύγκλισης κατά πιθανότητα, $\frac{S_n - E[S_n]}{n} \xrightarrow{P} 0$.

□

[12] Siméon Denis Poisson (1781-1840): Γάλλος μαθηματικός, γεωμέτρης, φυσικός, μέλος της Σουηδικής Ακαδημίας Επιστημών. Η συνεισφορά του στην επιστήμη είναι αναμφίβολη, ενδεικτικά εισήγαγε την κατανομή Poisson και την εξίσωση Poisson (η οποία προέκυψε από τη διόρθωση μιας εξίσωσης του Laplace), ενώ διαψεύστηκε ως προασπιστής της μη κυματικής φύσης του φωτός.

[13] Pafnuty Lvovich Chebyshev (1821-1894): Ρώσος μαθηματικός, ιδιαίτερα δραστήριος σε κλάδους των μαθηματικών όπως η θεωρία πιθανοτήτων, η στατιστική, η μηχανική, η θεωρία αριθμών, γνωστός για την ανισότητα Chebyshev, η οποία φέρει το όνομά του και χρησιμοποιείται στην απόδειξη των Α.Ν.Μ.Α.

Θα αποδείξουμε το Θεώρημα 1.4.1. :

Απόδειξη Θεωρήματος 1.4.1. Εφόσον οι τυχαίες μεταβλητές $X_i \sim B(p), i \in \mathbb{N}$, άρα $E[X_n] = p$ και $Var[X_n] = p(1-p)$.

Οπότε, $E[S_n] = E[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} E[X_i] = np$. (1)

Μπορούμε να εφαρμόσουμε το Λήμμα 1.4.1. διότι

$\sigma_i^2 = Var[X_i] = p(1-p) \le \frac{1}{4} < \infty$ (γιατί $0 \le p \le 1$). Άρα

$\sigma_i^2 = Var[X_i] < \infty$ και $\lim_{n\to\infty} \frac{1}{n^2} \sum_{i=1}^{n} \sigma_i^2 = 0$. Οπότε από Λήμμα 1.4.1. προκύπτει

$\frac{S_n - E[S_n]}{n} \xrightarrow{P} 0 \overset{(1)}{\Longrightarrow} \frac{S_n - np}{n} \xrightarrow{P} 0 \Longrightarrow \frac{S_n}{n} \xrightarrow{P} p$.

□

Θεώρημα 1.4.2. ***(Klinchin[14])*** *Θεωρούμε ακολουθία ανεξάρτητων και ισόνομων τυχαίων μεταβλητών $\{X_n : n = 1, 2, ...\}$ με πεπερασμένη μέση τιμή μ και διασπορά $\sigma^2 < \infty$. Έστω $S_n = X_1 + X_2 + ... X_n$. Τότε $\frac{S_n}{n} \xrightarrow{P} \mu$.*

Απόδειξη. Από τον ορισμό της κατά πιθανότητα σύγκλισης,

θα δείξουμε ότι αν $E[(X_n - \mu)]^2 \longrightarrow 0 \Longrightarrow X_n \xrightarrow{P} \mu$.

Από ανισότητα Chebychev (Λήμμα 1.2.2.) ισχύει ότι $\forall \alpha > 0$,

$E[(X_n - \mu)^2] \ge \alpha^2 P[|X_n - \mu| > \alpha]$. Οπότε, για $n \to \infty$

αν $E[(X_n - \mu)^2] \longrightarrow 0 \Longrightarrow P[|X_n - \mu| > \alpha] \longrightarrow 0 \Longrightarrow X_n \xrightarrow{P} \mu$.

Άρα αρκεί να δείξουμε ότι $E[(\frac{S_n}{n} - \mu)^2] \longrightarrow 0$καθώς $n \to \infty$.

Έχουμε $E[(\frac{S_n}{n} - \mu)^2] = Var[\frac{S_n}{n}] = \frac{1}{n^2} n\sigma^2 = \frac{\sigma^2}{n} \longrightarrow 0$ καθώς $n \to \infty$.

□

Ισχυροί Νόμοι Μεγάλων Αριθμών(Ι.Ν.Μ.Α.)

Θεώρημα 1.4.3. ***(Kolmogorov [15]για ανεξάρτητες τ.μ)*** *Θεωρούμε ακολουθία ανεξάρτητων τυχαίων μεταβλητών $\{X_i : i = 1, 2, ...\}$ με μέσες τιμές $\mu_i = E[X_i]$ και διασπορές $\sigma_i^2 = Var[X_i]$ πεπερασμένες τέτοιες ώστε να ισχύει $\sum_{i=1}^{\infty} \frac{\sigma_i^2}{i^2} < \infty$. Έστω $S_n = X_1 + X_2 + ... + X_n$. Τότε $\frac{S_n - E[S_n]}{n} \xrightarrow{a.s} 0$.*

[14]Aleksandr Yakovlevich Khinchin(1894–1959): Ρώσος μαθηματικός, ο οποίος καταπιάστηκε με τομείς των μαθηματικών όπως η θεωρία πιθανοτήτων, η πραγματική ανάλυση, η θεωρία μέτρου, η θεωρία αριθμών και η στατιστική φυσική.

[15]Andrey Nikolaevich Kolmogorov(1903-1987): Ρώσος μαθηματικός, ο οποίος δραστηριοποιήθηκε σε κλάδους τον μαθηματικών όπως η τοπολογία, η θεωρία πιθανοτήτων, η κλασική μηχανική, ενώ εισήγαγε την θεωρία πολυπλοκότητας.

Απόδειξη. Για την παρακάτω απόδειξη θα χρησιμοποιήσουμε την ανισότητα Hajek-Renyi[16] Έστω $X_i, i = 1, \dots, n$ ανεξάρτητες τυχαίες μεταβλητές με $E[X_i] = \mu_i, Var[X_i] = \sigma_i^2$ και έστω σταθερές $\alpha_i, i = 1, \dots, n$ με $\alpha_1 \geq \dots \geq \alpha_n > 0$. Τότε, $\forall c > 0$, έπεται ότι

$P[\max\limits_{r \leq k \leq n} \alpha_k |\sum\limits_{i=1}^{k}(X_i - \mu_i)| \geq c] \leq \frac{1}{c^2}(\alpha_r^2 \sum\limits_{i=1}^{r} \sigma_i^2 + \sum\limits_{i=r+1}^{n} \alpha_i^2 \sigma_i^2).$

Για $\alpha_i = \frac{1}{i}, i = r, r+1, \dots, n, c = \varepsilon$ παίρνουμε

$P[\max\limits_{r \leq k \leq n} |\frac{S_{k_n} - E[S_{k_n}]}{k_n}| \geq \varepsilon] \leq \frac{1}{\varepsilon^2}(\frac{1}{r^2} \sum\limits_{i=1}^{r} \sigma_i^2 + \sum\limits_{i=r+1}^{n} \frac{\sigma_i^2}{i^2}), r = 1, 2, \dots, n-1.$

Άρα για $n \to \infty$ προκύπτει ότι

$P[\sup\limits_{k \geq r} |\frac{S_{k_n} - E[S_{k_n}]}{k_n}| \geq \varepsilon] \leq \frac{1}{\varepsilon^2}(\frac{1}{r^2} \sum\limits_{i=1}^{r} \sigma_i^2 + \sum\limits_{i=r+1}^{n} \frac{\sigma_i^2}{i^2})$

Όμως,

$$
\begin{aligned}
\frac{1}{r^2} \sum_{i=1}^{r} \sigma_i^2 &= \frac{1}{r^2} \sum_{i=1}^{r} \frac{\sigma_i^2}{i^2} i^2 = \frac{1}{r^2} \sum_{i=1}^{r} \frac{\sigma_i^2}{i^2} (\sum_{j=i}^{r} j^2 - \sum_{j=i+1}^{r} j^2) \\
&= \frac{1}{r^2}(\frac{\sigma_1^2}{1^2}1^2 + \frac{\sigma_2^2}{2^2}2^2 + \dots + \frac{\sigma_r^2}{r^2}r^2) = \frac{1}{r^2}[1^2\frac{\sigma_1^2}{1^2} + 2^2(\frac{\sigma_1^2}{1^2} + \frac{\sigma_2^2}{2^2}) - 2^2(\frac{\sigma_1^2}{1^2}) + \\
&\qquad \dots + r^2(\frac{\sigma_1^2}{1^2} + \dots + \frac{\sigma_r^2}{r^2}) - r^2(\frac{\sigma_1^2}{1^2} + \dots + \frac{\sigma_{r-1}^2}{r-1^2})] \\
&= \frac{1}{r^2}[\sum_{i=1}^{r} i \sum_{j=1}^{r} \frac{\sigma_j^2}{j^2} - \sum_{i=2}^{r} i^2 \sum_{j=1}^{r-1} \frac{\sigma_j^2}{j^2}] = \frac{1}{r^2} \sum_{i=1}^{r} i \sum_{j=1}^{i} \frac{\sigma_j^2}{j^2} - \frac{1}{r^2} \sum_{i=2}^{r} i^2 \sum_{j=1}^{i-1} \frac{\sigma_j^2}{j^2} \\
&= \frac{1}{r^2}[\frac{\sigma_1^2}{1^2} + 2^2(\frac{\sigma_1^2}{1^2} + \frac{\sigma_2^2}{2^2}) + \dots + r^2(\frac{\sigma_1^2}{1^2} + \dots + \frac{\sigma_r^2}{r^2})] - \frac{1}{r^2} \sum_{i=2}^{r} i^2 \sum_{j=1}^{i-1} \frac{\sigma_j^2}{j^2} \\
&= \frac{1}{r^2}[r^2(\frac{\sigma_1^2}{1^2} \dots \frac{\sigma_r^2}{r^2})] + \frac{1}{r^2}[\frac{\sigma_1^2}{1^2} + \dots + (r-1)^2(\frac{\sigma_1^2}{1^2} + \dots + \frac{\sigma_{r-1}^2}{(r-1)^2})] - \frac{1}{r^2} \sum_{i=2}^{r} i^2 \sum_{j=1}^{i-1} \frac{\sigma_j^2}{j^2} \\
&= \sum_{j=1}^{r} \frac{\sigma_j^2}{j^2} + \frac{1}{r^2} \sum_{i=2}^{r} (i-1)^2 \sum_{j=1}^{i-1} \frac{\sigma_j^2}{j^2} - \frac{1}{r^2} \sum_{i=2}^{r} i^2 \sum_{j=1}^{i-1} \frac{\sigma_j^2}{j^2} \\
&= \sum_{j=1}^{r} \frac{\sigma_j^2}{j^2} + \frac{1}{r^2}[\sum_{i=2}^{r}((i-1)^2 - i^2)] \sum_{j=1}^{i-1} \frac{\sigma_j^2}{j^2} = \sum_{j=1}^{r} \frac{\sigma_j^2}{j^2} - \frac{1}{r^2} \sum_{i=2}^{r} (2i-1) \sum_{j=1}^{i-1} \frac{\sigma_j^2}{j^2}.
\end{aligned}
$$

και επειδή από υπόθεση $\sum\limits_{i=1}^{\infty} \frac{\sigma_i^2}{i^2} < \infty$ συμπεραίνουμε ότι $\lim\limits_{r \to \infty} \frac{1}{r^2} \sum\limits_{i=1}^{r} \sigma_i^2 = 0$.

Επιπροσθέτως, $\lim\limits_{r \to \infty} \sum\limits_{i=r+1}^{\infty} \frac{\sigma_i^2}{i^2} = \sum\limits_{i=1}^{\infty} \frac{\sigma_i^2}{i^2} - \lim\limits_{r \to \infty} \sum\limits_{i=1}^{r} \frac{\sigma_i^2}{i^2} \overset{\sum_{i=1}^{\infty} \frac{\sigma_i^2}{i^2} < \infty}{=} 0$. Επομένως, από Κριτήριο Παρεμβολής καταλήγουμε ότι $\lim\limits_{r \to \infty} P[\sup\limits_{r \leq k_n} |\frac{S_{k_n} - E[S_{k_n}]}{k_n}| \geq \varepsilon] = 0$ οπότε,

$P[\lim\limits_{n \to \infty} |\frac{S_n - E[S_n]}{n}| = 0] = 1$, άρα από Ορισμό 1.2.3. $\frac{S_n - E[S_n]}{n} \xrightarrow{a.s} 0$.

□

[16] Η απόδειξή της δύναται να βρεθεί στο [22].

Θεώρημα 1.4.4. ***(Kolmogorov για ανεξάρτητες και ισόνομες τυχαίες μεταβλητές)*** *Θεωρούμε ακολουθία ανεξάρτητων και ισόνομων τυχαίων μεταβλητών* $\{X_n : n = 1, 2, ...\}$ *με πεπερασμένη κοινή μέση τιμή* $\mu = E[X_1]$ *και διασπορά* $\sigma^2 = Var[X_1] < \infty$. *Έστω* $S_n = X_1 + X_2 + ... + X_n$. *Τότε* $\frac{S_n}{n} \xrightarrow{a.s} \mu$.

Απόδειξη. Αφού οι τυχαίες μεταβλητές έχουν κοινή διασπορά $Var[X_n] = \sigma^2 < \infty$ άρα η ποσότητα $\sum\limits_{n=1}^{\infty} \frac{\sigma_n^2}{n^2} = \sigma^2 \sum\limits_{n=1}^{\infty} \frac{1}{n^2} < \infty$. Επομένως, αφού $\{X_n : n = 1, 2, ...\}$ ανεξάρτητες τυχαίες μεταβλητές από υπόθεση, παρατηρούμε ότι καλύπτονται οι προϋποθέσεις του Θεωρήματος 1.4.3. και άρα
$\frac{S_n - E[S_n]}{n} \xrightarrow{a.s} 0 \overset{E[S_n]=n\mu}{\Rightarrow} \frac{S_n}{n} \xrightarrow{a.s} \mu$. □

Παρακάτω, αναφέρουμε μια επιπλέον απόδειξη κάνοντας χρήση του Λήμματος Borel-Cantelli:

Απόδειξη. 2^{η}. Έστω $\{X_n : n = 1, 2, ...\}$. Τότε, μπορούμε να ορίσουμε $\{X_n^+ : n = 1, 2, ...\}$ και $\{X_n^- : n = 1, 2, ...\}$ τέτοιες ώστε
$X_i^+ = X_i \vee 0 = max(X_i, 0)$, $X_i^- = (-X_i) \vee 0 = max(-X_i, 0)$.
Οπότε, μπορούμε να γράψουμε $X_i = X_i^+ - X_i^-$ και παρατηρούμε ότι $X_i \geq 0$. Επομένως, ορίζονται οι μέσες τιμές τους. Επιπλέον, θεωρούμε νέα τυχαία μεταβλητή $\Upsilon_i = X_i 1_i$, όπου

$$1_i = \begin{cases} 1 & \text{εάν } X_i \leq i \\ 0 & \text{εάν } X_i > i \end{cases} \tag{1.4.1}$$

η δείκτρια συνάρτηση και $S_n^* = \sum\limits_{i=1}^{n} Y_i$.Έστω $S_{k_n}^*$ υπακολουθια της S_n^*. Θα εφαρμόσουμε την ανισότητα του Chebyshev (Λήμμα 1.2.2.):
$\forall \alpha > 0, E[(X - E[X])^2] \geq \alpha^2 P[|X - E[X]| > \alpha]$ για την $\frac{S_{k_n}^*}{k_n}$.Έχουμε,$\forall \varepsilon > 0$,

$$\sum_{n=1}^{\infty} \{P[|\frac{S_{k_n}^* - E[S_{k_n}^*]}{k_n}| > \varepsilon]\} \leq \frac{1}{\varepsilon^2} \sum_{n=1}^{\infty} E[(\frac{S_{k_n}^* - E[S_{k_n}^*]}{k_n})^2]$$

$$= \frac{1}{\varepsilon^2} \sum_{n=1}^{\infty} Var[\frac{S_{k_n}^*}{k_n}] = \frac{1}{\varepsilon^2} \sum_{n=1}^{\infty} \frac{1}{k_n^2} \sum_{i=1}^{k_n} Var[Y_i] \leq \frac{1}{\varepsilon^2} \sum_{n=1}^{\infty} \frac{1}{k_n^2} \sum_{i=1}^{k_n} E[Y_i^2]$$

$$= \frac{1}{\varepsilon^2}[\frac{1}{k_1^2} E[Y_1^2] + \frac{1}{k_2^2}(E[Y_1^2] + E[Y_2^2]) + ...]$$

$$\leq \frac{1}{\varepsilon^2}[1E[Y_1^2] + \frac{1}{2^2}(E[Y_1] + E[Y_2]) + ...] = \frac{1}{\varepsilon^2} \sum_{n=1}^{\infty} \frac{1}{n^2} \int_0^n x^2 dF(x)$$

$$= \frac{1}{\varepsilon^2}[1 \int_0^1 x^2 dF(x) + \frac{1}{2^2} \int_0^2 x^2 dF(x) + ...]$$

$$= \frac{1}{\varepsilon^2}[1 \int_0^1 x^2 dF(x) + \frac{1}{2^2}\{\int_0^1 x^2 dF(x) + \int_1^2 x^2 dF(x)] + ...]$$

$$= \frac{1}{\varepsilon^2} \sum_{n=1}^{\infty} \frac{1}{n^2} \sum_{i=0}^{n-1} \int_i^{i+1} x^2 dF(x)$$
$$\leq \frac{1}{\varepsilon^2}[1 \int_0^1 x^2 dF(x) + \frac{1}{2} \int_0^2 x^2 dF(x) + ...]$$
$$= \frac{1}{\varepsilon^2} \sum_{i=0}^{\infty} \frac{1}{i+1} \int_i^{i+1} x^2 dF(x) \leq \frac{1}{\varepsilon^2} \sum_{i=0}^{\infty} \int_i^{i+1} |x| dF(x)$$
$$= \frac{1}{\varepsilon^2} E[X_1] < \infty, \quad \text{από υπόθεση,}$$

διότι $X_1 \leq X_2$ στον $\Lambda \Rightarrow \int_\Lambda X_1 dP \leq \int_\Lambda X_2 dP$ άρα
για $x \in [i, i+1]$, $\frac{x^2}{i+1} \leq \frac{(i+1)^2}{i+1} = i+1$ και $|x| \leq i+1$ οπότε διαιρώντας κατά μέλη $\frac{x^2}{i+1} \leq |x| \Rightarrow \int_i^{i+1} \frac{x^2}{i+1} dF(x) \leq \int_i^{i+1} |x| dF(x)$.
Επιπλέον, κάναμε χρήση της σχέσης $E[|X|] = \sum_{n=0}^{\infty} \int_n^{n+1} |X| dP$, αφού
$E[X] = \int_\Lambda X dP = \int_{\cup_n \Lambda_n} X dP = \sum_{n=0}^{\infty} \int_{\Lambda_n} X dP$.
Από Λήμμα Borel-Cantelli $\sum_{n=1}^{\infty} P[A_n] < \infty \Longrightarrow P[\lim_{n\to\infty} \sup A_n] = 0$ (βλ. Πρόταση 1.3.4.), έχουμε

$$\sum_{n=1}^{\infty} \{P[|\frac{S^*_{k_n} - E[S^*_{k_n}]}{k_n}| > \varepsilon]\} < \infty \Longrightarrow P[\lim_{n\to\infty} sup|\frac{S^*_{k_n} - E[S^*_{k_n}]}{k_n}| > \varepsilon] = 0$$
$$\Longleftrightarrow P[\lim_{n\to\infty} sup|\frac{S^*_{k_n} - E[S^*_{k_n}]}{k_n}| \leq \varepsilon] = 1$$
$$\overset{\varepsilon>0}{\Longleftrightarrow} P[\lim_{n\to\infty} |\frac{S^*_{k_n} - E[S^*_{k_n}]}{k_n}| = 0] = 1$$
$$\Longleftrightarrow P[\lim_{n\to\infty} |\frac{S^*_{k_n} - E[X_1]}{k_n}| = 0] = 1$$
$$\overset{Ορισμος 1.2.4}{\Longleftrightarrow} \frac{S^*_{k_n}}{k_n} \xrightarrow{a.s} E[X_1]$$

αφού $E[X_1] = \lim_{n\to\infty} \int_0^n x dF(x) = \lim_{n\to\infty} E[Y_n] = \lim_{n\to\infty} E[\frac{S^*_{k_n}}{k_n}]$.
Επιπλέον, $\sum_{n=1}^{\infty} P[Y_n \neq X_n] = \sum_{n=1}^{\infty} P[X_n > n] = \sum_{n=1}^{\infty} \int_n^{\infty} dF(x) = \sum_{n=1}^{\infty} \sum_{i=n}^{\infty} \int_i^{i+1} dF(x) =$

$$= \sum_{n=1}^{\infty} n \int_n^{n+1} dF(x)$$
$$= 1 \int_1^2 dF(x) + 2[\int_1^2 dF(x) + \int_2^3 dF(x)] + ...$$
$$\leq \sum_{n=1}^{\infty} \int_n^{n+1} x dF(x) \leq \sum_{n=0}^{\infty} \int_n^{n+1} x dF(x)$$
$$= E[X_1] < \infty.$$

Επομένως, από Λήμμα Borel-Cantelli,
$\sum_{n=1}^{\infty} P[Y_n \neq X_n] < \infty \Longrightarrow P[\lim_{n\to\infty} sup\{Y_n \neq X_n\}] = 0$. Οπότε μόνο για πεπερασμένες τιμές του n ισχύει ότι $Y_n \neq X_n$ και άρα αφού δείξαμε ότι
$\frac{S^*_{k_n}}{k_n} \xrightarrow{a.s} E[X_1] \Longleftrightarrow \frac{S_{k_n}}{k_n} \xrightarrow{a.s} E[X_1]$. Και επειδή $\{S_{k_n}\}$ συγκλίνουσα υπακολου-

θία της $\{S_n\}$ άρα $\frac{1}{c}E[X_1] \leq \underline{\lim}_{n\to\infty} \frac{S_n}{n} \leq \overline{\lim}_{n\to\infty} \frac{S_n}{n} \leq cE[X_1]$, για $c > 1$ και άρα καταλήγουμε ότι $\frac{S_n}{n} \xrightarrow{a.s} E[X_1]$ ή $\frac{S_n}{n} \xrightarrow{a.s} \mu$.

□

Παρατήρηση 1.4.1. Αφού από Θεώρημα 1.2.2. αποδείξαμε ότι η σχεδόν βέβαιη σύγκλιση συνεπάγεται τη σύγκλιση κατά πιθανότητα, είναι άμεσο ότι οι (Ι.Ν.Μ.Α) συνεπάγονται τους (Α.Ν.Μ.Α).

Θεώρημα 1.4.5. ***Κεντρικό Οριακό Θεώρημα***[17] ***(Lindeberg-Lévy)*** *Θεωρούμε ακολουθία ανεξάρτητων και ισόνομων τυχαίων μεταβλητών $\{X_n : n = 1, 2, ...\}$ με πεπερασμένη μέση τιμή $\mu = E[X_n]$ και διασπορά $\sigma^2 = Var[X_n]$.*
Έστω $S_n = X_1 + X_2 + ... + X_n$. Τότε $\frac{S_n - n\mu}{\sigma\sqrt{n}} \xrightarrow{D} Z \sim N(0, 1)$
ή ισοδύναμα $\sqrt{n}(\overline{X} - \mu) \xrightarrow{D} N(0, \sigma^2)$.

Απόδειξη. Θεωρούμε τυχαία μεταβλητή
$\Upsilon_n = \frac{S_n - n\mu}{\sigma\sqrt{n}} = \frac{\sum_{i=1}^{n} X_i - n\mu}{\sigma\sqrt{n}} = \frac{\sum_{i=1}^{n}(X_i - \mu)}{\sigma\sqrt{n}} = \frac{1}{\sigma\sqrt{n}}\sum_{i=1}^{n}(X_i - \mu), n = 1, 2, ...$
Έστω φ η χαρακτηριστική συνάρτηση των τυχαίων μεταβλητών $X_i - \mu$.

Λήμμα 1.4.1. *Έστω φ_X η χαρακτηριστική συνάρτηση της τυχαίας μεταβλητής X.*
Αν $\Upsilon = aX + b$ με $a, b \in \mathbb{R} \Rightarrow \varphi_Y(t) = \varphi_X(at)e^{ibt}$.

Απόδειξη. Λήμματος 1.4.1 Έχουμε,

$$\varphi_\Upsilon(t) = E[e^{itY}] = E[e^{it(aX+b)}] = E[e^{itaX}e^{itb}] = E[e^{i(ta)X}]e^{itb} = \varphi_X(at)e^{ibt}.$$

□

Λήμμα 1.4.2. *Έστω $\{X_i\}_{i=1...n}$ ακολουθία ανεξάρτητων τυχαίων μεταβλητών με χαρακτηριστικές συναρτήσεις $\varphi_i, i = 1, ... n$ αντιστοίχως.*
Αν $\Upsilon = \sum_{i=1}^{n} X_i \Rightarrow \varphi_\Upsilon(t) = \prod_{i=1}^{n} \varphi_i(t)$.

[17] Το Κεντρικό Οριακό Θεώρημα διατυπώθηκε για πρώτη φορά το 1733 από το Γάλλο μαθηματικό Abraham de Moivre. Παρά ταύτα, λόγω της καινοτομίας του, δεν του δόθηκε η ανάλογη σημασία από τους μαθηματικούς της εποχής του. Το ίδιο συνέβη και το 1812, όταν ο Γάλλος μαθηματικός Pierre-Simon Laplace το επανέφερε στο μαθηματικό προσκήνιο δίνοντάς του την τελική του μορφή. Εν τέλει, ήταν το 1901, μετά την απόδειξη της μαθηματικής λειτουργίας του Θεωρήματος από το Ρώσο μαθηματικό Aleksandr Lyapunov, που το Θεώρημα αποτέλεσε κυρίαρχο θεώρημα της Θεωρίας Πιθανοτήτων. Η πλέον διαδεδομένη του διατύπωση είναι αυτή των Lindeberg-Lévy.

Απόδειξη. Λήμματος 1.4.2. Παρατηρούμε ότι

$$\varphi_\Upsilon(t) = E[e^{itY}] = E[e^{it(\sum\limits_{i=1}^{n} X_i)}] = E[\prod_{i=1}^{n} e^{itX_i}] \overset{X_i \alpha\nu\varepsilon\xi}{=} \prod_{i=1}^{n} E[e^{itX_i}] = \prod_{i=1}^{n} \varphi_i(t).$$

□

Λήμμα 1.4.3. ***(Lévy)***[18]

Θεωρούμε $\{X_i\}_{i=1,2,\dots}$ ακολουθία τυχαίων μεταβλητών με αντίστοιχες χαρακτηριστικές συναρτήσεις $\varphi_i, i = 1, 2, \dots$.
α)$X_i \xrightarrow{D} X \Leftrightarrow \varphi_n(t) \to \varphi(t), \forall t \in \mathbb{R}$.
β)Έστω ότι $\forall t \in \mathbb{R}, \exists \lim\limits_{n\to\infty} \varphi_n(t)$ και η συνάρτηση $\lim\limits_{n\to\infty} \varphi_n(t)$ είναι συνεχής στο $t = 0$. Τότε, η συνάρτηση $\lim\limits_{n\to\infty} \varphi_n$ είναι η χαρακτηριστική συνάρτηση μιας τυχαίας μεταβλητής X και επιπλέον $X_i \xrightarrow{D} X$.

Συνέχεια Απόδειξης Θεωρήματος 1.4.5.
Από Λήμμα 1.4.2. έχουμε $\varphi_{\Sigma(X_i-\mu)} = \prod\limits_{i=1}^{n} \varphi(t) = [\varphi(t)]^n$ και από Λήμμα 1.4.1. η χαρακτηριστική συνάρτηση της Y_n είναι $\varphi_n(t) = [\varphi(\frac{t}{\sigma\sqrt{n}})]^n$, $n \in \mathbb{N}$.
Αποδεικνύεται ότι αν η τυχαία μεταβλητή X έχει πεπερασμένη ροπή n-τάξης, τότε η n-οστή παράγωγος της χαρακτηριστικής συνάρτησης φ_X υπάρχει και υπολογίζεται ως εξής:
$\varphi_X^{(n)} = \frac{d^n}{dt^n} E[e^{itX}] = E[\frac{d^n}{dt^n} e^{itX}] = E[i^n X^n e^{itX}] = i^n E[X^n e^{itX}]$.
Αφού από υπόθεση $\mu, \sigma^2 < \infty \Rightarrow E[X_i - \mu] < \infty, E[(X_i - \mu)^2] < \infty$, άρα υπάρχουν οι παράγωγοι πρώτης και δεύτερης τάξης της χαρακτηριστικής συνάρτησης φ των τυχαίων μεταβλητών $X_i - \mu$. Συγκεκριμένα, για $t = 0$,
$\varphi^{(1)}(0) = i^1 E[(X_i - \mu)^1 e^0] = iE[X_i - \mu] = 0$.
$\varphi^{(2)}(0) = i^2 E[(X_i - \mu)^2 e^0] = i^2 E[(X_i - \mu)^2] = i^2\sigma^2 = -\sigma^2$.
Θα αναπτύξουμε τη χαρακτηριστική συνάρτηση φ σε σειρά Taylor γύρω από το μηδέν: $\varphi(t) = \varphi(0) + \varphi^{(1)}(0)t + \frac{1}{2}\varphi^{(2)}(0)t^2 + o(t^2)$

$$= E[e^0] + 0 + \tfrac{1}{2}(-\sigma^2)t^2 + o(t^2) = 1 - \tfrac{1}{2}\sigma^2 t^2 + o(t^2) \quad , \lim_{t\to 0} \frac{o(t^2)}{t^2} \to 0.$$

Επομένως, για τη χαρακτηριστική συνάρτηση της Y_n:
$\varphi_n(t) = [\varphi(\frac{t}{\sigma\sqrt{n}})]^n = [1 - \frac{1}{2}\sigma^2(\frac{t}{\sigma\sqrt{n}})^2 + o(\frac{t^2}{\sigma^2 n})]^n = [1 - \frac{t^2}{2n} + \frac{1}{\sigma^2} o(\frac{t^2}{n})]^n$.
Και $\lim\limits_{n\to\infty} \frac{o(\frac{t^2}{n})}{\frac{t^2}{n}} = 0$, άρα $\lim\limits_{n\to\infty} \varphi_n(t) = e^{-\frac{1}{2}t^2}$.
Παρατηρούμε ότι η συνάρτηση $\lim\limits_{n\to\infty} \varphi_n(t) = e^{-\frac{1}{2}t^2}$ είναι συνεχής στο $t = 0$.

[18] Paul Pierre Lévy (1886–1971): Γάλλος μαθηματικός, μέλος της Γαλλικής Ακαδημίας Επιστημών καθώς και της Μαθηματικής Εταιρείας του Λονδίνου. Η ενασχόλησή του πρωτίστως με τη Θεωρία Πιθανοτήτων είχε ως αποτέλεσμα πληθώρα μαθηματικών εργαλείων και θεωρημάτων, πολλά από τα οποία φέρουν το όνομά του.

Επομένως, από Λήμμα 1.4.3.(β) $\Upsilon_n = \frac{S_n - n\mu}{\sigma\sqrt{n}} \xrightarrow{D} Z$, όπου Z μια τυχαία μεταβλητή με χαρακτηριστική συνάρτηση $\varphi_Z = \lim_{n\to\infty} \varphi_n(t) = e^{-\frac{1}{2}t^2}$.
Αποδεικνύεται ότι αν $Z \sim N(0,1) \Leftrightarrow \varphi(t) = e^{-\frac{t^2}{2}}$.
Επομένως, συμπεραίνουμε ότι $\frac{S_n - n\mu}{\sigma\sqrt{n}} \xrightarrow{D} Z \sim N(0,1)$.

□

Θεώρημα 1.4.6. ***(Πολυδιάστατο Κεντρικό Οριακό Θεώρημα)*** *Θεωρούμε τα ανεξάρτητα και ισόνομα τυχαία διανύσματα* $\boldsymbol{X}^{(j)} = (X_1^{(j)}, X_2^{(j)}, \dots, X_k^{(j)}), j = 1, \dots, n$ *και* $\boldsymbol{X} = (X_1, X_2, \dots, X_k)$ *με μέση τιμή* $\boldsymbol{\mu} = (\mu_1, \mu_2, \dots, \mu_k)$ *και πεπερασμένο πίνακα διασποράς* $\boldsymbol{\Sigma} = [\sigma_{ij}] \in \mathbb{R}^{kxk}$. *Εάν* $\overline{X}_i = \frac{1}{n}(X_i^{(1)} + \dots + X_i^{(n)}) = \frac{1}{n}\sum_{k=1}^{n} X_i^{(k)}$, *τότε έπεται ότι* $(\sqrt{n}(\overline{X}_1 - \mu_1), \dots, \sqrt{n}(\overline{X}_k - \mu_k)) \xrightarrow{D} N(\boldsymbol{0}, \boldsymbol{\Sigma})$.

Απόδειξη. Για την απόδειξη του Θεωρήματος θα κάνουμε χρήση του παρακάτω Λήμματος:

Λήμμα 1.4.4. ***(Τέχνασμα Cramer-Wold)*** *Έστω τα τυχαία διανύσματα* $\boldsymbol{X}^{(n)} = (X_1^{(n)}, X_2^{(n)}, \dots, X_k^{(n)})$, $\boldsymbol{X} = (X_1, X_2, \dots, X_k)$. *Αναγκαία και ικανή συνθήκη ώστε να ισχύει ότι* $(X_1^{(n)}, X_2^{(n)}, \dots, X_k^{(n)}) \xrightarrow{D} (X_1, X_2, \dots, X_k)$
είναι να αληθεύει ότι $\sum_{i=1}^{k} t_i X_i^{(n)} \xrightarrow{D} \sum_{i=1}^{k} t_i X_i \quad , \forall (t_1, t_2, \dots, t_k) \in \mathbb{R}^k$.

Απόδειξη. Λήμματος 1.4.4. Θα χρησιμοποιήσουμε το Θεώρημα συνέχειας του Levy (βλ. Λήμμα 1.4.3.(α)). Έχουμε,

$$\begin{aligned}
\mathbf{X}^{(n)} \xrightarrow{D} \mathbf{X} &\Leftrightarrow \varphi_n(\mathbf{t}) \to \varphi(\mathbf{t}), \mathbf{t} = (t_1, \dots, t_k)^T \in \mathbb{R}^k \Leftrightarrow E[e^{i\mathbf{t}^T\mathbf{X}^n}] \to E[e^{i\mathbf{t}^T\mathbf{X}}] \\
&\Leftrightarrow E[e^{i\sum_{i=1}^{k} t_i X_i^{(n)}}] \to E[e^{i\sum_{i=1}^{k} t_i X_i}] \\
&\Leftrightarrow \sum_{i=1}^{k} t_i X_i^{(n)} \xrightarrow{D} \sum_{i=1}^{k} t_i X_i \quad , \forall \mathbf{t} = (t_1, \dots, t_k)^T \in \mathbb{R}^k.
\end{aligned}$$

□

Συνέχεια Απόδειξης Πολυδιάστατου Κ.Ο.Θ. Σύμφωνα με το τέχνασμα Cramer-Wold που αναφέρθηκε, αρκεί να αποδείξουμε ότι $\sqrt{n}\sum_{i=1}^{k} t_i(\overline{X}_i - \mu_i) \xrightarrow{D} N(0, \mathbf{t}'\boldsymbol{\Sigma}\mathbf{t})$, $\forall \mathbf{t} = (t_1, \dots, t_k)$. Όμως, $\sqrt{n}\sum_{i=1}^{k} t_i(\overline{X}_i - \mu_i) = \sqrt{n}[\frac{1}{n}\sum_{j=1}^{n}(\sum_{i=1}^{k} t_i X_i^{(j)}) - \sum_{i=1}^{k} t_i\mu_i]$.
Οι τυχαίες μεταβλητές $Y_i = \sum_{i=1}^{k} t_i X_i^{(j)}, j = 1, 2, \dots$ είναι ανεξάρτητες και ισό-

νομες με μέση τιμή $\sum_{i=1}^{k} t_i\mu_i$ και διασπορά $\mathbf{t}'\mathbf{\Sigma}\mathbf{t}$. Άρα από μονοδιάστατο Κεντρικό Οριακό Θεώρημα προκύπτει το ζητούμενο.

□

1.5 Χρήσιμοι Ορισμοί-Θεωρήματα

Ορισμός 1.5.1. Θεωρούμε την ακολουθία εκτιμητριών $\{T_n : n = 1, 2, ...\}$ της παραμετρικής συνάρτησης $\alpha(\theta)$. Η εκτιμήτρια T_n ονομάζεται *συνεπής* όταν $\forall \varepsilon > 0$ ισχύει ότι $\lim_{n\to\infty} P[\| T_n - \alpha(\theta) \| > \varepsilon] = 0 \overset{O\rho 1.2.3.}{\Leftrightarrow} T_n \xrightarrow{P} g(\theta)$.

Ορισμός 1.5.2. Θεωρούμε την ακολουθία εκτιμητριών $\{T_n : n = 1, 2, ...\}$ (ή εκτιμήτρια T_n) της παραμετρικής συνάρτησης $\alpha(\theta)$. Η εκτιμήτρια T_n ονομάζεται *ισχυρά συνεπής* όταν ισχύει ότι
$P[\lim_{n\to\infty} \| T_n - \alpha(\theta) \| = 0] = 1 \overset{O\rho 1.2.4.}{\Leftrightarrow} T_n \xrightarrow{a.s} \alpha(\theta)$.

Ορισμός 1.5.3. Η εκτιμήτρια T_nτης παραμετρικής συνάρτησης $\alpha(\theta)$ ονομάζεται *αμερόληπτη* όταν ισχύει ότι $E[T_n] = \alpha(\theta)$, ενώ λέγεται *ασυμπτωτικά αμερόληπτη* όταν $E[T_n] \to \alpha(\theta)$, καθώς $n \to \infty$.

Ορισμός 1.5.4. Η εκτιμήτρια T_n μιας παραμέτρου θ ονομάζεται *ασυμπτωτικά Κανονική* όταν $\exists \sigma^2 = \sigma^2(\theta) : \forall \theta \in \Theta \quad \sqrt{n}(T_n - \theta) \xrightarrow{D} N(0, \sigma^2)$.

Θεώρημα 1.5.1. ***(Θεώρημα Slutsky[19])*** *Έστω $X_n \xrightarrow{D} X$, $A_n \xrightarrow{P} a$ και $B_n \xrightarrow{P} b$. Τότε, αληθεύει ότι $A_n + B_nX_n \xrightarrow{D} a + bX$.*

Θεώρημα 1.5.2. ***(Θεώρημα Skorohod[20])*** *Έστω τα μέτρα πιθανότητας P_n, P στον $(\mathbb{R}^k, \mathfrak{R}^k)$ και $\mu_n \to \mu$ ή ισοδύναμα έστω $\boldsymbol{X}_n, \boldsymbol{X}$ τυχαία διανύσματα τέτοια ώστε $\boldsymbol{X}_n \xrightarrow{D} \boldsymbol{X}$. Τότε, υπάρχουν τυχαία διανύσματα $\boldsymbol{Y}_n, \boldsymbol{Y}$ στο χώρο πιθανότητας (Ω,F,P) τέτοια ώστε $\boldsymbol{Y}_n, \boldsymbol{Y}$ να έχουν κατανομές P_n, P αντιστοίχως και $\boldsymbol{Y}_n(\omega) \to \boldsymbol{Y}(\omega), \forall \omega \in \Omega$.*

[19]Evgeny Slutsky (1880-1948): Ρώσος μαθηματικός και οικονομολόγος, ο οποίος ασχολήθηκε κυρίως με τη στατιστική και τα οικονομικά. Τα πορίσματά του, όπως η εξίσωση Slutsky, εφαρμόστηκαν σε διάφορους επιστημονικούς κλάδους όπως τα οικονομικά.

[20]Anatoliy Volodymyrovych Skorokhod (1930–2011): Ουκρανός μαθηματικός, ακαδημαϊκός στην Εθνική Ακαδημία Επιστημών της Ουκρανίας, καθηγητής του Kyiv University και του Michigan State University. Η επιστημονική του δραστηριότητα είναι έκδηλη από το πλήθος των δημοσιεύσεων του, οι οποίες ανέρχονται στις 450 και αναφέρονται κυρίως στις στοχαστικές διαφορικές εξισώσεις, τις Markov ανελίξεις, τις κατανομές απειροδιάστατων χώρων.

1.6 Σχέσεις $O(.)$ και $o(.)$

Ορισμός 1.6.1. Έστω $\{\alpha_n\}_{n\in\mathbb{N}}$,$\{b_n\}_{n\in\mathbb{N}}$ ακολουθίες πραγματικών αριθμών.
•Λέμε ότι οι $\{\alpha_n\}_{n\in\mathbb{N}}$,$\{b_n\}_{n\in\mathbb{N}}$ είναι (ασυμπτωτικά) ισοδύναμες και συμβολίζουμε $\alpha_n \sim b_n$ καθώς $n \to \infty \Leftrightarrow \frac{\alpha_n}{b_n} \to 1$.
•Λέμε ότι $\alpha_n = o(b_n)$ καθώς $n \to \infty$ αν αληθεύει $\frac{\alpha_n}{b_n} \to 0$.
•Λέμε ότι $\alpha_n = O(b_n)$ αν $\exists M$ και n_0 τέτοια ώστε $\forall n > n_0, \quad |\frac{\alpha_n}{b_n}| < M$.
(δηλαδή $\{\frac{\alpha_n}{b_n}\}$ φραγμένη).

Λήμμα 1.6.1. *Έστω $\{\alpha_n\}_{n\in\mathbb{N}}$,$\{b_n\}_{n\in\mathbb{N}}$ ακολουθίες πραγματικών αριθμών.*
(i) $\alpha_n = o(1) \Leftrightarrow \alpha_n \to 0$.
(ii) $\alpha_n = O(1) \Leftrightarrow \exists M, n_0 : \forall n > n_0 \quad |\alpha_n| < M$ (α_n φραγμένη).

Ορισμός 1.6.2. Έστω $\{A_n\}_{n\in\mathbb{N}}$,$\{B_n\}_{n\in\mathbb{N}}$ ακολουθίες τυχαίων μεταβλητών.
•Λέμε ότι $A_n = o_P(B_n)$ αν $\frac{A_n}{B_n} \xrightarrow{P} 0$.
•Λέμε ότι $A_n = O_P(B_n)$ αν
$\forall \varepsilon > 0, \exists M, n_0 : \forall n > n_0 \quad P[|A_n| \leq M|B_n|] \geq 1 - \varepsilon$, (δηλαδή, $\{\frac{A_n}{B_n}\}$ φραγμένη κατά πιθανότητα).

Λήμμα 1.6.2. *Έστω $\{A_n\}_{n\in\mathbb{N}}$,$\{B_n\}_{n\in\mathbb{N}}$ ακολουθίες τυχαίων μεταβλητών.*
(i) $A_n = o_P(1) \Leftrightarrow A_n \xrightarrow{P} 0$.
(ii) $A_n = O_P(1) \Leftrightarrow \exists M, n_0 : \forall n > n_0 \quad P[|A_n| \leq M] \geq 1 - \varepsilon$ (A_n φραγμένη κατά πιθανότητα).

Θεώρημα 1.6.1. *Εάν $X_n \xrightarrow{D} H$, όπου H συνάρτηση, τότε η X_n είναι φραγμένη κατά πιθανότητα, δηλαδή $X_n = O_P(1)$.*

Κεφάλαιο 2

Σειρές Taylor

2.1 Ιστορική Αναδρομή

Το βασικότερο αντικείμενο μελέτης του Απειροστικού Λογισμού[1] αποτελούν οι συναρτήσεις και πιο συγκεκριμένα, η διαφόριση και ολοκλήρωσή τους. Παρά ταύτα, η πολυπλοκότητα του τύπου μιας συνάρτησης καθιστά πολλές φορές περίπλοκο έως αδύνατο τον υπολογισμό των τιμών της και τη γενικότερη διαχείρισή της. Η δυσκολία αυτή μπορεί να παρακαμφθεί, προσεγγίζοντας την εν λόγω συνάρτηση μέσω μιας νέας, πιο απλουστευμένης παράστασης, όπως για παράδειγμα ένα πολυώνυμο, λαμβάνοντας πάντα υπόψη το σφάλμα που δημιουργείται. Η πλέον γνωστή και ευρέως διαδεδομένη προσέγγιση μιας συνάρτησης είναι η ανάπτυξή της σε **σειρά Taylor**.

Ο **Brook Taylor** γεννήθηκε στο Canterbury στις 18 Αυγούστου 1685 από εύπορη οικογένεια. Η κατ' οίκον διδασκαλία που έλαβε κατά τα νεανικά του χρόνια του επέτρεψε την εισαγωγή στο Cambridge (1701), όπου διηύρυνε τους ορίζοντές του τόσο στα νομικά όσο και στα μαθηματικά ζητήματα. Ωστόσο, τα δεύτερα φαίνεται να κέρδισαν το ενδιαφέρον του, καθώς δεν εφήρμοσε ποτέ την ιδιότητα του δικηγόρου. Άμεσα μετά την αποφοίτησή του επιλέχθηκε να συμμετέχει στη Βασιλική Ένωση, γεγονός που του έδωσε τη δυνατότητα να έρθει σε επαφή με άλλους μαθηματικούς της εποχής του και τελικά να αποτελέσει μέλος της ομάδας του Isaak Newton. [2]

[1]Κλάδος των μαθηματικών που ορίστηκε αρχικά ως ο λογισμός των σχέσεων μεταξύ ροών και ρευσών, έννοιες που εισήχθησαν από τον Isaak Newton και εκφράζουν την ταχύτητα αύξησης μιας ποσότητας και την τελική ποσότητα μετά τη μεταβολή αντίστοιχα. Η έννοια της συνάρτησης εισάγεται πρώτη φορά το 1692 από τον αυτοδίδακτο μαθηματικό G.W.Leibniz (1646-1716). Οι προαναφερόμενοι ασχολήθηκαν κυρίως με την πράξη της διαφόρισης ενώ η ολοκλήρωση αναπτύχθηκε μεταγενέστερα από τους Riemman,Baire,Lebesgue.

[2]Isaak Newton (1643–1727): Άγγλος φυσικός, μαθηματικός, αστρονόμος, φιλόσοφος, αλχημιστής και θεολόγος. Η συνεισφορά του, ιδιαίτερα στους κλάδους της Κλασικής Φυσικής και του

Το πλήθος των δημοσιεύσεών του διαφωτίζει το εύρος των μαθηματικών θεμάτων που τον απασχόλησαν όπως οι σειρές, η ολοκλήρωση συναρτήσεων (απέδειξε τον τύπο παραγοντικής ολοκλήρωσης), η θεωρία πιθανοτήτων, κάποια προβλήματα θερμότητας και μαγνητισμού ενώ η πρώτες δημοσιεύσεις του έγιναν το 1713 στο **«Philosophical Transactions»** και αφορούσαν στο κέντρο μιας ταλάντωσης.

Η εργασία που καταξίωσε τον Taylor φέρει τον τίτλο **«The Methodus Incrementorum directa et traversa» (ευθεία και αντίστροφη μέθοδος των προσαυξήσεων)** και δημοσιεύτηκε στο Λονδίνο το 1715. Η προαναφερθείσα δημοσίευση αναφέρεται στη θεωρία των ροών του Isaak Newton, ενώ σημαντικά στοιχεία αποτελούν η **πρόταση 7-πόρισμα 2**, που αφορά στις σειρές Taylor καθώς και η πρόταση 11 στην οποία το ολοκλήρωμα εκφράζεται ως άπειρη σειρά.

Η πεποίθηση του Johann Bernoulli[3] ότι ο Taylor στην εργασία του οικειοποιήθηκε δικές του ιδέες, ήδη δημοσιευμένες καθώς και η πυκνογραμμένη, δυσνόητη γραφή του τελευταίου υπήρξαν ανασταλτικοί παράγοντες στην υιοθέτηση των πορισμάτων του. Παρ' όλα αυτά, πληθώρα μαθηματικών σύγχρονων ή μεταγενέστερων του Taylor (όπως οι James Gregory[4], Isaak Newton,James Stirling[5], Abraam De Moivre[6]) ασχολήθηκαν διεξοδικά με την εργασία του συντελώντας στη διαμόρφωση της σειράς Taylor και του Θεωρήματος Taylor στη σημερινή τους μορφή.

2.2 Πολυώνυμο Taylor

Θεωρούμε ότι η συνάρτηση f είναι άπειρες φορές διαφορίσιμη σε μια γειτονιά ενός (εσωτερικού) σημείου α του πεδίου ορισμού της. Η προσέγγιση της συνάρτησης f είναι επιτυχής εάν η νέα συνάρτηση μας διευκολύνει στους αριθμητικούς υπολογισμούς. Οι πλέον απλοποιημένες συναρτήσεις που ενδείκνυνται για τη διαδικασία αυτή αφού χαρακτηρίζονται από εύκολους αλγεβρικούς υπολογισμούς και ολοκλήρωση σε πεπερασμένο αριθμό βημάτων είναι τα πολυώνυμα.

Συνεπώς, στόχος μας είναι να εκτιμήσουμε ένα πολυώνυμο, έστω $P_n(x)$ το οποίο να έχει όσο το δυνατόν περισσότερα κοινά χαρακτηριστικά με τη συνάρ-

---

Διαφορικού λογισμού, είναι αναμφισβήτητη με χαρακτηρίζεται από πληθώρα ευρημάτων και νεωτερισμών καθιστώντας τον θεμελιωτή τους.

[3]Johann Bernoulli(1667-1748): Σουηδός μαθηματικός ο οποίος συνέβαλλε στην ανάπτυξη του Απειροστικού Λογισμού και υπήρξε καθηγητής του Leonhard Euler (1707-1783).

[4]James Gregory (1638-1675): Σκοτσέζος μαθηματικός και αστρονόμος, ανέπτυξε άπειρες σειρές για κάποιες τριγωνομετρικές συναρτήσεις ασχολήθηκε ευρέως με την τριγωνομετρία.

[5]James Stirling (1692-1770): Σκοτσέζος μαθηματικός, γνωστός για την προσέγγιση(ή τύπο) Stirling.

[6]Abraham de Moivre (1667–1754): Γάλλος μαθηματικός ασχολήθηκε ιδιαίτερα με τη θεωρία πιθανοτήτων και εισήγαγε την Κανονική Κατανομή.

τηση f,δηλαδή να την προσεγγίζει καλύτερα γύρω από το σημείο α.

Υπενθυμίζουμε ότι αν $f(x)$ είναι παραγωγίσιμη συνάρτηση στο σημείο $x = x_0$ τότε η $f(x)$ προσεγγίζεται από τη γραμμική συνάρτηση
$f_1(x) = f(x_0) + f'(x_0)(x - x_0)$,
η οποία γεωμετρικά αντιστοιχεί στην εφαπτομένη της συνάρτησης f στην περιοχή του x_0. Για την επίτευξη μεγαλύτερης ακρίβειας (και δεδομένου ότι fείναι δύο φορές παραγωγίσιμη) θα προσεγγίσουμε την f μέσω μιας νέας συνάρτησης f_2 η οποία ταυτίζεται στο σημείο x_0 με την f, έχει την ίδια κλίση με την f στο σημείο αυτό (οι δύο αυτές συνθήκες ισχύουν και για την συνάρτηση f_1) ενώ επιπλέον θα καμπυλώνει στην ίδια κατεύθυνση με την f στο x_0. Και επειδή η καμπυλότητα προέρχεται από τη μεταβολή της κλίσης ενός γραφήματος άρα ζητάμε επιπλέον $f_2'(x_0) = f'(x_0)$. Έτσι, προκύπτει νέα συνάρτηση
$f_2 = f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2}f''(x_0)(x - x_0)^2$.

Συνεπώς, αφού η γεωμετρία της συνάρτησης f σε ένα σημείο α σχετίζεται με τις τιμές των παραγώγων της στο σημείο α, θέλουμε να βρούμε ένα πολυώνυμο του οποίου η τάξη επαφής[7] του γραφήματός του με το γράφημα της f να είναι k, για όσο το δυνατόν μεγαλύτερο k. Δηλαδή, ζητάμε ένα πολυώνυμο $P_n(x)$ βαθμού n τέτοιο ώστε να ισχύει $P_n^{(k)}(\alpha) = f^{(k)}(\alpha)$, για όσο το δυνατόν περισσότερες τιμές του k.

Θεωρούμε ένα πολυώνυμο $P_n(x) = a_0 + a_1x + a_2x^2 + ..., +a_nx^n$.
Τότε, για κάθε $k \geq n + 1$ είναι άμεσο ότι $P_n^{(k)}(x) = 0$ εφόσον
$P_n^{(n)}(x) = a_n n(n-1) \dots 1 = n!\, a_n$. Επομένως, αρκεί να βρούμε ένα πολυώνυμο $P_n(x)$ τέτοιο ώστε να ισχύει

$$P_n^{(k)}(\alpha) = f^{(k)}(\alpha) \tag{2.2.1}$$

για κάθε $k = 0, \dots, n$. Έπειτα από διαδοχικές διαιρέσεις του πολυωνύμου με $(x - \alpha)$μπορούμε να γράψουμε το παραπάνω στην ισοδύναμη μορφή
$P_n(x) = a_0 + a_1(x - \alpha) + a_2(x - \alpha)^2 + ..., +a_n(x - \alpha)^n$.

Θα δείξουμε ότι

$$P_n^{(k)}(x) = k!\, a_k + (x - \alpha)Q_{n-k-1}(x), \forall k \in 0, \dots, n, \tag{2.2.2}$$

όπου $Q_{n-k-1}(x)$ ένα πολυώνυμο $n - k - 1$ βαθμού της μορφής

[7] Δύο καμπύλες K_1, K_2 με παραστάσεις $\overline{r_1} = \overline{r_1}(s)$ και $\overline{r_2} = \overline{r_2}(t)$, αντίστοιχα, με παράμετρο τα μήκη τους, λέμε ότι έχουν επαφή τάξης k όταν ισχύει $\overline{r_1} = \overline{r_2}, \overline{r_1}' = \overline{r_2}', \dots \overline{r_1}^{(k)} = \overline{r_2}^{(k)}$ για κάποιες τιμές των παραμέτρων s, t.

$$Q_{n-k-1}(x) = (k+1)!\,a_{k+1} + \tfrac{(k+2)!}{2!}a_{k+2}(x-\alpha) + \tfrac{(k+3)!}{3!}a_{k+3}(x-\alpha)^2 + \ldots + \tfrac{n!}{n-k!}a_n(x-\alpha)^{n-k-1} = \sum_{i=1}^{n-k} \tfrac{(k+i!)}{i!}a_{k+i}(x-\alpha)^{i-1}.$$

Απόδειξη. Για την απόδειξη θα εργαστούμε με Μαθηματική Επαγωγή. Έχουμε

- για $k=1$ τότε $P_n^{(1)}(x) = a_1+2a_2(x-\alpha)+3a_3(x-\alpha)^2+\ldots+na_n(x-\alpha)^{n-1}$
$= a_1+(x-\alpha)[2a_2+3a_3(x-\alpha)+\ldots+na_n(x-\alpha)^{n-2}]$
$= 1!\,a_1 + (x-\alpha)Q_{n-2}$

  , όπου $Q_{n-2}(x) = 2a_2 + 3a_3(x-\alpha) + \ldots + na_n(x-\alpha)^{n-2}$
$= \sum_{i=1}^{n-1} \tfrac{(1+i!)}{i!}a_{1+i}(x-\alpha)^{i-1}.$

- Υποθέτουμε ότι η παραπάνω ισότητα ισχύει για τυχαίο k, δηλαδή ισχύει ότι $P_n^{(k)}(x) = k!\,a_k + (x-\alpha)Q_{n-k-1}(x)$, όπου

$$Q_{n-k-1}(x) = (k+1)!\,a_{k+1} + \frac{(k+2)!}{2!}a_{k+2}(x-\alpha) + \ldots + \frac{n!}{n-k!}a_n(x-\alpha)^{n-k-1} \tag{2.2.3}$$

  $= \sum_{i=1}^{n-k} \tfrac{(k+i!)}{i!}a_{k+i}(x-\alpha)^{i-1}$ ένα πολυώνυμο $n-k-1$ βαθμού.

- Αρκεί να δείξουμε ότι η ισότητα ισχύει για k+1. Έχουμε

$$\begin{aligned} P_n^{(k+1)}(x) &= (P_n^{(k)}(x))' = (k!\,a_k + (x-\alpha)Q_{n-k-1}(x))' = ((x-\alpha)Q_{n-k-1}(x))' \\ &= Q_{n-k-1}(x) + (x-\alpha)(Q_{n-k-1}(x))' \\ &= (k+1)!\,a_{k+1} + \tfrac{(k+2)!}{2!}a_{k+2}(x-\alpha) + \ldots + \tfrac{n!}{n-k!}a_n(x-\alpha)^{n-k-1} + \\ &\qquad (x-\alpha)[\tfrac{(k+2)!}{2!}a_{k+2} + \ldots + \tfrac{n!}{n-k!}a_n(x-\alpha)^{n-k-2}] \\ &= (k+1)!\,a_{k+1} + 2(x-\alpha)[\tfrac{(k+2)!}{2!}a_{k+2} + \ldots + \tfrac{n!}{n-k!}a_n(x-\alpha)^{n-k-2}] \\ &= (k+1)!\,a_{k+1} + (x-\alpha)Q_{n-k-2}(x) \end{aligned}$$

  Επομένως, ως συνέπεια της μαθηματικής επαγωγής προκύπτει το ζητούμενο.

$\square$

Άρα η σχέση (2.2.2) γίνεται

$$P_n^{(k)}(x) = k!\, a_k + (x-\alpha)Q_{n-k-1}(x) \tag{2.2.4}$$

Από τις σχέσεις (2.2.2), (2.2.3) βλέπουμε ότι
$f^{(k)}(\alpha) = P_n^{(k)}(\alpha) = k!\, a_k + (\alpha-\alpha)Q_{n-k-1}(x) = k!\, a_k$ και άρα $a_k = \frac{f^{(k)}(\alpha)}{k!}$.

Συνεπώς, το ζητούμενο πολυώνυμο είναι
$P_n(x) = f(\alpha) + \frac{f^{(1)}(\alpha)}{1!}(x-\alpha) + \frac{f^{(2)}(\alpha)}{2!}(x-\alpha)^2 + ... + \frac{f^{(n)}(\alpha)}{n!}(x-\alpha)^n$.
Το παραπάνω πολυώνυμο ονομάζεται πολυώνυμο Taylor, n βαθμού, κέντρου α, για τη συνάρτηση f, ενώ για την ειδική περίπτωση όπου α=0 ονομάζεται πολυώνυμο Maclaurin[8].

2.3 Θεώρημα Taylor

Η μέθοδος Δέλτα, η οποία θα αναλυθεί διεξοδικά στο επόμενο κεφάλαιο, αποδεικνύεται μέσω του Θεωρήματος Taylor. Το τελευταίο, επιτρέπει σε μια περίπλοκη συνάρτηση να προσεγγιστεί με τη βοήθεια του πολυωνύμου Taylor και δίνει μια έκφραση για το σφάλμα της προσέγγισης.

Παρατήρηση 2.3.1. Προκειμένου να εντοπίσουμε πότε μια συνάρτηση f ισούται με τη σειρά Taylor της f σε ένα διάστημα γύρω από το α, σχηματίζουμε τη διαφορά $R_n(x) = f(x) - P_n(x) = f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(\alpha)}{k!}(x-\alpha)^k$. Η διαφορά $R_n(x)$ καλείται συνάρτηση υπόλοιπο και η απόλυτη τιμή της λέγεται σφάλμα της προσέγγισης. Η συνάρτηση f προσεγγίζεται από το πολυώνυμο Taylor $P_n(x)$ με σφάλμα $R_n(x)$ για μια κατάλληλα μεγάλη τιμή του n.

Μια συνάρτηση f είναι ίση με τη σειρά Taylor της f σε ένα διάστημα γύρω από το α, όταν $R_n(x) \to 0$ καθώς $n \to \infty$.

Θεώρημα 2.3.1. ***(Ολοκληρωτική μορφή του υπολοίπου** $R_n(x)$ **)*** *Έστω η συνάρτηση $f \in C^{n+1}(I)$, όπου I ένα διάστημα το οποίο περιέχει το σημείο α. Τότε, $\forall x \in I$ ισχύει ότι*

$$R_n(x) = f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(\alpha)}{k!}(x-\alpha)^k = \frac{1}{n!}\int_{\alpha}^{x} (x-t)^n f^{(n+1)}(t)dt. \tag{2.3.1}$$

[8]Colin Maclaurin (1698–1746): Σουηδός μαθηματικός με έντονη συνεισφορά στην άλγεβρα και τη γεωμετρία.

Απόδειξη. Η απόδειξη κάνει χρήση της Μαθηματικής Επαγωγής.

- για n=1 εξ' ορισμού έχουμε

$$\begin{aligned} & f(x) = f(\alpha) + f'(\alpha)(x-\alpha) + R_1(x) \\ \Longleftrightarrow\ & R_1(x) = f(x) - f(\alpha) - f'(\alpha)(x-\alpha) \\ & = \int_\alpha^x f'(t)dt - \int_\alpha^x f'(\alpha)dt \\ & = \int_\alpha^x [f'(t)dt - f'(\alpha)]dt \\ & = \int_\alpha^x [f'(t)dt - f'(\alpha)]d(t-x) \\ & = [f'(x) - f'(\alpha)](t-x)|_\alpha^x - \int_\alpha^x (t-x)f''(t)dt \\ & = 0 + \int_\alpha^x (x-t)f^{(2)}(t)dt \\ & = \int_\alpha^x (x-t)f^{(2)}(t)dt. \end{aligned}$$

Καταλήξαμε ότι η (2.3.1) αληθεύει για n=1.

- Υποθέτουμε ότι η (2.3.1) ισχύει για n=k, δηλαδή

$$R_k(x) = \frac{1}{k!}\int_\alpha^x (x-t)f^{(k+1)}(t)dt \tag{2.3.2}$$

- Θα δείξουμε ότι η (2.3.1) ισχύει για n=k+1. Έχουμε,

$$\begin{aligned} & f(x) = \sum_{n=0}^{k+1} \tfrac{f^{(n)}(\alpha)}{n!}(x-\alpha)^n + R_{k+1}(x) \\ \Longleftrightarrow\ & R_{k+1}(x) = f(x) - \sum_{n=0}^{k+1} \tfrac{f^{(n)}(\alpha)}{n!}(x-\alpha)^n \\ & = f(x) - \sum_{n=0}^{k} \tfrac{f^{(n)}(\alpha)}{n!}(x-\alpha)^n - \tfrac{f^{(k+1)}(\alpha)}{(k+1)!}(x-\alpha)^{k+1} \\ & = R_k(x) - \tfrac{f^{(k+1)}(\alpha)}{(k+1)!}(x-\alpha)^{k+1} \\ & = \tfrac{1}{k!}\int_\alpha^x (x-t)f^{(k+1)}(t)dt - \tfrac{f^{(k+1)}(\alpha)}{(k+1)!}(x-\alpha)^{k+1} \\ & = \tfrac{1}{k!}\int_\alpha^x (x-t)^k f^{(k+1)}(t)dt - \tfrac{1}{k!}\int_\alpha^x f^{(k+1)}(\alpha)(x-t)^k dt \\ & = \tfrac{1}{k!}\int_\alpha^x (x-t)^k [f^{(k+1)}(t) - f^{(k+1)}(\alpha)]dt \\ & = \tfrac{1}{k!}\int_\alpha^x [f^{(k+1)}(t) - f^{(k+1)}(\alpha)]d(\tfrac{-(x-t)^{k+1}}{k+1}) \\ & = \tfrac{1}{k!}[f^{(k+1)}(t) - f^{(k+1)}(\alpha)](\tfrac{-(x-t)^{k+1}}{k+1})|_\alpha^x - \tfrac{1}{k!}\int_\alpha^x (\tfrac{-(x-t)^{k+1}}{k+1})f^{(k+2)}(t)dt \\ & = 0 + \tfrac{1}{(k+1)!}\int_\alpha^x (x-t)^{k+1} f^{(k+2)}(t)dt \\ & = \tfrac{1}{(k+1)!}\int_\alpha^x (x-t)^{k+1} f^{(k+2)}(t)dt. \end{aligned}$$

Η (2.3.1) αληθεύει για n=k+1. Συνεπώς, η απόδειξη ολοκληρώθηκε.

□

Παρατήρηση 2.3.2. Ο ολοκληρωτικός τύπος (2.3.1) προϋποθέτει ότι $f \in C^{n+1}(I)$, όπου I ένα διάστημα που περιέχει το α, ενώ για την ανάπτυξή της σε σειρά Taylor αρκεί $f \in C^n(I)$. Έτσι,ο παραπάνω τύπος είναι περιοριστικός, αφού απαιτεί την ύπαρξη μιας επιπλέον παραγώγου. Αντίθετα, ο παρακάτω ολοκληρωτικός τύπος είναι εφαρμόσιμος για συναρτήσεις $f \in C^n(I)$.

Θεώρημα 2.3.2. ***(Ολοκληρωτική μορφή του υπολοίπου*** $R_n(x)$ ***)*** *Έστω η συνάρτηση* $f \in C^n(I)$*, όπου I ένα διάστημα το οποίο περιέχει το σημείο α. Τότε,* $\forall x \in I$ *ισχύει ότι*

$$R_n(x) = f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-\alpha)^k = \frac{1}{(n-1)!}\int_{\alpha}^{x}(x-t)^{n-1}[f^{(n)}(t) - f^{(n)}(\alpha)]dt. \tag{2.3.3}$$

Απόδειξη. Θα δείξουμε ότι

$$f(x) - \sum_{k=0}^{n} \tfrac{f^{(k)}(\alpha)}{k!}(x-\alpha)^k = \tfrac{1}{(n-1)!}\int_{\alpha}^{x}(x-t)^{n-1}[f^{(n)}(t) - f^{(n)}(\alpha)]dt$$

$$\Leftrightarrow f(x) - \sum_{k=0}^{n} \tfrac{f^{(k)}(a)}{k!}(x-\alpha)^k = \tfrac{1}{(n-1)!}\int_{\alpha}^{x}(x-t)^{n-1}f^{(n)}(t)dt$$
$$-\tfrac{1}{(n-1)!}\int_{\alpha}^{x}(x-t)^{n-1}f^{(n)}(\alpha)dt$$

$$\overset{\Theta\varepsilon\omega\rho.2.3.1.}{\Leftrightarrow} f(x) - \sum_{k=0}^{n} \tfrac{f^{(k)}(\alpha)}{k!}(x-\alpha)^k = f(x) - \sum_{k=0}^{n-1} \tfrac{f^{(k)}(\alpha)}{k!}(x-\alpha)^k$$
$$-\tfrac{1}{(n-1)!}\int_{\alpha}^{x}(x-t)^{n-1}f^{(n)}(\alpha)dt$$

$$\Leftrightarrow \tfrac{f^{(n)}(\alpha)}{n!}(x-\alpha)^n = \tfrac{1}{(n-1)!}\int_{\alpha}^{x}(x-t)^{n-1}f^{(n)}(\alpha)dt.$$

Πράγματι έχουμε,

$$\begin{aligned}\tfrac{1}{(n-1)!}\int_{\alpha}^{x}(x-t)^{n-1}f^{(n)}(\alpha)dt &= \tfrac{f^{(n)}(\alpha)}{(n-1)!}\int_{\alpha}^{x}(x-t)^{n-1}dt \\ &= \tfrac{f^{(n)}(\alpha)}{(n-1)!}\left(-\tfrac{(x-t)^n}{n}\right)\Big|_{\alpha}^{x} = \tfrac{f^{(n)}(\alpha)}{(n-1)!}\tfrac{(x-\alpha)^n}{n} \\ &= \tfrac{f^{(n)}(\alpha)}{n!}(x-\alpha)^n.\end{aligned}$$

□

2.3.1 Μονοδιάστατη Μορφή Θεωρήματος Taylor

Θεώρημα 2.3.3. ***(Θεώρημα Taylor)****Έστω συνάρτηση* $f \in C^n(I)$*, όπου I ένα ανοικτό διάστημα που περιέχει το κλειστό διάστημα* $[\alpha, x]$*. Τότε, υπάρχει σημείο* $\xi = \xi(x) \in (\alpha, x)$*, τέτοιο ώστε να ικανοποιείται η παρακάτω ισότητα*

$$f(x) = f(\alpha) + \frac{f^{(1)}(\alpha)}{1!}(x-\alpha) + \frac{f^{(2)}(\alpha)}{2!}(x-\alpha)^2 + ... + \frac{f^{(n-1)}(\alpha)}{(n-1)!}(x-\alpha)^{(n-1)} + \frac{f^{(n)}(\xi)}{n!}(x-\alpha)^n. \tag{2.3.4}$$

Επιπλέον, αληθεύει ότι $\lim\limits_{x\to\alpha} \frac{f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(\alpha)}{k!}(x-\alpha)^k}{(x-\alpha)^n} = 0.$

Απόδειξη. Θα αποδείξουμε την περίπτωση όπου x>α.
Θέτουμε $m = min\{f^{(n)}(t) : \alpha \leq t \leq x\}$ και $M = max\{f^{(n)}(t) : \alpha \leq t \leq x\}$.
Τότε, προκύπτει ότι

$$m\int_\alpha^x (x-t)^{n-1}dt \leq \int_\alpha^x (x-t)^{n-1} f^{(n)}(t)dt \leq M\int_\alpha^x (x-t)^{n-1}dt$$

και επειδή $\int_\alpha^x (x-t)^{n-1}dt = \frac{(x-\alpha)^n}{n}$ παίρνουμε

$$m\frac{(x-\alpha)^n}{n} \leq \int_\alpha^x (x-t)^{n-1} f^{(n)}(t)dt \leq M\frac{(x-\alpha)^n}{n}$$

$$\Rightarrow \frac{1}{(n-1)!} m\frac{(x-\alpha)^n}{n} \leq \frac{1}{(n-1)!}\int_\alpha^x (x-t)^{n-1} f^{(n)}(t)dt \leq \frac{1}{(n-1)!} M\frac{(x-\alpha)^n}{n}$$

$$\Rightarrow m\frac{(x-\alpha)^n}{n!} \leq \frac{1}{(n-1)!}\int_\alpha^x (x-t)^{n-1} f^{(n)}(t)dt \leq M\frac{(x-\alpha)^n}{n!}$$

$$\Rightarrow m\frac{(x-\alpha)^n}{n!} \leq R_{n-1}(x) \leq M\frac{(x-\alpha)^n}{n!}$$

Εφαρμόζοντας το Θεώρημα ενδιαμέσων τιμών, υπάρχει σημείο $\xi \in (\alpha, x)$, τέτοιο ώστε $R_{n-1}(x) = \frac{f^{(n)}(\xi)}{n!}(x-\alpha)^n$. Έτσι καταλήγουμε ότι

$$f(x) = f(\alpha) + \frac{f^{(1)}(\alpha)}{1!}(x-\alpha) + ... + \frac{f^{(n-1)}(\alpha)}{(n-1)!}(x-\alpha)^{(n-1)} + R_{n-1}(x)$$

$$= f(\alpha) + \frac{f^{(1)}(\alpha)}{1!}(x-\alpha) + ... + \frac{f^{(n-1)}(\alpha)}{(n-1)!}(x-\alpha)^{(n-1)} + \frac{f^{(n)}(\xi)}{n!}(x-\alpha)^n.$$

Για το δεύτερο σκέλος του θεωρήματος εργαζόμαστε ως εξής: Αποδείξαμε ότι

$$f(x) = f(\alpha) + \frac{f^{(1)}(\alpha)}{1!}(x-\alpha) + ... + \frac{f^{(n-1)}(\alpha)}{(n-1)!}(x-\alpha)^{(n-1)} + \frac{f^{(n)}(\xi)}{n!}(x-\alpha)^n$$

$$= \sum_{k=0}^{n-1} \frac{f^{(k)}(\alpha)}{k!}(x-\alpha)^k + \frac{f^{(n)}(\xi)}{n!}(x-\alpha)^n$$

$$= \sum_{k=0}^{n} \frac{f^{(k)}(\alpha)}{k!}(x-\alpha)^k + \frac{f^{(n)}(\xi) - f^{(n)}(\alpha)}{n!}(x-\alpha)^n$$

$$\Rightarrow f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(\alpha)}{k!}(x-\alpha)^k = \frac{f^{(n)}(\xi) - f^{(n)}(\alpha)}{n!}(x-\alpha)^n.$$

Καθώς $x \to \alpha$ παρατηρούμε ότι $\xi \in (\alpha, x) \Rightarrow \alpha < \xi < x \to \alpha$

$$\Rightarrow \xi \to \alpha \Rightarrow f^{(n)}(\xi) - f^{(n)}(\alpha) \to 0. \text{ (f συνεχής)}$$

Επομένως, καταλήγουμε ότι

$$\lim_{x\to\alpha} \frac{f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(\alpha)}{k!}(x-\alpha)^k}{(x-\alpha)^n} = \lim_{x\to\alpha} \frac{f^{(n)}(\xi) - f^{(n)}(\alpha)}{n!(x-\alpha)^n}(x-\alpha)^n = \lim_{x\to\alpha} \frac{f^{(n)}(\xi) - f^{(n)}(\alpha)}{n!} = 0.$$

□

Παρατήρηση 2.3.3. Σύμφωνα με το Θεώρημα Taylor και χρησιμοποιώντας το πολυώνυμο Taylor προκύπτει ότι $f(x) = P_{n-1}(x) + R_{n-1}(x)$, όπου το υπόλοιπο $R_{n-1}(x) = \frac{f^n(\xi)}{n!}(x-\alpha)^n$ ονομάζεται υπόλοιπο Lagrange.

2.3.2 Πολυδιάστατη Μορφή Θεωρήματος Taylor

Θεώρημα 2.3.4. ***(Πολυδιάστατο Θεώρημα Taylor)*** *Θεωρούμε ένα ανοικτό σύνολο $D \subset \mathbb{R}^k$, το οποίο περιέχει το δίσκο $D(\boldsymbol{\alpha}, r)$ κέντρου $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, ..., \alpha_k)$*

και ακτίνας $r > 0$. *Έστω* $f : D \to \mathbb{R}$ *συνάρτηση τέτοια ώστε* $f \in C^{(n)}(D)$. *Αν θεωρήσουμε* j_i *το πλήθος των παραγωγίσεων ως προς τη μεταβλητή* x_i, *οι μερικές παράγωγοι της f, δηλαδή για* $\boldsymbol{x} = (x_1, x_2, \dots, x_k) \in D$ *οι συναρτήσεις* $f_{j_1,j_2,\dots,j_k}(\boldsymbol{x}) = f_{j_1,j_2,\dots,j_k}(x_1, x_2, \dots, x_k) = \frac{\partial^{j_1+j_2+\dots+j_k} f(\boldsymbol{x})}{\partial x_1^{j_1}, \partial x_2^{j_2}, \dots, \partial x_k^{j_k}}$ *είναι συνεχείς,* $\forall j_1, j_2, \dots, j_k \in \{0, 1, \dots\}$, *όπου* $j = j_1 + j_2 + \dots + j_k \leq n$. *Τότε, ορίζοντας για* $j = 0, 1, \dots, n$

$$
\begin{aligned}
&f^{(j)}(\boldsymbol{\alpha})(\boldsymbol{x}-\boldsymbol{\alpha})^j := D^j f(\boldsymbol{\alpha})(\boldsymbol{x}-\boldsymbol{\alpha})^j \\
&= \sum_{j_1+j_2+\dots+j_k=j} \frac{j!}{j_1! j_2! \dots j_k!} f_{j_1,j_2,\dots,j_k}(\boldsymbol{\alpha})(x_1-\alpha_1)^{j_1}(x_2-\alpha_2)^{j_2} \dots (x_k-\alpha_k)^{j_k},
\end{aligned}
$$

παίρνουμε
$f(\boldsymbol{x}) = f(\boldsymbol{\alpha}) + \frac{D^1 f(\boldsymbol{\alpha})}{1!}(\boldsymbol{x}-\boldsymbol{\alpha})^1 + \dots + \frac{D^{n-1} f(\boldsymbol{\alpha})}{(n-1)!}(\boldsymbol{x}-\boldsymbol{\alpha})^{n-1} + \frac{D^n f(\boldsymbol{\alpha}+\xi(\boldsymbol{x}-\boldsymbol{\alpha}))}{n!}(\boldsymbol{x}-\boldsymbol{\alpha})^n$, *και*

ισχύει ότι $\lim\limits_{\boldsymbol{x}\to\boldsymbol{\alpha}} \frac{f(\boldsymbol{x}) - f(\boldsymbol{\alpha}) - \frac{D^1 f(\boldsymbol{\alpha})}{1!}(\boldsymbol{x}-\boldsymbol{\alpha})^1 - \dots - \frac{D^n f(\boldsymbol{\alpha})}{n!}(\boldsymbol{x}-\boldsymbol{\alpha})^n}{\sum\limits_{i-1}^{k} |x_i - \alpha_i|^n} = \lim\limits_{\boldsymbol{x}\to\boldsymbol{\alpha}} \frac{f(\boldsymbol{x}) - \sum\limits_{i=0}^{n} \frac{D^i f(\boldsymbol{\alpha})}{i!}(\boldsymbol{x}-\boldsymbol{\alpha})^i}{\sum\limits_{i-1}^{k} |x_i - \alpha_i|^n} = 0.$

Παρατήρηση 2.3.4. Πριν προβούμε στην απόδειξη του Θεωρήματος, προκειμένου να επιτευχθεί η βαθύτερη κατανόησή του, θα παραθέσουμε τις ποσότητες $D^i f(\boldsymbol{\alpha})(\mathbf{x}-\boldsymbol{\alpha})^i$, για i=1,2 καθώς και το διδιάστατο θεώρημα Taylor.

Για το διδιάστατο θεώρημα Taylor, θεωρούμε δίσκο D=($\boldsymbol{\alpha}$,r) κέντρου $\boldsymbol{\alpha} = (\alpha_1, \alpha_2) \in \mathbb{R}$. Έστω j_1, j_2 το πλήθος των παραγωγίσεων ως προς τη μεταβλητή x_1 ή x_2 και έστω συνάρτηση $f : D \to \mathbb{R}, f \in C^n(D)$ με $f_{j_1,j_2} = \frac{\partial^{j_1+j_2} f(x_1,x_2)}{\partial x_1 \partial x_2}$ συνεχείς για $\mathbf{x} = (x_1, x_2) \in D, \forall j_1, j_2 \in \{0, 1, \dots\}$. Τότε, για $j = j_1 + j_2 \leq n$ ισχύει ότι $\lim\limits_{\mathbf{x}\to\boldsymbol{\alpha}} \frac{f(\mathbf{x}) - \sum\limits_{i=0}^{n} \frac{D^i f(\boldsymbol{\alpha})}{i!}(\mathbf{x}-\boldsymbol{\alpha})^i}{\sum\limits_{i-1}^{2} |x_i - \alpha_i|^n} = 0$, όπου

$$
\begin{aligned}
f^{(j)}(\boldsymbol{\alpha})(\mathbf{x}-\boldsymbol{\alpha})^j &= D^j f(\boldsymbol{\alpha})(\mathbf{x}-\boldsymbol{\alpha})^j \\
&= \sum_{j_1+j_2=j} \frac{j!}{j_1! + j_2!} f_{j_1,j_2}(\alpha_1, \alpha_2)(x_1-\alpha_1)^{j_1}(x_2-\alpha_2)^{j_2}.
\end{aligned}
$$

Δηλαδή, για $j = 1, 2, 3$ παίρνουμε

$$
\begin{aligned}
f^{(1)}(\boldsymbol{\alpha})(\mathbf{x}-\boldsymbol{\alpha})^1 &= D^1 f(\boldsymbol{\alpha})(\mathbf{x}-\boldsymbol{\alpha})^1 \\
&= f_{0,1}(\alpha_1, \alpha_2)(x_1-\alpha_1)^0(x_2-\alpha_2)^1 + f_{1,0}(\alpha_1, \alpha_2)(x_1-\alpha_1)^1(x_2-\alpha_2)^0 \\
&= f_{0,1}(\alpha_1, \alpha_2)(x_2-\alpha_2) + f_{1,0}(\alpha_1, \alpha_2)(x_1-\alpha_1). \\
f^{(2)}(\boldsymbol{\alpha})(\mathbf{x}-\boldsymbol{\alpha})^2 &= D^2 f(\boldsymbol{\alpha})(\mathbf{x}-\boldsymbol{\alpha})^2 \\
&= f_{0,2}(\alpha_1, \alpha_2)(x_1-\alpha_1)^0(x_2-\alpha_2)^2 \\
&\quad + 2f_{1,1}(\alpha_1, \alpha_2)(x_1-\alpha_1)^1(x_2-\alpha_2)^1 + f_{2,0}(\alpha_1, \alpha_2)(x_1-\alpha_1)^2(x_2-\alpha_2)^0 \\
&= f_{0,2}(\alpha_1, \alpha_2)(x_2-\alpha_2)^2 + 2f_{1,1}(\alpha_1, \alpha_2)(x_1-\alpha_1)(x_2-\alpha_2) \\
&\quad + f_{2,0}(\alpha_1, \alpha_2)(x_1-\alpha_1)^2. \\
f^{(3)}(\boldsymbol{\alpha})(\mathbf{x}-\boldsymbol{\alpha})^3 &= D^3 f(\boldsymbol{\alpha})(\mathbf{x}-\boldsymbol{\alpha})^3 \\
&= f_{0,3}(\alpha_1, \alpha_2)(x_1-\alpha_1)^0(x_2-\alpha_2)^3 + 3f_{1,2}(\alpha_1, \alpha_2)(x_1-\alpha_1)^1(x_2-\alpha_2)^2
\end{aligned}
$$

$$+3f_{2,1}(\alpha_1,\alpha_2)(x_1-\alpha_1)^2(x_2-\alpha_2)^1+f_{3,0}(\alpha_1,\alpha_2)(x_1-\alpha_1)^3(x_2-\alpha_2)^0$$
$$= f_{0,3}(\alpha_1,\alpha_2)(x_2-\alpha_2)^3 + 3f_{1,2}(\alpha_1,\alpha_2)(x_1-\alpha_1)(x_2-\alpha_2)^2$$
$$+3f_{2,1}(\alpha_1,\alpha_2)(x_1-\alpha_1)^2(x_2-\alpha_2) + f_{3,0}(\alpha_1,\alpha_2)(x_1-\alpha_1)^3.$$

Αντίστοιχα, στην πολυδιάστατη περίπτωση, για j=1,2 οι ποσότητες $f^{(j)}(\boldsymbol{\alpha})(\mathbf{x}-\boldsymbol{\alpha})^j$ υπολογίζονται ως εξής:

$$f^{(1)}(\boldsymbol{\alpha})(\mathbf{x}-\boldsymbol{\alpha})^1 = D^1 f(\boldsymbol{\alpha})(\mathbf{x}-\boldsymbol{\alpha})^1$$
$$= f_{1,0,\dots,0}(\boldsymbol{\alpha})(x_1-\alpha_1) + f_{0,1,\dots,0}(\boldsymbol{\alpha})(x_2-\alpha_2) + \dots$$
$$+f_{0,\dots,1,0,\dots,0}(\boldsymbol{\alpha})(x_i-\alpha_i) + \dots + f_{0,\dots,0,1}(\boldsymbol{\alpha})(x_k-\alpha_k)$$
$$= \tfrac{\partial f(\mathbf{x})}{\partial x_1}|_{\mathbf{x}=\boldsymbol{\alpha}}(x_1-\alpha_1) + \tfrac{\partial f(\mathbf{x})}{\partial x_2}|_{\mathbf{x}=\boldsymbol{\alpha}}(x_2-\alpha_2) + \dots + \tfrac{\partial f(\mathbf{x})}{\partial x_k}|_{\mathbf{x}=\boldsymbol{\alpha}}(x_k-\alpha_k)$$
$$= \sum_{i=1}^{k} \tfrac{\partial f(\mathbf{x})}{\partial x_i}|_{\mathbf{x}=\boldsymbol{\alpha}}(x_i-\alpha_i) = \begin{bmatrix} \tfrac{\partial f(\mathbf{x})}{\partial x_1}|_{\mathbf{x}=\boldsymbol{\alpha}} & \tfrac{\partial f(\mathbf{x})}{\partial x_2}|_{\mathbf{x}=\boldsymbol{\alpha}} & \cdots & \tfrac{\partial f(\mathbf{x})}{\partial x_k}|_{\mathbf{x}=\boldsymbol{\alpha}} \end{bmatrix} \begin{bmatrix} x_1-\alpha_1 \\ x_2-\alpha_2 \\ \dots \\ x_k-\alpha_k \end{bmatrix}$$
$$= [\nabla f(\boldsymbol{\alpha})]^T(\mathbf{x}-\boldsymbol{\alpha}).$$

Θα δείξουμε ότι $f^{(2)}(\boldsymbol{\alpha})(\mathbf{x}-\boldsymbol{\alpha})^2 = D^2 f(\boldsymbol{\alpha})(\mathbf{x}-\boldsymbol{\alpha})^2 = [\mathbf{x}-\boldsymbol{\alpha}]^T H(\boldsymbol{\alpha})[\mathbf{x}-\boldsymbol{\alpha}]$, όπου $H(\boldsymbol{\alpha}) = D^2 f(\boldsymbol{\alpha})$ πίνακας Hessian.

Πράγματι, έχουμε ότι

$$f^{(2)}(\boldsymbol{\alpha})(\mathbf{x}-\boldsymbol{\alpha})^2 = D^2 f(\boldsymbol{\alpha})(\mathbf{x}-\boldsymbol{\alpha})^2$$
$$= \sum_{j_1+\dots+j_k=2} \tfrac{2!}{j_1!\dots j_k!} f_{j_1,j_2,\dots,j_k}(x_1-\alpha_1)^{j_1}\dots(x_k-\alpha_k)^{j_k}$$
$$= \sum_{i=1}^{k} f_{0,\dots,2,\dots,0}(x_i-\alpha_i)^2 + \sum_{i=1}^{k}\sum_{i<j\leq k} 2f_{0,\dots,1,\dots,1,\dots,0}(x_i-\alpha_i)(x_j-\alpha_j)$$
$$= \sum_{i=1}^{k} \tfrac{\partial^2 f(\mathbf{x})}{\partial x_i^2}|_{\mathbf{x}=\boldsymbol{\alpha}}(x_i-\alpha_i)^2 + 2\sum_{i<j} \tfrac{\partial^2 f(\mathbf{x})}{\partial x_i \partial x_j}|_{\mathbf{x}=\boldsymbol{\alpha}}(x_i-\alpha_i)(x_j-\alpha_j).$$

Επιπροσθέτως,

$$[\mathbf{x}-\boldsymbol{\alpha}]^T H(\boldsymbol{\alpha})[\mathbf{x}-\boldsymbol{\alpha}] =$$
$$= \begin{bmatrix} x_1-\alpha_1 & \dots & x_k-\alpha_k \end{bmatrix} \begin{bmatrix} \tfrac{\partial^2 f(\mathbf{x})}{\partial x_1^2}|_{\mathbf{x}=\boldsymbol{\alpha}} & \tfrac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_2}|_{\mathbf{x}=\boldsymbol{\alpha}} & \cdots & \tfrac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_k}|_{\mathbf{x}=\boldsymbol{\alpha}} \\ \tfrac{\partial^2 f(\mathbf{x})}{\partial x_2 \partial x_1}|_{\mathbf{x}=\boldsymbol{\alpha}} & \tfrac{\partial^2 f(\mathbf{x})}{\partial x_2^2}|_{\mathbf{x}=\boldsymbol{\alpha}} & \cdots & \tfrac{\partial^2 f(\mathbf{x})}{\partial x_2 \partial x_k}|_{\mathbf{x}=\boldsymbol{\alpha}} \\ \dots & \dots & \dots & \dots \\ \tfrac{\partial^2 f(\mathbf{x})}{\partial x_k \partial x_1}|_{\mathbf{x}=\boldsymbol{\alpha}} & \tfrac{\partial^2 f(\mathbf{x})}{\partial x_k \partial x_2}|_{\mathbf{x}=\boldsymbol{\alpha}} & \cdots & \tfrac{\partial^2 f(\mathbf{x})}{\partial x_k^2}|_{\mathbf{x}=\boldsymbol{\alpha}} \end{bmatrix} \begin{bmatrix} x_1-\alpha_1 \\ x_2-\alpha_2 \\ \dots \\ x_k-\alpha_k \end{bmatrix}$$
$$= \begin{bmatrix} \sum_{i=1}^{k}(x_i-\alpha_i)\tfrac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_i}|_{\mathbf{x}=\boldsymbol{\alpha}} & \cdots & \sum_{i=1}^{k}(x_i-\alpha_i)\tfrac{\partial^2 f(\mathbf{x})}{\partial x_k \partial x_i}|_{\mathbf{x}=\boldsymbol{\alpha}} \end{bmatrix} \begin{bmatrix} x_1-\alpha_1 \\ x_2-\alpha_2 \\ \dots \\ x_k-\alpha_k \end{bmatrix}$$
$$= \sum_{i=1}^{k}(x_i-\alpha_i)\tfrac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_i}|_{\mathbf{x}=\boldsymbol{\alpha}}(x_1-\alpha_1) + \dots + \sum_{i=1}^{k}(x_i-\alpha_i)\tfrac{\partial^2 f(\mathbf{x})}{\partial x_k \partial x_i}|_{\mathbf{x}=\boldsymbol{\alpha}}(x_k-\alpha_k)$$
$$= \sum_{i=1}^{k} \tfrac{\partial^2 f(\mathbf{x})}{\partial x_i^2}|_{\mathbf{x}=\boldsymbol{\alpha}}(x_i-\alpha_i)^2 + 2\sum_{i<j} \tfrac{\partial^2 f(\mathbf{x})}{\partial x_i \partial x_j}|_{\mathbf{x}=\boldsymbol{\alpha}}(x_i-\alpha_i)(x_j-\alpha_j).$$

Συνεπώς, $f^{(2)}(\boldsymbol{\alpha})(\mathbf{x}-\boldsymbol{\alpha})^2 = D^2 f(\boldsymbol{\alpha})(\mathbf{x}-\boldsymbol{\alpha})^2 = [\mathbf{x}-\boldsymbol{\alpha}]^T H(\boldsymbol{\alpha})[\mathbf{x}-\boldsymbol{\alpha}]$.

Απόδειξη. Θεωρούμε τη συνάρτηση $h : I \to \mathbb{R}$, όπου I ένα ανοικτό διάστημα που περιέχει το [0,1], τέτοια ώστε $h(t) = f(\boldsymbol{\alpha} + t(\mathbf{x}-\boldsymbol{\alpha}))$. Παρατηρούμε ότι $h(0) = f(\boldsymbol{\alpha})$ και $h(1) = f(\mathbf{x})$. Από μονοδιάστατο θεώρημα Taylor για την h, για x=1 και α=0, υπάρχει $\xi \in (0,1)$ τέτοιο ώστε

$$h(1) = h(0) + \frac{h'(0)}{1!}1 + \frac{h''(0)}{2!}1^2 + ... + \frac{h^{(n-1)}(0)}{(n-1)!}1^{n-1} + \frac{h^n(\xi)}{n!}1^n. \quad (2.3.5)$$

Θα αποδείξουμε με μαθηματική επαγωγή ότι
$h^{(j)}(t) = D^j f(\boldsymbol{\alpha} + t(\mathbf{x}-\boldsymbol{\alpha}))(\mathbf{x}-\boldsymbol{\alpha})^j$.

Για n=1

$$\begin{aligned} h^{(1)}(t) &= \tfrac{\partial f(\boldsymbol{\alpha}+t(\mathbf{x}-\boldsymbol{\alpha}))}{\partial(\alpha_1+t(x_1-\alpha_1))}(x_1-\alpha_1) + ... + \tfrac{\partial f(\boldsymbol{\alpha}+t(\mathbf{x}-\boldsymbol{\alpha}))}{\partial(\alpha_k+t(x_k-\alpha_k))}(x_k-\alpha_k) \\ &= \sum_{i=1}^{k} \tfrac{\partial f(\boldsymbol{\alpha}+t(\mathbf{x}-\boldsymbol{\alpha}))}{\partial(\alpha_i+t(x_i-\alpha_i))}(x_i-\alpha_i) \\ &= D^1 f(\boldsymbol{\alpha} + t(\mathbf{x}-\boldsymbol{\alpha}))(\mathbf{x}-\boldsymbol{\alpha})^1. \end{aligned}$$

Έστω ότι ισχύει για n=j, δηλαδή

$$\begin{aligned} h^{(j)}(t) &= D^j f(\boldsymbol{\alpha} + t(\mathbf{x}-\boldsymbol{\alpha}))(\mathbf{x}-\boldsymbol{\alpha})^j \\ &= \sum_{j_1+...+j_k=j} \tfrac{j!}{j_1!...j_k!} f_{j_1,...,j_k}(\boldsymbol{\alpha} + t(\mathbf{x}-\boldsymbol{\alpha}))(x_1-\alpha_1)^{j_1} ... (x_k-\alpha_k)^{j_k}. \end{aligned}$$

Τότε, για n=j+1

$$\begin{aligned} h^{(j+1)}(t) &= [h^{(j)}(t)]' \\ &= \sum_{j_1+...+j_k=j} \tfrac{j!}{j_1!...j_k!}[f_{j_1,...,j_k}(\boldsymbol{\alpha} + t(\mathbf{x}-\boldsymbol{\alpha}))]'(x_1-\alpha_1)^{j_1} ... (x_k-\alpha_k)^{j_k} \\ &= \sum_{j_1+...+j_k=j} \tfrac{j!}{j_1!...j_k!}[\tfrac{\partial f_{j_1,...,j_k}(\boldsymbol{\alpha}+t(\mathbf{x}-\boldsymbol{\alpha}))}{\partial(\alpha_1+t(x_1-\alpha_1))}(x_1-\alpha_1) + ... + \tfrac{\partial f_{j_1,...,j_k}(\boldsymbol{\alpha}+t(\mathbf{x}-\boldsymbol{\alpha}))}{\partial(\alpha_k+t(x_k-\alpha_k))}(x_k-\alpha_k)] \\ &\qquad (x_1-\alpha_1)^{j_1} ... (x_k-\alpha_k)^{j_k} \\ &= \sum_{i=1}^{k}(x_i-\alpha_i) \sum_{j_1+...+j_k=j} \tfrac{j!}{j_1!...j_k!} f_{j_1,...,j_{i+1},...,j_k}(\boldsymbol{\alpha} + t(\mathbf{x}-\boldsymbol{\alpha})) \\ &\qquad (x_1-\alpha_1)^{j_1} ... (x_k-\alpha_k)^{j_k} \\ &= D^{j+1} f(\boldsymbol{\alpha} + t(\mathbf{x}-\boldsymbol{\alpha}))(\mathbf{x}-\boldsymbol{\alpha})^{j+1}. \end{aligned}$$

Για t=0 αληθεύουν
$h(0) = f(\boldsymbol{\alpha})$.
$h'(0) = D^1 f(\boldsymbol{\alpha})(\mathbf{x}-\boldsymbol{\alpha}) = [\nabla f(\boldsymbol{\alpha})]^T[\mathbf{x}-\boldsymbol{\alpha}]$.
$h^{(j)}(0) = D^j f(\boldsymbol{\alpha})(\mathbf{x}-\boldsymbol{\alpha})^j$.

Αντικαθιστώντας στην σχέση (2.3.5) καταλήγουμε ότι

$$\begin{aligned} f(\mathbf{x}) &= f(\boldsymbol{\alpha}) + \tfrac{D^1 f(\boldsymbol{\alpha})}{1!}(\mathbf{x}-\boldsymbol{\alpha}) + ... + \tfrac{D^{n-1} f(\boldsymbol{\alpha})}{(n-1)!}(\mathbf{x}-\boldsymbol{\alpha})^{n-1} \\ &\qquad + \tfrac{D^n f(\boldsymbol{\alpha}-\xi(\mathbf{x}-\boldsymbol{\alpha}))}{n!}(\mathbf{x}-\boldsymbol{\alpha})^n \\ &= \sum_{i=0}^{n-1} \tfrac{D^i f(\boldsymbol{\alpha})}{i!}(\mathbf{x}-\boldsymbol{\alpha})^i + \tfrac{D^n f(\boldsymbol{\alpha}-\xi(\mathbf{x}-\boldsymbol{\alpha}))}{n!}(\mathbf{x}-\boldsymbol{\alpha})^n \end{aligned}$$

$$= \sum_{i=0}^{n} \frac{D^i f(\boldsymbol{\alpha})}{i!}(\mathbf{x}-\boldsymbol{\alpha})^i + \frac{D^n f(\boldsymbol{\alpha}-\xi(\mathbf{x}-\boldsymbol{\alpha}))-D^n f(\boldsymbol{\alpha})}{n!}(\mathbf{x}-\boldsymbol{\alpha})^n$$

Επομένως,

$$f(x) - \sum_{i=0}^{n} \frac{D^i f(\boldsymbol{\alpha})}{i!}(\mathbf{x}-\boldsymbol{\alpha})^i = \frac{D^n f(\boldsymbol{\alpha}-\xi(\mathbf{x}-\boldsymbol{\alpha}))-D^n f(\boldsymbol{\alpha})}{n!}(\mathbf{x}-\boldsymbol{\alpha})^n$$

$$= \frac{\sum\limits_{j_1+\ldots+j_k=n} \frac{n!}{j_1!\ldots j_k!}[f_{j_1,j_2,\ldots,j_k}(\boldsymbol{\alpha}+\xi(\mathbf{x}-\boldsymbol{\alpha}))-f_{j_1,j_2,\ldots,j_k}(\boldsymbol{\alpha})]}{n!}(x_1-\alpha_1)^{j_1}\ldots(x_k-\alpha_k)^{j_k}$$

και άρα,

$$n!\,\frac{|f(x)-\sum\limits_{i=0}^{n}\frac{D^i f(\boldsymbol{\alpha})}{i!}(\mathbf{x}-\boldsymbol{\alpha})^i|}{\sum\limits_{i=1}^{k}|x_i-\alpha_i|^n}$$

$$= \frac{\sum\limits_{j_1+\ldots+j_k=n} \frac{n!}{j_1!\ldots j_k!}|f_{j_1,\ldots,j_k}(\boldsymbol{\alpha}+\xi(\mathbf{x}-\boldsymbol{\alpha}))-f_{j_1,\ldots,j_k}(\boldsymbol{\alpha})|}{\sum\limits_{i=1}^{k}|x_i-\alpha_i|^n}|x_1-\alpha_1|^{j_1}\ldots|x_k-\alpha_k|^{j_k}$$

$$\leq \frac{\sum\limits_{j_1+\ldots+j_k=n} \frac{n!}{j_1!\ldots j_k!}|x_1-\alpha_1|^{j_1}\ldots|x_k-\alpha_k|^{j_k}}{\sum\limits_{i=1}^{k}|x_i-\alpha_i|^n}\max_{j_1,\ldots,j_k}\{|f_{j_1,\ldots,j_k}(\boldsymbol{\alpha}+\xi(\mathbf{x}-\boldsymbol{\alpha}))-f_{j_1,\ldots,j_k}(\boldsymbol{\alpha})|\}.$$

Παρατηρούμε ότι η ποσότητα $\max\limits_{j_1,\ldots,j_k}\{|f_{j_1,\ldots,j_k}(\boldsymbol{\alpha}+\xi(\mathbf{x}-\boldsymbol{\alpha}))-f_{j_1,\ldots,j_k}(\boldsymbol{\alpha})|\} \to 0$, καθώς $\mathbf{x} \to \boldsymbol{\alpha}$.

Ακόμα, θα δείξουμε ότι η ποσότητα $\frac{\sum\limits_{j_1+\ldots+j_k=n} \frac{n!}{j_1!\ldots j_k!}|x_1-\alpha_1|^{j_1}\ldots|x_k-\alpha_k|^{j_k}}{\sum\limits_{i=1}^{k}|x_i-\alpha_i|^n}$ είναι φραγμένη. Έχουμε

$$\begin{aligned}
|x_1-\alpha_1|^{j_1}\ldots|x_k-\alpha_k|^{j_k} &\leq (\max\{|x_1-\alpha_1|,\ldots,|x_k-\alpha_k|\})^{j_1}\ldots \\
&\quad (\max\{|x_1-\alpha_1|,\ldots,|x_k-\alpha_k|\})^{j_k} \\
&= (\max\{|x_1-\alpha_1|,\ldots,|x_k-\alpha_k|\})^{j_1+\ldots+j_k} \\
&= (\max\{|x_1-\alpha_1|,\ldots,|x_k-\alpha_k|\})^{n} \\
&= (\max\{|x_1-\alpha_1|^n,\ldots,|x_k-\alpha_k|^n\}) \\
&\leq \sum_{i=1}^{k}|x_i-\alpha_i|^n.
\end{aligned}$$

$$\Rightarrow \sum_{j_1+\ldots+j_k=n}\frac{n!}{j_1!\ldots j_k!}|x_1-\alpha_1|^{j_1}\ldots|x_k-\alpha_k|^{j_k} \leq k^n\sum_{i=1}^{k}|x_i-\alpha_i|^n.$$

Καταληκτικά,

$$\frac{\sum\limits_{j_1+\ldots+j_k=n} \frac{n!}{j_1!\ldots j_k!}|x_1-\alpha_1|^{j_1}\ldots|x_k-\alpha_k|^{j_k}}{\sum\limits_{i=1}^{k}|x_i-\alpha_i|^n} \leq \frac{k^n\sum\limits_{i=1}^{k}|x_i-\alpha_i|^n}{\sum\limits_{i=1}^{k}|x_i-\alpha_i|^n} = k^n.$$

Τελικώς, δεδομένου ότι το όριο φραγμένης επί μηδενικής συνάρτησης είναι μηδέν, παίρνουμε

$$\lim_{\mathbf{x}\to\boldsymbol{\alpha}}\frac{f(\mathbf{x})-f(\boldsymbol{\alpha})-\frac{D^1 f(\boldsymbol{\alpha})}{1!}(\mathbf{x}-\boldsymbol{\alpha})^1-\ldots-\frac{D^n f(\boldsymbol{\alpha})}{n!}(\mathbf{x}-\boldsymbol{\alpha})^n}{\sum\limits_{i-1}^{k}|x_i-\alpha_i|^n} = \lim_{\mathbf{x}\to\boldsymbol{\alpha}}\frac{f(\mathbf{x})-\sum\limits_{i=0}^{n}\frac{D^i f(\boldsymbol{\alpha})}{i!}(\mathbf{x}-\boldsymbol{\alpha})^i}{\sum\limits_{i-1}^{k}|x_i-\alpha_i|^n} = 0. \qquad \square$$

Κεφάλαιο 3

Μέθοδος Δέλτα

Μία θεμελιώδης συνέπεια του θεωρήματος Taylor, το οποίο αναλύθηκε στο προηγούμενο κεφάλαιο, είναι η Μέθοδος Δέλτα. Πρόκειται για τη διαδικασία προσέγγισης μιας σύνθετης συνάρτησης από μια γραμμική συνάρτηση και τον υπολογισμό της αντίστοιχης διασποράς. Μέσω της μεθόδου δέλτα είναι ξεκάθαρο ότι συνάρτηση μιας ασυμπτωτικά κανονικής εκτιμήτριας ακολουθεί ασυμπτωτικά κανονική κατανομή. Η εν λόγω μέθοδος διατυπώθηκε από τον Harald Cramér[1] το 1946, στο βιβλίο του « Μαθηματικές μέθοδοι της Στατιστικής » και αποτελεί μια γενίκευση του Κεντρικού Οριακού Θεωρήματος. Ο τελευταίος ανέπτυξε τη μέθοδο για τη διδιάστατη συνάρτηση $g(m_{i,n}, m_{j,n})$, με μεταβλητές τις δειγματικές κεντρικές ροπές γύρω από τον μέσο, όμως κατέστησε σαφές ότι η ίδια τεχνική είναι εφαρμόσιμη και για συναρτήσεις με περισσότερες μεταβλητές. Εν συνεχεία, ο Hurt το 1976 επέκτεινε τα αποτελέσματα του Cramér εφαρμόζοντας τη μέθοδο και σε συναρτήσεις με μεταβλητές διάφορες των δειγματικών κεντρικών ροπών γύρω από τον μέσο. Η μέθοδος αυτή μας επιτρέπει τη μελέτη της ασυμπτωτικής συμπεριφοράς του διανύσματος των δειγματικών κεντρικών ροπών $(M_{1,n}, \ldots, M_{k,n})$, η οποία αναλύεται ενδελεχώς στο τέταρτο κεφάλαιο.

[1] Harald Cramér (1893-1985): Σουηδός μαθηματικός, χημικός, αναλογιστής, στατιστικός. Με σπουδές τόσο στα μαθηματικά όσο και στη χημεία, στο Πανεπιστήμιο της Στοκχόλμης, προτίμησε τελικώς τα πρώτα ως επιστημονικό κλάδο και ξεχώρισε με τη διπλωματική εργασία του στην οποία επιβλέπων ήταν ο μαθηματικός Marcel Riesz. Η αρχική του μαθηματική δραστηριότητα αναφέρεται στην αναλυτική θεωρία αριθμών, ενώ η εντατική ενασχόλησή του με τη θεωρία πιθανοτήτων δεν ξεκινά παρά στα τέλη της δεκαετίας των '20 με επιρροές από μαθηματικούς όπως οι Kolmogorov, Lévy, Bernstein, Khinchin. Παράλληλα, η επιστημονική του δραστηριότητα καλύπτει και άλλους τομείς όπως η αναλογιστική επιστήμη και η θεωρία κινδύνου. Ως καθηγητής του Πανεπιστημίου της Στοκχόλμης, υπήρξε ο υπεύθυνος καθηγητής των διδακτορικών διατριβών των Herman Wold και Kai Lai Chung.

3.1 Μέθοδος Δέλτα $1^{ης}$ τάξης

Μονοδιάστατη μέθοδος Δέλτα πρώτης τάξης από τον $\mathbb{R}$ στον $\mathbb{R}$

Θεώρημα 3.1.1. *Θεωρούμε ακολουθία τυχαίων μεταβλητών $(T_n)_{n\geq 1}$, η οποία είναι ασυμπτωτικά Κανονική εκτιμήτρια μιας παραμέτρου θ (βλ. Ορισμό 1.5.4).*
Δηλαδή, $\sqrt{n}(T_n - \theta) \xrightarrow{D} N(0, \sigma^2)$.
Τότε, ισχύει ότι $\sqrt{n}(f(T_n) - f(\theta)) \xrightarrow{D} N(0, \sigma^2[f'(\theta)]^2)$, εφ'όσον η $f'(\theta)$ υπάρχει και είναι διάφορη του μηδενός.

Απόδειξη. Από τη μονοδιάστατη μορφή του Θεωρήματος Taylor (βλ. Θεώρημα 2.3.3.), για x=T_n, αναπτύσσουμε τη συνάρτηση f σε σειρα Taylor πρώτης τάξης γύρω από το α=θ. Έχουμε, $f(T_n) = f(\theta) + \frac{f^{(1)}(\theta)}{1!}(T_n - \theta)^1 + R_1$,
όπου $R_1 = o_P((T_n - \theta)) \overset{Ορ.1.6.2.}{\Leftrightarrow} \frac{R_1}{T_n - \theta} \xrightarrow{P} 0$. Συνεπώς,

$$f(T_n) - f(\theta) = f'(\theta)(T_n - \theta) + R_1 \text{ και}$$

$$\sqrt{n}(f(T_n) - f(\theta)) = \sqrt{n}f'(\theta)(T_n - \theta) + \sqrt{n}R_1.$$

Όμως, από υπόθεση $\sqrt{n}(T_n - \theta) \xrightarrow{D} N(0, \sigma^2)$.
Άρα, $\sqrt{n}f'(\theta)(T_n - \theta) \xrightarrow{D} N(0, \sigma^2[f'(\theta)]^2)$.

Ακόμη, $\sqrt{n}R_1 = \sqrt{n}o_P((T_n - \theta)) = o_P(\sqrt{n}(T_n - \theta)) \Rightarrow \frac{\sqrt{n}R_1}{\sqrt{n}(T_n - \theta)} \xrightarrow{P} 0$.
Και από υπόθεση $\sqrt{n}(T_n - \theta) \xrightarrow{D} N(0, \sigma^2)$. Οπότε, από Θεώρημα 1.6.1. έπεται ότι $\sqrt{n}(T_n - \theta) = O_P(1)$. Επομένως, $\sqrt{n}R_1 \xrightarrow{P} 0$.
Τέλος, εφαρμόζοντας το Θεώρημα Slutsky (Θεώρημα 1.5.1.) παίρνουμε το ζητούμενο, δηλαδή $\sqrt{n}(f(T_n) - f(\theta)) \xrightarrow{D} N(0, \sigma^2[f'(\theta)]^2)$.

□

Πολυδιάστατη μέθοδος Δέλτα πρώτης τάξης από τον $\mathbb{R}^k$ στον $\mathbb{R}^s$

Θεώρημα 3.1.2. *Θεωρούμε ακολουθία $\boldsymbol{X}_n$ k-διάστατων τυχαίων διανυσμάτων τέτοια ώστε $\sqrt{n}(\boldsymbol{X}_n - \boldsymbol{\theta}) \xrightarrow{D} \boldsymbol{W}$, όπου $\boldsymbol{\theta} = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \dots \\ \theta_k \end{bmatrix} \in \mathbb{R}^k$ και $\boldsymbol{W} = \begin{bmatrix} W_1 \\ W_2 \\ \dots \\ W_k \end{bmatrix}$ k-διάστατο τυχαίο διάνυσμα.Έστω συνάρτηση $\boldsymbol{f} : \mathbb{R}^k \to \mathbb{R}^s$, με $\boldsymbol{f} = \begin{bmatrix} f_1 \\ f_2 \\ \dots \\ f_s \end{bmatrix}$, τέτοια ώστε $f_1, f_2, \dots, f_s$ έχουν συνεχείς μερικές παραγώγους πρώτης τάξης σε μια περιοχή γύρω από το $\boldsymbol{\theta}$. Τότε, ισχύει ότι $\sqrt{n}(\boldsymbol{f}(\boldsymbol{X}_n) - \boldsymbol{f}(\boldsymbol{\theta})) \xrightarrow{D} H(\boldsymbol{\theta})\boldsymbol{W}$, όπου $H(\boldsymbol{\theta})$ είναι (s x k) Hessian πίνακας. Δηλαδή,*

$$\sqrt{n} \begin{bmatrix} f_1(\boldsymbol{X}_n) - f_1(\boldsymbol{\theta}) \\ f_2(\boldsymbol{X}_n) - f_2(\boldsymbol{\theta}) \\ \dots \\ f_s(\boldsymbol{X}_n) - f_s(\boldsymbol{\theta}) \end{bmatrix} \xrightarrow{D} \begin{bmatrix} \frac{\partial f_1(\boldsymbol{x})}{\partial x_1}|_{\boldsymbol{x}=\boldsymbol{\theta}} & \frac{\partial f_1(\boldsymbol{x})}{\partial x_2}|_{\boldsymbol{x}=\boldsymbol{\theta}} & \cdots & \frac{\partial f_1(\boldsymbol{x})}{\partial x_k}|_{\boldsymbol{x}=\boldsymbol{\theta}} \\ \frac{\partial f_2(\boldsymbol{x})}{\partial x_1}|_{\boldsymbol{x}=\boldsymbol{\theta}} & \frac{\partial f_2(\boldsymbol{x})}{\partial x_2}|_{\boldsymbol{x}=\boldsymbol{\theta}} & \cdots & \frac{\partial f_2(\boldsymbol{x})}{\partial x_k}|_{\boldsymbol{x}=\boldsymbol{\theta}} \\ \dots & \dots & \dots & \dots \\ \frac{\partial f_s(\boldsymbol{x})}{\partial x_1}|_{\boldsymbol{x}=\boldsymbol{\theta}} & \frac{\partial f_s(\boldsymbol{x})}{\partial x_k}|_{\boldsymbol{x}=\boldsymbol{\theta}} & \cdots & \frac{\partial f_s(\boldsymbol{x})}{\partial x_k}|_{\boldsymbol{x}=\boldsymbol{\theta}} \end{bmatrix} \begin{bmatrix} W_1 \\ W_2 \\ \dots \\ W_k \end{bmatrix}$$

$$= \begin{bmatrix} \frac{\partial f_1(\boldsymbol{x})}{\partial x_1}|_{\boldsymbol{x}=\boldsymbol{\theta}} W_1 + \frac{\partial f_1(\boldsymbol{x})}{\partial x_2}|_{\boldsymbol{x}=\boldsymbol{\theta}} W_2 + \dots + \frac{\partial f_1(\boldsymbol{x})}{\partial x_k}|_{\boldsymbol{x}=\boldsymbol{\theta}} W_k \\ \frac{\partial f_2(\boldsymbol{x})}{\partial x_1}|_{\boldsymbol{x}=\boldsymbol{\theta}} W_1 + \frac{\partial f_2(\boldsymbol{x})}{\partial x_2}|_{\boldsymbol{x}=\boldsymbol{\theta}} W_2 + \dots + \frac{\partial f_2(\boldsymbol{x})}{\partial x_k}|_{\boldsymbol{x}=\boldsymbol{\theta}} W_k \\ \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \\ \frac{\partial f_s(\boldsymbol{x})}{\partial x_1}|_{\boldsymbol{x}=\boldsymbol{\theta}} W_1 + \frac{\partial f_s(\boldsymbol{x})}{\partial x_k}|_{\boldsymbol{x}=\boldsymbol{\theta}} W_2 + \dots + \frac{\partial f_s(\boldsymbol{x})}{\partial x_k}|_{\boldsymbol{x}=\boldsymbol{\theta}} W_k \end{bmatrix}.$$

Απόδειξη. Για την απόδειξη, θα κάνουμε χρήση του τεχνάσματος Cramer-Wold (βλ. Λήμμα 1.4.4.), σύμφωνα με το οποίο $\mathbf{X}_n \xrightarrow{D} \mathbf{X} \Leftrightarrow \mathbf{c}^T\mathbf{X}_n \xrightarrow{D} \mathbf{c}^T\mathbf{X}$, $\forall \mathbf{c} \in \mathbb{R}^k$.

Οπότε, για να δείξουμε ότι $\sqrt{n}(\mathbf{f}(\mathbf{X}_n) - \mathbf{f}(\boldsymbol{\theta})) \xrightarrow{D} H(\boldsymbol{\theta})\mathbf{W}$, αρκεί να δείξουμε ότι $\mathbf{c}^T(\sqrt{n}(\mathbf{f}(\mathbf{X}_n) - \mathbf{f}(\boldsymbol{\theta}))) \xrightarrow{D} \mathbf{c}^T H(\boldsymbol{\theta})\mathbf{W}, \forall c \in \mathbb{R}^s$

$$\Leftrightarrow \begin{bmatrix} c_1 & c_2 & \dots & c_k \end{bmatrix} \begin{bmatrix} \sqrt{n}(f_1(\mathbf{X}_n) - f_1(\boldsymbol{\theta})) \\ \sqrt{n}(f_2(\mathbf{X}_n) - f_2(\boldsymbol{\theta})) \\ \dots \\ \sqrt{n}(f_s(\mathbf{X}_n) - f_s(\boldsymbol{\theta})) \end{bmatrix} \xrightarrow{D} \begin{bmatrix} c_1 & c_2 & \dots & c_k \end{bmatrix} \begin{bmatrix} \sum_{i=1}^k \frac{\partial f_1(\mathbf{x})}{\partial x_i}|_{\mathbf{x}=\boldsymbol{\theta}} W_i \\ \sum_{i=1}^k \frac{\partial f_2(\mathbf{x})}{\partial x_i}|_{\mathbf{x}=\boldsymbol{\theta}} W_i \\ \dots \\ \sum_{i=1}^k \frac{\partial f_s(\mathbf{x})}{\partial x_i}|_{\mathbf{x}=\boldsymbol{\theta}} W_i \end{bmatrix}$$

$$\Leftrightarrow c_1[\sqrt{n}(f_1(\mathbf{X}_n) - f_1(\boldsymbol{\theta}))] + \dots + c_s[\sqrt{n}(f_s(\mathbf{X}_n) - f_s(\boldsymbol{\theta}))] \xrightarrow{D} c_1 \sum_{i=1}^k \frac{\partial f_1(\mathbf{x})}{\partial x_i}|_{\mathbf{x}=\boldsymbol{\theta}} W_i + \dots + c_s \sum_{i=1}^k \frac{\partial f_s(\mathbf{x})}{\partial x_i}|_{\mathbf{x}=\boldsymbol{\theta}} W_i$$

Θεωρούμε τη συνάρτηση $f(\mathbf{x}) = c_1 f_1(\mathbf{x}) + c_2 f_2(\mathbf{x}) + ... + c_s f_s(\mathbf{x})$. Άρα προκύπτει $\frac{\partial f(\mathbf{x})}{\partial x_i} = c_1 \frac{\partial f_1(\mathbf{x})}{\partial x_i} + c_2 \frac{\partial f_2(\mathbf{x})}{\partial x_i} + ... + c_s \frac{\partial f_s(\mathbf{x})}{\partial x_i} = \sum_{j=1}^{s} c_j \frac{\partial f_j(\mathbf{x})}{x_i}$.

Από την πολυδιάστατη μέθοδο Δέλτα ανώτερης τάξης από τον $\mathbb{R}^k$ στον $\mathbb{R}$, η οποία παρατίθεται και αποδεικνύεται στην παράγραφο 3.3 (Θεώρημα 3.3.1.) έχουμε ότι αν
$\sqrt{n}(\mathbf{X}_n - \boldsymbol{\theta}) \xrightarrow{D} \mathbf{W}$, τότε παίρνουμε
$n^{\frac{m}{2}}(f(\mathbf{X}_n) - f(\mathbf{X})) \xrightarrow{D} \frac{1}{m!}[D^m f(\boldsymbol{\theta})]\mathbf{W}^m = \sum_{j_1+...+j_k=m} f_{j_1,...,j_k}(\boldsymbol{\theta}) \frac{W_1^{j_1}}{j_1} ... \frac{W_k^{j_k}}{j_k}$.

Άρα, αφού από υπόθεση $\sqrt{n}(\mathbf{X}_n - \boldsymbol{\theta}) \xrightarrow{D} \mathbf{W}$ και f μια φορά διαφορίσιμη ως προς τη μεταβλητή $x_i, \forall i = 1, ..., k$ εφαρμόζοντας την πολυδιάστατη μέθοδο Δέλτα ανώτερης τάξης από τον $\mathbb{R}^k$ στον $\mathbb{R}$ για τη συνάρτηση f, για m=1, προκύπτει ότι $\sqrt{n}(f(\mathbf{X}_n) - f(\mathbf{X})) \xrightarrow{D} [D^1 f(\boldsymbol{\theta})]\mathbf{W}^1 = [\nabla f(\boldsymbol{\theta})]^T \mathbf{W}$

$$\begin{aligned}
&= \begin{bmatrix} \frac{\partial f(\mathbf{x})}{x_1}|_{\mathbf{x}=\boldsymbol{\theta}} & \frac{\partial f(\mathbf{x})}{x_2}|_{\mathbf{x}=\boldsymbol{\theta}} & \cdots & \frac{\partial f(\mathbf{x})}{x_k}|_{\mathbf{x}=\boldsymbol{\theta}} \end{bmatrix} \begin{bmatrix} W_1 \\ W_2 \\ \cdots \\ W_k \end{bmatrix} \\
&= \frac{\partial f(\mathbf{x})}{x_1}|_{\mathbf{x}=\boldsymbol{\theta}} W_1 + \frac{\partial f(\mathbf{x})}{x_2}|_{\mathbf{x}=\boldsymbol{\theta}} W_2 + ... + \frac{\partial f(\mathbf{x})}{x_k}|_{\mathbf{x}=\boldsymbol{\theta}} W_k \\
&= \sum_{i=1}^{k} \frac{\partial f(\mathbf{x})}{x_i}|_{\mathbf{x}=\boldsymbol{\theta}} W_i \\
&= (c_1 \frac{\partial f_1(\mathbf{x})}{\partial x_1}|_{\mathbf{x}=\boldsymbol{\theta}} + ... + c_s \frac{\partial f_s(\mathbf{x})}{\partial x_1}|_{\mathbf{x}=\boldsymbol{\theta}}) W_1 + ... \\
&\quad + (c_1 \frac{\partial f_1(\mathbf{x})}{\partial x_k}|_{\mathbf{x}=\boldsymbol{\theta}} + ... + c_s \frac{\partial f_s(\mathbf{x})}{\partial x_k}|_{\mathbf{x}=\boldsymbol{\theta}}) W_k \\
&= c_1 \sum_{i=1}^{k} \frac{\partial f_1(\mathbf{x})}{\partial x_i}|_{\mathbf{x}=\boldsymbol{\theta}} W_i + ... + c_k \sum_{i=1}^{k} \frac{\partial f_s(\mathbf{x})}{\partial x_i}|_{\mathbf{x}=\boldsymbol{\theta}} W_i.
\end{aligned}$$

Όμως,

$$\begin{aligned}
\sqrt{n}[f(\mathbf{X}_n) - f(\mathbf{X})) &= \sqrt{n}[c_1 f_1(\mathbf{X}_n) + ... + c_s f_s(\mathbf{X}_n) - c_1 f_1(\mathbf{X}) + ... + c_s f_s(\mathbf{X})] \\
&= \sqrt{n}[c_1(f_1(\mathbf{X}_n) - f_1(\mathbf{X})) + ... + c_s(f_s(\mathbf{X}_n) - f_s(\mathbf{X}))] \\
&= \sum_{i=1}^{s} c_i[\sqrt{n}(f_i(\mathbf{X}_n) - f_i(\mathbf{X}))]
\end{aligned}$$

Καταλήξαμε ότι

$$\sum_{i=1}^{s} c_i[\sqrt{n}((f_i(\mathbf{X}_n) - f_i(\mathbf{X})))] \xrightarrow{D} c_1 \sum_{i=1}^{k} \frac{\partial f_1(\mathbf{x})}{\partial x_i}|_{\mathbf{x}=\boldsymbol{\theta}} W_i + ... + c_k \sum_{i=1}^{k} \frac{\partial f_s(\mathbf{x})}{\partial x_i}|_{\mathbf{x}=\boldsymbol{\theta}} W_i,$$

ή ισοδύναμα, $\forall c \in \mathbb{R}^k$,

$$\mathbf{c}^T(\sqrt{n}(\mathbf{f}(\mathbf{X}_n) - \mathbf{f}(\boldsymbol{\theta}))) \xrightarrow{D} \mathbf{c}^T H(\boldsymbol{\theta})\mathbf{W} \overset{Cr-Wold}{\Leftrightarrow} \sqrt{n}(\mathbf{f}(\mathbf{X}_n) - \mathbf{f}(\boldsymbol{\theta})) \xrightarrow{D} H(\boldsymbol{\theta})\mathbf{W}.$$

□

3.2 Μέθοδος Δέλτα $2^{ης}$ τάξης

Μονοδιάστατη μέθοδος Δέλτα δεύτερης τάξης από τον $\mathbb{R}$ στον $\mathbb{R}$

Θεώρημα 3.2.1. *Θεωρούμε ακολουθία τυχαίων μεταβλητών $(T_n)_{n\geq 1}$, η οποία είναι ασυμπτωτικά Κανονική εκτιμήτρια μιας παραμέτρου θ (βλ. Ορισμό 1.5.4). Δηλαδή, $\sqrt{n}(T_n-\theta) \xrightarrow{D} N(0,\sigma^2)$.*
Τότε, ισχύει ότι $n(f(T_n)-f(\theta)) \xrightarrow{D} \frac{\sigma^2 f''(\theta)}{2}\chi_1^2$, εφ' όσον η συνάρτηση $f:\mathbb{R}\to\mathbb{R}$ είναι Borel-μετρήσιμη συνάρτηση τέτοια ώστε να υπάρχουν οι παράγωγοι της f έως και δεύτερης τάξης, με $f'(\theta)=0$ και $f''(\theta)\neq 0$.

Απόδειξη. Από τη μονοδιάστατη μορφή του Θεωρήματος Taylor (βλ. Θεώρημα 2.3.3.), για x=T_n, αναπτύσσουμε τη συνάρτηση f σε σειρά Taylor δεύτερης τάξης γύρω από το α=θ. Έχουμε, $f(T_n)=f(\theta)+\frac{f^{(1)}(\theta)}{1!}(T_n-\theta)^1+\frac{f^{(2)}(\theta)}{2!}(T_n-\theta)^2+R_2$, όπου $R_2=o_P((T_n-\theta^2)) \overset{O\rho.1.6.2.}{\Leftrightarrow} \frac{R_2}{(T_n-\theta)^2}\xrightarrow{P} 0$. Συνεπώς, αφού από υπόθεση $f'(\theta)=0$ έπεται ότι $f(T_n)=f(\theta)+\frac{f''(\theta)}{2}(T_n-\theta)^2+R_2$. Δηλαδή,

$$f(T_n)-f(\theta)=\frac{f''(\theta)}{2}(T_n-\theta)^2+R_2 \text{ και}$$

$$n(f(T_n)-f(\theta))=n\frac{f''(\theta)}{2}(T_n-\theta)^2+nR_2.$$

Όμως, από υπόθεση $\sqrt{n}(T_n-\theta)\xrightarrow{D} N(0,\sigma^2)$.

$$\Rightarrow \frac{\sqrt{n}(T_n-\theta)}{\sigma}\xrightarrow{D} N(0,1)$$

$\Rightarrow (\frac{\sqrt{n}(T_n-\theta)}{\sigma})^2\xrightarrow{D} N^2(0,1)\equiv\chi_1^2$, όπου χ_1^2 είναι η χ-κατανομή με ένα βαθμό ελευθερίας.

Επομένως, προκύπτει ότι $n\frac{f''(\theta)}{2}(T_n-\theta)^2=\frac{f''(\theta)}{2}\sigma^2\frac{n(T_n-\theta)^2}{\sigma^2}\xrightarrow{D}\frac{\sigma^2 f''(\theta)}{2}\chi_1^2$.

Επιπροσθέτως, $nR_2=no_P((T_n-\theta^2))=o_P(n(T_n-\theta^2))\Leftrightarrow\frac{nR_2}{n(T_n-\theta)^2}\xrightarrow{P}0$

Από υπόθεση $\sqrt{n}(T_n-\theta)\xrightarrow{D}N(0,\sigma^2)\Rightarrow n(T_n-\theta)^2\xrightarrow{D}\sigma^2\chi_1^2$. Άρα, από Θεώρημα 1.6.1. η $n(T_n-\theta)^2$ είναι φραγμένη κατά πιθανότητα, ή ισοδύναμα $n(T_n-\theta)^2=O_P(1)$. Οπότε, $nR_2\xrightarrow{P}0$.
Συνοψίζοντας, με εφαρμογή του Θεωρήματος Slutsky (Θεώρημα 1.5.1.) προκύπτει το ζητούμενο, $n(f(T_n)-f(\theta))\xrightarrow{D}\frac{\sigma^2 f''(\theta)}{2}\chi_1^2$.

□

Πολυδιάστατη μέθοδος Δέλτα δεύτερης τάξης από τον $\mathbb{R}^k$ στον $\mathbb{R}$

Θεώρημα 3.2.2. *Θεωρούμε ακολουθία $\boldsymbol{X}_n$ k-διάστατων τυχαίων διανυσμάτων τέτοια ώστε $\sqrt{n}(\boldsymbol{X}_n - \boldsymbol{\theta}) \xrightarrow{D} \boldsymbol{W}$, όπου $\boldsymbol{\theta} = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \dots \\ \theta_k \end{bmatrix} \in \mathbb{R}^k$, $\boldsymbol{W} = \begin{bmatrix} W_1 \\ W_2 \\ \dots \\ W_k \end{bmatrix} \in \mathbb{R}^k$. Έστω $f : \mathbb{R}^k \to \mathbb{R}$ μια Borel-μετρήσιμη συνάρτηση με συνεχείς μερικές παραγώγους έως δεύτερης τάξης γύρω από το $\boldsymbol{\theta}$. Εάν $[\nabla f(\boldsymbol{\theta})] = \boldsymbol{0}$, τότε αληθεύει ότι*

$n(f(\boldsymbol{X}_n) - f(\boldsymbol{\theta})) \xrightarrow{D} \frac{1}{2}\boldsymbol{W}^T H_k(\boldsymbol{\theta})\boldsymbol{W}$, όπου θέσαμε $[\nabla f(\boldsymbol{\theta})] = \begin{bmatrix} \frac{\partial f(\boldsymbol{x})}{x_1}|_{\boldsymbol{x}=\boldsymbol{\theta}} \\ \frac{\partial f(\boldsymbol{x})}{x_2}|_{\boldsymbol{x}=\boldsymbol{\theta}} \\ \dots \\ \frac{\partial f(\boldsymbol{x})}{x_k}|_{\boldsymbol{x}=\boldsymbol{\theta}} \end{bmatrix}$ και

$H_k(\boldsymbol{\theta})$ ο (k x k) Hessian πίνακας : $H_k(\boldsymbol{\theta}) = \begin{bmatrix} \frac{\partial^2 f(\boldsymbol{x})}{\partial x_1^2}|_{\boldsymbol{x}=\boldsymbol{\theta}} & \frac{\partial^2 f(\boldsymbol{x})}{\partial x_1 \partial x_2}|_{\boldsymbol{x}=\boldsymbol{\theta}} & \cdots & \frac{\partial^2 f(\boldsymbol{x})}{\partial x_1 \partial x_k}|_{\boldsymbol{x}=\boldsymbol{\theta}} \\ \frac{\partial^2 f(\boldsymbol{x})}{\partial x_2 \partial x_1}|_{\boldsymbol{x}=\boldsymbol{\theta}} & \frac{\partial^2 f(\boldsymbol{x})}{\partial x_2^2}|_{\boldsymbol{x}=\boldsymbol{\theta}} & \cdots & \frac{\partial^2 f(\boldsymbol{x})}{\partial x_2 \partial x_k}|_{\boldsymbol{x}=\boldsymbol{\theta}} \\ \dots & \dots & \dots & \dots \\ \frac{\partial^2 f(\boldsymbol{x})}{\partial x_k \partial x_1}|_{\boldsymbol{x}=\boldsymbol{\theta}} & \frac{\partial^2 f(\boldsymbol{x})}{\partial x_k \partial x_2}|_{\boldsymbol{x}=\boldsymbol{\theta}} & \cdots & \frac{\partial^2 f(\boldsymbol{x})}{\partial x_k^2}|_{\boldsymbol{x}=\boldsymbol{\theta}} \end{bmatrix}$.

Απόδειξη. Από Πολυδιάστατη μορφή Θεωρήματος Taylor(βλ. Θεώρημα 2.3.4.), αναπτύσσουμε τη συνάρτηση f σε σειρά Taylor δεύτερης τάξης γύρω από το α=θ. Παίρνουμε $f(\mathbf{X}_n) = f(\boldsymbol{\theta}) + \frac{D^1(\boldsymbol{\theta})}{1!}(\mathbf{X}_n - \boldsymbol{\theta})^1 + \frac{D^2(\boldsymbol{\theta})}{2!}(\mathbf{X}_n - \boldsymbol{\theta})^2 + R_2$.

Από Παρατήρηση 2.3.4. αποδείξαμε ότι
$D^1 f(\boldsymbol{\alpha})(\mathbf{x} - \boldsymbol{\alpha}) = [\nabla f(\boldsymbol{\alpha})]^T(\mathbf{x} - \boldsymbol{\alpha})$ και
$D^2 f(\boldsymbol{\alpha})(\mathbf{x} - \boldsymbol{\alpha})^2 = [\mathbf{x} - \boldsymbol{\alpha}]^T H(\boldsymbol{\alpha})[\mathbf{x} - \boldsymbol{\alpha}]$.
Άρα έχουμε $f(\mathbf{X}_n) = f(\boldsymbol{\theta}) + [\nabla f(\boldsymbol{\theta})]^T(\mathbf{X}_n - \boldsymbol{\theta}) + \frac{1}{2}[\mathbf{X}_n - \boldsymbol{\theta}]^T H(\boldsymbol{\theta})[\mathbf{X}_n - \boldsymbol{\theta}] + R_2$, όπου $R_2 = o_P((\mathbf{X}_n - \boldsymbol{\theta})^T(\mathbf{X}_n - \boldsymbol{\theta})) \Leftrightarrow \frac{R_2}{(\mathbf{X}_n - \boldsymbol{\theta})^T(\mathbf{X}_n - \boldsymbol{\theta})} \xrightarrow{P} 0$. Επομένως,

$$f(\mathbf{X}_n) - f(\boldsymbol{\theta}) = [\nabla f(\boldsymbol{\theta})]^T(\mathbf{X}_n - \boldsymbol{\theta}) + \tfrac{1}{2}[\mathbf{X}_n - \boldsymbol{\theta}]^T H(\boldsymbol{\theta})[\mathbf{X}_n - \boldsymbol{\theta}] + R_2$$
$$\overset{[\nabla f(\theta)]=0}{\Rightarrow} n(f(\mathbf{X}_n) - f(\boldsymbol{\theta})) = \tfrac{n}{2}[\mathbf{X}_n - \boldsymbol{\theta}]^T H(\boldsymbol{\theta})[\mathbf{X}_n - \boldsymbol{\theta}] + nR_2$$

Γνωρίζουμε ότι αν $\mathbf{X}_n \xrightarrow{D} \mathbf{X}$ και f συνεχής συνάρτηση $\Rightarrow f(\mathbf{X}_n) \xrightarrow{D} f(\mathbf{X})$.

Από υπόθεση,$\sqrt{n}(\mathbf{X}_n - \boldsymbol{\theta}) \xrightarrow{D} \mathbf{W}$

$\Rightarrow n(\mathbf{X}_n - \boldsymbol{\theta})^T H(\boldsymbol{\theta})(\mathbf{X}_n - \boldsymbol{\theta}) = \sqrt{n}(\mathbf{X}_n - \boldsymbol{\theta})^T H(\boldsymbol{\theta})\sqrt{n}(\mathbf{X}_n - \boldsymbol{\theta}) \xrightarrow{D} \mathbf{W}^T H(\boldsymbol{\theta})\mathbf{W}$.

Άρα, $\frac{n}{2}(\mathbf{X}_n - \boldsymbol{\theta})^T H(\boldsymbol{\theta})(\mathbf{X}_n - \boldsymbol{\theta}) \xrightarrow{D} \frac{1}{2}\mathbf{W}^T H(\boldsymbol{\theta})\mathbf{W}$.

Από Θεώρημα 1.6.1 $n(\mathbf{X}_n - \boldsymbol{\theta})^T H(\boldsymbol{\theta})(\mathbf{X}_n - \boldsymbol{\theta}) = O_P(1)$.
Επιπλέον, $nR_2 = no_P((\mathbf{X}_n - \boldsymbol{\theta})^T(\mathbf{X}_n - \boldsymbol{\theta})) = o_P(n(\mathbf{X}_n - \boldsymbol{\theta})^T(\mathbf{X}_n - \boldsymbol{\theta})) \Leftrightarrow \frac{nR_2}{n(\mathbf{X}_n - \boldsymbol{\theta})^T(\mathbf{X}_n - \boldsymbol{\theta})} \xrightarrow{P} 0$. Οπότε, $nR_2 \xrightarrow{P} 0$.

Καταληκτικά, από Θεώρημα Slutsky (βλ. Θεώρημα 1.5.1.) προκύπτει το ζητούμενο, δηλαδή $n(f(\mathbf{X}_n) - f(\boldsymbol{\theta})) \xrightarrow{D} \frac{1}{2}\mathbf{W}^T H_k(\boldsymbol{\theta})\mathbf{W}$. $\square$

3.3 Πολυδιάστατη Μέθοδος Δέλτα ανώτερης τάξης

Πολυδιάστατη μέθοδος Δέλτα ανώτερης τάξης από τον $\mathbb{R}^k$ στον $\mathbb{R}$

Θεώρημα 3.3.1. *Θεωρούμε ακολουθία $\boldsymbol{X}_n$ k-διάστατων τυχαίων διανυσμάτων τέτοια ώστε $\sqrt{n}(\boldsymbol{X}_n - \boldsymbol{\theta}) \xrightarrow{D} \boldsymbol{W}$, όπου $\boldsymbol{\theta} = \begin{bmatrix}\theta_1\\ \theta_2\\ \dots\\ \theta_k\end{bmatrix} \in \mathbb{R}^k$, $\boldsymbol{W} = \begin{bmatrix}W_1\\ W_2\\ \dots\\ W_k\end{bmatrix} \in \mathbb{R}^k$. Έστω $f : \mathbb{R}^k \to \mathbb{R}$ μια Borel-μετρήσιμη συνάρτηση με m συνεχείς μερικές παραγώγους γύρω από το $\boldsymbol{\theta}$. Ορίζουμε j_i το πλήθος των παραγωγίσεων ως προς τη μεταβλητή x_i, άρα οι μερικές παράγωγοι της f, δηλαδή για $\boldsymbol{x} = (x_1, x_2, \dots, x_k) \in D$ οι συναρτήσεις $f_{j_1,j_2,\dots,j_k}(\boldsymbol{x}) = f_{j_1,j_2,\dots,j_k}(x_1, x_2, \dots, x_k) = \frac{\partial^{j_1+j_2+\dots+j_k} f(\boldsymbol{x})}{\partial x_1^{j_1}, \partial x_2^{j_2}, \dots, \partial x_k^{j_k}}$ είναι συνεχείς, $\forall j_1, j_2, \dots, j_k \in \{0, 1, \dots\}$, όπου $j = j_1 + j_2 + \dots + j_k \leq m$. Τότε, υποθέτοντας ότι*

$f_{j_1,j_2,\dots,j_k}(\boldsymbol{\theta}) = 0, \forall j_1, j_2, \dots, j_k \in \{0, 1, \dots\}$, με $1 \leq j_1+j_2+\dots+j_k \leq m-1$, αληθεύει ότι

$n^{\frac{m}{2}}(f(\boldsymbol{X}_n) - f(\boldsymbol{\theta})) \xrightarrow{D} \frac{1}{m!}[D^m f(\boldsymbol{\theta})]\boldsymbol{W}^m$, όπου θέσαμε $[D^m f(\boldsymbol{\theta})]\boldsymbol{W}^m = \sum\limits_{j_1+j_2+\dots+j_k=m} \frac{m!}{j_1!\dots j_k!} f_{j_1,j_2,\dots,j_k}(\boldsymbol{\theta}) W_1^{j_1} W_2^{j_2} \dots W_k^{j_k}$. Ισοδύναμα, γράφουμε

$n^{\frac{m}{2}}(f(\boldsymbol{X}_n) - f(\boldsymbol{\theta})) \xrightarrow{D} \sum\limits_{j_1+j_2+\dots+j_k=m} f_{j_1,j_2,\dots,j_k}(\boldsymbol{\theta}) \frac{W_1^{j_1}}{j_1} \frac{W_2^{j_2}}{j_2} \dots \frac{W_k^{j_k}}{j_k}$.

Απόδειξη. Από υπόθεση,

$f_{j_1,j_2,\dots,j_k}(\boldsymbol{\theta}) = 0, \forall j = j_1 + \dots + j_k \leq m - 1, j \geq 1,$

$\Rightarrow [D^j f(\boldsymbol{\theta})](\mathbf{x}-\boldsymbol{\theta})^j = \sum\limits_{j_1+\dots+j_k=j} \frac{j!}{j_1!\dots j_k!} f_{j_1,\dots,j_k}(\boldsymbol{\theta})(x_1-\theta_1)^{j_1} \dots (x_k-\theta_k)^{j_k} = 0,$

$\forall 1 \leq j = j_1 + \dots + j_k \leq m - 1$. Άρα το Πολυδιάστατο θεώρημα Taylor (βλ. Θεώρημα 2.3.4.) θα είναι

$\lim\limits_{\mathbf{x}\to\boldsymbol{\theta}} \frac{f(\mathbf{x})-f(\boldsymbol{\theta})-\frac{D^1 f(\boldsymbol{\theta})}{1!}(\mathbf{x}-\boldsymbol{\theta})^1-\dots-\frac{D^m f(\boldsymbol{\theta})}{m!}(\mathbf{x}-\boldsymbol{\theta})^m}{\sum\limits_{i-1}^{k} |x_i-\theta_i|^m}$ ή ισοδύναμα

$$\lim_{\mathbf{x}\to\boldsymbol{\theta}} \frac{f(\mathbf{x}) - f(\boldsymbol{\theta}) - \frac{D^m f(\boldsymbol{\theta})}{m!}(\mathbf{x} - \boldsymbol{\theta})^m}{\sum\limits_{i-1}^{k} |x_i - \theta_i|^m} = 0. \tag{3.3.1}$$

Επιπλέον, έχουμε

$$n^{\frac{m}{2}}(f(\mathbf{X}_n)-f(\boldsymbol{\theta})) = n^{\frac{m}{2}}[f(\mathbf{X}_n)-f(\boldsymbol{\theta})-\tfrac{D^m f(\boldsymbol{\theta})}{m!}(\mathbf{X}_n-\boldsymbol{\theta})^m + \tfrac{D^m f(\boldsymbol{\theta})}{m!}(\mathbf{X}_n-\boldsymbol{\theta})^m]$$

$$= n^{\frac{m}{2}}[f(\mathbf{X}_n)-f(\boldsymbol{\theta})-\frac{D^m f(\boldsymbol{\theta})}{m!}(\mathbf{X}_n-\boldsymbol{\theta})^m]+n^{\frac{m}{2}}\frac{D^m f(\boldsymbol{\theta})}{m!}(\mathbf{X}_n-\boldsymbol{\theta})^m. \quad (3.3.2)$$

Ακόμα,

$$n^{\frac{m}{2}}[f(\mathbf{X}_n)-f(\boldsymbol{\theta})-\tfrac{D^m f(\boldsymbol{\theta})(\mathbf{X}_n-\boldsymbol{\theta})^m}{m!}] = \frac{n^{\frac{m}{2}}[f(\mathbf{X}_n)-f(\boldsymbol{\theta})-\frac{D^m f(\boldsymbol{\theta})}{m!}(\mathbf{X}_n-\boldsymbol{\theta})^m]\sum_{i-1}^{k}|X_{n,i}-\theta_i|^m}{\sum_{i-1}^{k}|X_{n,i}-\theta_i|^m}$$

$$= \frac{[f(\mathbf{X}_n)-f(\boldsymbol{\theta})-\frac{D^m f(\boldsymbol{\theta})}{m!}(\mathbf{X}_n-\boldsymbol{\theta})^m]}{\sum_{i-1}^{k}|X_{n,i}-\theta_i|^m}\sum_{i-1}^{k}(\sqrt{n}|X_{n,i}-\theta_i|)^m. \quad (3.3.3)$$

Όμως, $\mathbf{X}_n \xrightarrow{P} \boldsymbol{\theta} \overset{\Theta\varepsilon\omega\rho.1.2.1.}{\Longrightarrow} \mathbf{X}_n \xrightarrow{D} \boldsymbol{\theta} \overset{\Theta\varepsilon\omega\rho.1.5.2.}{\Longrightarrow} \exists \mathbf{Y}_n : \mathbf{Y}_n(\omega) \to \boldsymbol{\theta}, \forall \omega \in \Omega,$ με $\mathbf{Y}_n \overset{D}{=} \mathbf{X}_n$

$$\Longrightarrow P[\lim_{n\to\infty}|\mathbf{Y}_n-\theta|=0]=1 \Rightarrow \mathbf{Y}_n \xrightarrow{a.s} \boldsymbol{\theta}$$

$$\overset{(3.3.1)}{\Longrightarrow} \lim_{\mathbf{Y}_n\to\boldsymbol{\theta}} \frac{|f(\mathbf{Y}_n)-f(\boldsymbol{\theta})-\frac{D^m f(\boldsymbol{\theta})}{m!}(\mathbf{Y}_n-\boldsymbol{\theta})^m|}{\sum_{i-1}^{k}|Y_{n,i}-\theta_i|^m}=0$$

$$\Longrightarrow \frac{|f(\mathbf{Y}_n)-f(\boldsymbol{\theta})-\frac{D^m f(\boldsymbol{\theta})}{m!}(\mathbf{Y}_n-\boldsymbol{\theta})^m|}{\sum_{i-1}^{k}|Y_{n,i}-\theta_i|^m} \xrightarrow{a.s} 0$$

$$\overset{\Theta\varepsilon\omega\rho.1.2.2.}{\Longrightarrow} \frac{|f(\mathbf{Y}_n)-f(\boldsymbol{\theta})-\frac{D^m f(\boldsymbol{\theta})}{m!}(\mathbf{Y}_n-\boldsymbol{\theta})^m|}{\sum_{i-1}^{k}|Y_{n,i}-\theta_i|^m} \xrightarrow{P} 0. \quad (3.3.4)$$

$$\overset{\mathbf{Y}_n\overset{D}{=}\mathbf{X}_n}{\Longrightarrow} \frac{|f(\mathbf{X}_n)-f(\boldsymbol{\theta})-\frac{D^m f(\boldsymbol{\theta})}{m!}(\mathbf{X}_n-\boldsymbol{\theta})^m|}{\sum_{i-1}^{k}|X_{n,i}-\theta_i|^m} \xrightarrow{P} 0.$$

Και

$$\sqrt{n}(\mathbf{X}_n-\boldsymbol{\theta}) \xrightarrow{D} \mathbf{W} \Longrightarrow \sum_{i-1}^{k}(\sqrt{n}|X_{n,i}-\theta_i|)^m \xrightarrow{D} \sum_{i=1}^{k}|W_i|^m. \quad (3.3.5)$$

Επομένως, η (3.3.3),μέσω των (3.3.4) και (3.3.5) και του Θεωρήματος Slutsky (Θεώρημα 1.5.1.) γίνεται

$$n^{\frac{m}{2}}[f(\mathbf{X}_n)-f(\boldsymbol{\theta})-\frac{D^m f(\boldsymbol{\theta})(\mathbf{X}_n-\boldsymbol{\theta})^m}{m!}]\xrightarrow{D}0. \tag{3.3.6}$$

Επιπροσθέτως,
$n^{\frac{m}{2}}\frac{D^m f(\boldsymbol{\theta})}{m!}(\mathbf{X}_n-\boldsymbol{\theta})^m=\frac{D^m f(\boldsymbol{\theta})}{m!}[\sqrt{n}(\mathbf{X}_n-\boldsymbol{\theta})]^m$

$$=\frac{1}{m!}\sum_{j_1+\ldots+j_k=m}\frac{m!}{j_1!\ldots j_k!}f_{j_1,\ldots,j_k}(\boldsymbol{\theta})[\sqrt{n}(X_{n,1}-\theta_1)]^{j_1}\ldots[\sqrt{n}(X_{n,k}-\theta_k)]^{j_k}$$

$$\xrightarrow{D}\sum_{j_1+\ldots+j_k=m}f_{j_1,\ldots,j_k}(\boldsymbol{\theta})\frac{W_1^{j_1}}{j_1}\ldots\frac{W_k^{j_k}}{j_k}, \tag{3.3.7}$$

αφού $\sqrt{n}(\mathbf{X}_n-\boldsymbol{\theta})\xrightarrow{D}\mathbf{W}$.
Τελικώς, από Θεώρημα Slutsky (Θεώρημα 1.5.1.) και εφαρμόζοντας τις (3.3.6), (3.3.7) στην (3.3.2) καταλήγουμε ότι
$n^{\frac{m}{2}}(f(\mathbf{X}_n)-f(\boldsymbol{\theta}))\xrightarrow{D}\sum_{j_1+\ldots+j_k=m}f_{j_1,\ldots,j_k}(\boldsymbol{\theta})\frac{W_1^{j_1}}{j_1!}\ldots\frac{W_k^{j_k}}{j_k!}.$ $\square$

Κεφάλαιο 4

Οριακή συμπεριφορά δειγματικών κεντρικών ροπών

4.1 Συνέπεια δειγματικών κεντρικών ροπών

Έστω τυχαία μεταβλητή X, μη εκφυλισμένη, με συνάρτηση κατανομής F για την οποία ισχύει $E[X^{2k}] < \infty$, για κάποια σταθερά $k \in \{1, 2, ...\}$. Επιπλέον, θεωρούμε τυχαίο δείγμα $X_1, X_2, ..., X_n$ που ακολουθεί την ίδια κατανομή F.

Ορισμός 4.1.1. Μια τυχαία μεταβλητή X ονομάζεται *μη εκφυλισμένη* όταν δεν είναι σταθερή με πιθανότητα τη μονάδα.
Ισοδύναμα, λέμε ότι η τυχαία μεταβλητή X είναι μη εκφυλισμένη όταν το σύνολο των σημείων αύξησης της συνάρτησης κατανομής της,

$\{x \in \mathbb{R} : F(x+\varepsilon) - F(x-\varepsilon) > 0, \forall \varepsilon > 0\} \neq \emptyset$ δεν είναι μονοσύνολο.

Ορισμός 4.1.2. Έστω τυχαία μεταβλητή X. Λέγεται ότι υπάρχει η *κεντρική ροπή n-τάξης περί την αρχή* και συμβολίζουμε με $\mu'_n = E[X^n]$, αν και μόνο αν $E[|X|^n] < \infty$.

Παρατήρηση 1.4.1. Αν $E[|X|^n] < \infty \implies E[|X|^m] < \infty, \forall m < n$,όπου $m, n \in \mathbb{R}^+$.

Απόδειξη.
$$\begin{aligned} m < n \implies |X|^m < |X|^n &\implies |X|^m < |X|^n + 1 \\ &\implies E[|X|^m] < E[|X|^n + 1] \\ &\implies E[|X|^m] < E[|X|^n] + E[1] \\ &\implies E[|X|^m] < E[|X|^n] + 1 < \infty. \end{aligned}$$ □

Ορισμός 4.1.3. Έστω ότι η τυχαία μεταβλητή X έχει πεπερασμένη μέση τιμή

$\mu = E[X]$. Τότε, λέμε ότι υπάρχει η κεντρική ροπή j-τάξης και συμβολίζουμε με $\mu_j = E[(X-\mu)^j]$, αν και μόνο αν $E[|X-\mu|^j] < \infty, \quad j \in \mathbb{N}$.

Βάσει των ανωτέρω, συμπεραίνουμε ότι οι πρώτες $2k-$τάξης κεντρικές ροπές της τυχαίας μεταβλητής X γύρω από τον μέσο $\mu := E[X]$ είναι καλά ορισμένες, πεπερασμένες και υπολογίζονται από τον τύπο

$$\mu_j := E[(X-\mu)^j], \quad j = 0, 1, \dots, 2k. \tag{4.1.1}$$

Πιο συγκεκριμένα,έχουμε:

- για $j = 1$, $\quad \mu_1 = E[(X-\mu)^1] = E[X-\mu] = E[X] - \mu = 0$,

- για $j = 2$ $$\begin{aligned} \mu_2 &= E[(X-\mu)^2] = E[X^2 - 2X\mu + \mu^2] \\ &= E[X^2] - 2\mu E[X] + \mu^2 = E[X^2] - 2E[X]^2 + E[X]^2 \\ &= E[X^2] - E[X]^2 = Var[X] = \sigma^2 > 0, \end{aligned}$$

- για $j = 3$, $$\begin{aligned} \mu_3 &= E[(X-\mu)^3] = E[X^3 - 3X^2\mu + 3X\mu^2 - \mu^3] \\ &= E[X^3] - 3E[X^2]\mu + 3E[X]\mu^2 - \mu^3 \\ &= E[X^3] - 3E[X^2]E[X] + 3E[X]E[X]^2 - E[X]^3 \\ &= E[X^3] - 3E[X]\{E[X^2] - E[X]^2\} - E[X]^3 \\ &= E[X^3] - 3E[X]Var[X] - E[X]^3, \end{aligned}$$

(δεδομένου ότι η $E[X^3]$ υπάρχει) κ.ο.κ.

Σχόλιο 1.4.1. Η $\sigma^2 = Var[X] = E[(X-\mu)^2]$ είναι αυστηρά θετική διότι ισχύει $E[(X-\mu)^2] \geq 0$ και για να είναι $E[(X-\mu)^2] = 0 \Rightarrow X = \mu$ καταλήγουμε σε άτοπο διότι η X από υπόθεση είναι μη εκφυλισμένη, δηλαδή δεν παίρνει μοναδική τιμή με πιθανότητα τη μονάδα. Άρα, $\sigma^2 > 0$.

Η φυσιολογική εκτιμήτρια της $\mu_k = E[(X-\mu)^k]$, για $k \geq 2$, όταν η μέση τιμή $\mu = E[X]$ είναι άγνωστη (όπως συμβαίνει στις περισσότερες περιπτώσεις), είναι η δειγματική κεντρική ροπή k-τάξης, η οποία ορίζεται ως:

$$M_k = M_{k,n} := \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^k, \tag{4.1.2}$$

όπου $\overline{X}$, είναι ο δειγματικός μέσος, δηλαδή $\overline{X} = \overline{X}_n := \frac{1}{n}\sum\limits_{i=1}^{n} X_i$.

Παρακάτω, θα ελέγξουμε τη συνέπεια της εκτιμήτριας M_k της μ_k.

Ορισμός 4.1.4. Ορίζουμε ως *δειγματικές κεντρικές ροπές γύρω από τον (άγνωστο) μέσο* μ τις ποσότητες:

$$m_j = m_{j,n} := \frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^j, \qquad j=1,2,\ldots,k. \tag{4.1.3}$$

Παρακάτω, θα δείξουμε μέσω του πολυωνύμου Newton ότι οι δειγματικές κεντρικές ροπές M_k μπορούν να εκφραστούν μέσω των δειγματικών κεντρικών ροπών περί τη μέση τιμή, δηλαδή των m_j.

Για κάθε ζεύγος πραγματικών αριθμών a, b
(ή εναλλακτικά $\forall a, b$ τ.ω $ab = ba$) και $\forall n \in \mathbb{N}$ ισχύει ο τύπος του **Διωνύμου του Newton:**

$$(a+b)^n = \sum_{k=0}^{n}\binom{n}{k}a^k b^{n-k} = b^n + nab^{n-1} + \ldots + \binom{n}{k}a^k b^{n-k} + \ldots + na^{n-1}b + a^n. \tag{4.1.4}$$

Έχουμε, $m_1 = m_{1,n} = \frac{1}{n}\sum\limits_{i=1}^{n}(X_i-\mu) = \frac{1}{n}\sum\limits_{i=1}^{n}X_i - \frac{1}{n}n\mu = \overline{X}-\mu$

$\Rightarrow \overline{X} = \mu + m_1$.

Επομένως, η δειγματική κεντρική ροπή k-τάξης γράφεται:

$$\begin{aligned}
M_k &= M_{k,n} = \tfrac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^k = \tfrac{1}{n}\sum_{i=1}^{n}(X_i-\mu-m_1)^k \\
&\overset{4.1.4}{=} \tfrac{1}{n}\sum_{i=1}^{n}\Big[\sum_{j=0}^{k}\tbinom{k}{j}(X_i-\mu)^j(-m_1)^{k-j}\Big] = \tfrac{1}{n}\sum_{i=1}^{n}\Big[\sum_{j=0}^{k}\tbinom{k}{j}(X_i-\mu)^j(-1)^{k-j}m_1^{k-j}\Big] \\
&= \tfrac{1}{n}\sum_{i=1}^{n}\Big[\tbinom{k}{0}(X_i-\mu)^0(-1)^{k-0}m_1^{k-0} + \tbinom{k}{1}(X_i-\mu)^1(-1)^{k-1}m_1^{k-1} \\
&\qquad + \sum_{j=2}^{k-1}\tbinom{k}{j}(X_i-\mu)^j(-1)^{k-j}m_1^{k-j} + \tbinom{k}{k}(X_i-\mu)^k(-1)^{k-k}m_1^{k-k}\Big] \\
&= \tfrac{1}{n}\sum_{i=1}^{n}\Big[(-1)^k m_1^k + k(X_i-\mu)(-1)^{k-1}m_1^{k-1} \\
&\qquad + \sum_{j=2}^{k-1}\tbinom{k}{j}(X_i-\mu)^j(-1)^{k-j}m_1^{k-j} + (X_i-\mu)^k\Big] \\
&= \tfrac{1}{n}n[(-1)^k m_1^k] + k(-1)^{k-1}m_1^{k-1}\sum_{i=1}^{n}(X_i-\mu)] \\
&\qquad + \sum_{j=2}^{k-1}\tbinom{k}{j}(-1)^{k-j}m_1^{k-j}\Big[\sum_{i=1}^{n}(X_i-\mu)^j\Big] + \tfrac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^k \\
&\overset{4.1.3.}{=} (-1)^k m_1^k + k(-1)^{k-1}m_1^{k-1}m_1 + \sum_{j=2}^{k-1}\tbinom{k}{j}(-1)^{k-j}m_1^{k-j}m_j + m_k
\end{aligned}$$

$$= (-1)^k m_1^k + k(-1)^{k-1} m_1^k + \sum_{j=2}^{k-1} \tbinom{k}{j} (-1)^{k-j} m_1^{k-j} m_j + m_k$$
$$= (-1)^{k-1}(k-1) m_1^k + \sum_{j=2}^{k-1} \tbinom{k}{j} (-1)^{k-j} m_1^{k-j} m_j + m_k.$$

Συνεπώς, καταλήγουμε ότι:

$$M_k = M_{k,n} = (-1)^{k-1}(k-1)m_1^k + \sum_{j=2}^{k-1} \binom{k}{j} (-1)^{k-j} m_j m_1^{k-j} + m_k. \quad (4.1.5)$$

Θεωρούμε $\{Y_i = (X_i - \mu)^k : i = 1, 2, \dots, n\}$ ακολουθία ανεξάρτητων και ισόνομων τυχαίων μεταβλητών (αφού $X_i : i = 1, \dots n$ ανεξάρτητες και ισόνομες τυχαίες μεταβλητές). Δεδομένου ότι $E[|X_i|^k] < \infty \Rightarrow E[|Y_i|^k] < \infty$ άρα για $j \leq k$ μπορούμε να εφαρμόσουμε τον Ισχυρό Νόμο Μεγάλων Αριθμών (βλ. Θεώρημα 1.4.4.). Έστω $S_n = Y_1 + \dots + Y_n$.

Τότε, $\frac{S_n}{n} \xrightarrow{a.s} E[Y_i] \Longrightarrow \frac{1}{n} \sum_{i=1}^{n} (X_i - \mu)^j \xrightarrow{a.s} E[(X_i - \mu)^j]$, καθώς $n \to \infty$

$$\overset{(4.1.1.),(4.1.3.)}{\Longrightarrow} m_{j,n} \xrightarrow{a.s} \mu_j \text{ ,καθώς } n \to \infty. \qquad (4.1.5a)$$

για j=1: $m_{1,n} \xrightarrow{a.s} \mu_1 = 0$,καθώς $n \to \infty$. (4.1.5b)

Επομένως,(4.1.5) $\overset{(4.1.5a),(4.1.5b)}{\Longrightarrow}$

$M_k := M_{k,n} \xrightarrow{a.s} (-1)^{k-1}(k-1)0^k + \sum_{j=2}^{k-1} \tbinom{k}{j} (-1)^{k-j} \mu_j 0^{k-j} + \mu_k$ καθώς $n \to \infty$.

Δηλαδή, καταλήγουμε ότι $M_k = M_{k,n} \xrightarrow{a.s} \mu_k$, καθώς $n \to \infty$.

Παρατήρηση 4.1.2. Παρατηρούμε ότι οι δειγματικές κεντρικές ροπές γύρω από τον μέσο $m_j = m_{j,n} := \frac{1}{n} \sum_{i=1}^{n} (X_i - \mu)^j, \quad j = 1, 2, \dots, k$ είναι αθροίσματα ανεξάρτητων και ισόνομων τυχαίων μεταβλητών αφού $(X_i - \mu)^j$ ανεξάρτητες και ισόνομες τ.μ., $\forall i = 1, \dots n$, γεγονός το οποίο καθιστά εφαρμόσιμο το Κεντρικό Οριακό Θεώρημα για την προσέγγιση της κατανομής τους.

Αντίθετα, με την αλλαγή του μέσου μ από το δειγματικό μέσο $\overline{X}$, οι δειγματικές κεντρικές ροπές $M_k = M_{k,n} := \frac{1}{n} \sum_{i=1}^{n} (X_i - \overline{X})^k$ δεν είναι διαχειρίσιμες με τον ίδιο τρόπο αφού οι ποσότητες $\{(X_i - \overline{X})^k, i = 1, 2, \dots, n\}$ δεν είναι ανεξάρτητες. Παραδείγματος χάριν, η προσέγγιση της κατανομής της M_2, με βάση τη σχέση

$M_2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \overline{X}^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - [\frac{1}{n}\sum_{i=1}^{n} X_i]^2$, αντιμετωπίζεται πλέον ως η προσέγγιση της κατανομής του διδιάστατου διανύσματος $\left[\frac{1}{n}\sum_{i=1}^{n} X_i^2 \quad [\frac{1}{n}\sum_{i=1}^{n} X_i]^2\right]'$.
Προκειμένου να επιτευχθεί η μελέτη της συμπεριφοράς τους, οι M_k γράφονται ως συνάρτηση του διανύσματος $\mathbf{m}_n = (m_{1,n}, m_{2,n}, \ldots, m_{k,n})$ (βλ. 4.1.5.).Δηλαδή, όπως αποδείχθηκε παραπάνω,

$$M_k = M_{k,n} = g(\mathbf{m}_n) = (-1)^{k-1}(k-1)m_1^k + \sum_{j=2}^{k-1}\binom{k}{j}(-1)^{k-j} m_j m_1^{k-j} + m_k. \tag{4.1.6}$$

4.2 Οριακή κατανομή δειγματικών κεντρικών ροπών

Οριακή κατανομή της δειγματικής κεντρικής ροπής $M_{k,n}$

Έστω $k \geq 2$. Θεωρούμε το διάνυσμα όλων των δειγματικών κεντρικών ροπών περί την αρχή $\mathbf{m} = \mathbf{m}_n := (m_{1,n}, m_{2,n}, \ldots, m_{k,n})^T$, όπου $\forall j = 1, 2, \ldots, k$ ισχύει ότι $m_{j,n} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^j$.

Εφαρμόζοντας το Πολυδιάστατο Κεντρικό Οριακό Θεώρημα (βλ. Θεώρημα 1.4.6.) στα k- διάστατα ανεξάρτητα και ισόνομα τυχαία διανύσματα
$\mathbf{Y}_1 = \begin{bmatrix} X_1 - \mu \\ (X_1-\mu)^2 \\ \ldots \\ (X_1-\mu)^k \end{bmatrix}, \mathbf{Y}_2 = \begin{bmatrix} X_2 - \mu \\ (X_2-\mu)^2 \\ \ldots \\ (X_2-\mu)^k \end{bmatrix}, \ldots, \mathbf{Y}_n = \begin{bmatrix} X_n - \mu \\ (X_n-\mu)^2 \\ \ldots \\ (X_n-\mu)^k \end{bmatrix}$,
καταλήγουμε ότι καθώς $n \to \infty$ αληθεύει ότι

$$\sqrt{n}(\mathbf{m}_n - \boldsymbol{\mu}) \xrightarrow{D} N(\mathbf{0}, \boldsymbol{\Sigma}), \tag{4.2.1}$$

όπου $\mathbf{0} = (0, 0, \ldots, 0)' \in \mathbb{R}^k$, $\boldsymbol{\mu} = (\mu_1, \mu_2, \ldots, \mu_k) = (0, \sigma^2, \ldots, \mu_k)' \in \mathbb{R}^k$ και $\boldsymbol{\Sigma} = (\sigma_{ij}) \in \mathbb{R}^{kxk}$ είναι kxk πίνακας διασποράς του $\mathbf{Y}_1$ με στοιχεία $\sigma_{ij} = Cov[(X_1-\mu)^i, (X_1-\mu)^j] = \mu_{i+j} - \mu_i\mu_j$, δηλαδή
$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma^2 & \mu_3 & \cdots & \mu_{k+1} \\ \mu_3 & \mu_4 - \sigma^4 & \cdots & \mu_{k+2} - \sigma^2\mu_k \\ \ldots & \ldots & \ldots & \ldots \\ \mu_{k+1} & \mu_{k+2} - \sigma^2\mu_k & \cdots & \mu_{2k} - \mu_k^2 \end{bmatrix}.$$
Πράγματι, $\sqrt{n}(\mathbf{m}_n - \boldsymbol{\mu}) = \sqrt{n}(m_{1,n} - \mu_1, m_{2,n} - \mu_2, \ldots, m_{k,n} - \mu_k)$
$= (\sqrt{n}(\frac{1}{n}\sum_{i=1}^{n}(X_i - \mu) - \mu_1), \ldots, \sqrt{n}(\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^k - \mu_k))$

$$\xrightarrow{D} N(\mathbf{0}, \mathbf{\Sigma}).$$

Από τη σχέση (4.1.6) προέκυψε ότι

$M_k = M_{k,n} = g(\mathbf{m}_n) = (-1)^{k-1}(k-1)m_1^k + \sum_{j=2}^{k-1} \binom{k}{j}(-1)^{k-j} m_j m_1^{k-j} + m_k,$

και έτσι ορίζουμε τη συνάρτηση $g : \mathbb{R}^k \to \mathbb{R}$, η οποία δίνεται από τον τύπο

$g(\mathbf{x}) = g(x_1, \dots, x_k) = (-1)^{k-1}(k-1)x_1^k + \sum_{j=2}^{k-1} \binom{k}{j}(-1)^{k-j} x_j x_1^{k-j} + x_k,$

για $\mathbf{x} = (x_1, \dots, x_k)^T \in \mathbb{R}^k$.

Η συνάρτηση g έχει συνεχείς μερικές παραγώγους, αφού είναι πολυώνυμο, άρα μπορούμε να εφαρμόσουμε Πολυδιάστατη μέθοδο Δέλτα (βλ. Θεώρημα 3.1.2.).

Αφού $\sqrt{n}(\mathbf{m}_n - \boldsymbol{\mu}) \xrightarrow{D} \mathbf{W} \sim N(\mathbf{0}, \mathbf{\Sigma})$ και g έχει συνεχείς παραγώγους, άρα ισχύει

$$\sqrt{n}(g(\mathbf{m}_n) - g(\boldsymbol{\mu})) \xrightarrow{D} [D^1 g(\boldsymbol{\mu})]\mathbf{W} = [\nabla g(\boldsymbol{\mu})]\mathbf{W} \sim N(0, [\nabla g(\boldsymbol{\mu})]^T \mathbf{\Sigma} [\nabla g(\boldsymbol{\mu})]).$$

Παρακάτω, θα υπολογίσουμε την ποσότητα $[\nabla g(\boldsymbol{\mu})]^T \mathbf{\Sigma} [\nabla g(\boldsymbol{\mu})]$.

Έχουμε $\nabla g(\mathbf{x}) = \begin{bmatrix} \frac{\partial g(x_1,\dots,x_k)}{\partial x_1} \\ \frac{\partial g(x_1,\dots,x_k)}{\partial x_2} \\ \dots \\ \frac{\partial g(x_1,\dots,x_k)}{\partial x_k} \end{bmatrix}$ και

$\frac{\partial g(x_1,\dots,x_k)}{\partial x_1} = \frac{\partial}{\partial x_1}[(-1)^{k-1}(k-1)x_1^k + \sum_{j=2}^{k-1} \binom{k}{j}(-1)^{k-j} x_j x_1^{k-j} + x_k]$

$\qquad = x_1 \sum_{j=1}^{k-2} \binom{k}{j}(k-j) x_j x_1^{k-j-2} - k x_{k-1}.$

$\frac{\partial g(x_1,\dots,x_k)}{\partial x_i} = \frac{\partial}{\partial x_i}[(-1)^{k-1}(k-1)x_1^k + \sum_{j=2}^{k-1} \binom{k}{j}(-1)^{k-j} x_j x_1^{k-j} + x_k]$

$\qquad = (-1)^{k-i}\binom{k}{i} x_1^{k-i}, \quad i \in \{2, 3, \dots, k-1\}.$

$\frac{\partial g(x_1,\dots,x_k)}{\partial x_k} = \frac{\partial}{\partial x_k}[(-1)^{k-1}(k-1)x_1^k + \sum_{j=2}^{k-1} \binom{k}{j}(-1)^{k-j} x_j x_1^{k-j} + x_k]$

$\qquad = x_k.$

Και επειδή $\boldsymbol{\mu} = (\mu_1, \mu_2, \dots, \mu_k) = (0, \sigma^2, \dots, \mu_k)$ προκύπτει

$\frac{\partial g(x_1,\dots,x_k)}{\partial x_1}|_{\mathbf{x}=\boldsymbol{\mu}} = \mu_1 \sum_{j=1}^{k-2} \binom{k}{j}(k-j)\mu_j \mu_1^{k-j-2} - k\mu_{k-1} \overset{\mu_1=0}{=} -k\mu_{k-1}.$

$\frac{\partial g(x_1,\dots,x_k)}{\partial x_i}|_{\mathbf{x}=\boldsymbol{\mu}} = (-1)^{k-i}\binom{k}{i}\mu_1^{k-i} \overset{\mu_1=0}{=} 0, \quad i \in \{2, 3, \dots, k-1\}.$

$\frac{\partial g(x_1,\dots,x_k)}{\partial x_k}|_{\mathbf{x}=\boldsymbol{\mu}} = \mu_k.$

Επομένως, $\nabla g(\boldsymbol{\mu}) = \begin{bmatrix} \frac{\partial g(x_1,\ldots,x_k)}{\partial x_1}|_{\mathbf{x}=\boldsymbol{\mu}} \\ \frac{\partial g(x_1,\ldots,x_k)}{\partial x_2}|_{\mathbf{x}=\boldsymbol{\mu}} \\ \ldots \\ \frac{\partial g(x_1,\ldots,x_k)}{\partial x_k}|_{\mathbf{x}=\boldsymbol{\mu}} \end{bmatrix} = \begin{bmatrix} -k\mu_{k-1} \\ 0 \\ \ldots \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} g_1 \\ 0 \\ \ldots \\ 0 \\ g_k \end{bmatrix}$.

Άρα, $[\nabla g(\boldsymbol{\mu})]^T \boldsymbol{\Sigma} [\nabla g(\boldsymbol{\mu})] = [g_1 \quad 0 \quad \ldots \quad 0 \quad g_k] \begin{bmatrix} \sigma_{11} & \sigma_{12} & \ldots & \sigma_{1k} \\ \sigma_{12} & \sigma_{22} & \ldots & \sigma_{2k} \\ \ldots & \ldots & \ldots & \ldots \\ \sigma_{1k} & \sigma_{2k} & \ldots & \sigma_{kk} \end{bmatrix} \begin{bmatrix} g_1 \\ 0 \\ \ldots \\ 0 \\ g_k \end{bmatrix}$

$$= [g_1\sigma_{11} + g_k\sigma_{1k} \quad g_2\sigma_{12} + g_k\sigma_{2k} \quad \ldots \quad g_1\sigma_{1k} + g_k\sigma_{kk}] \begin{bmatrix} g_1 \\ 0 \\ \ldots \\ 0 \\ g_k \end{bmatrix}$$

$$\begin{aligned} &= (g_1\sigma_{11} + g_k\sigma_{1k})g_1 + (g_1\sigma_{1k} + g_k\sigma_{kk})g_k \\ &= g_1^2\sigma_{11} + 2g_1g_k\sigma_{1k} + g_k^2\sigma_{kk}. \end{aligned}$$

Επιπλέον, $\sigma_{ij} = \mu_{i+j} - \mu_i\mu_j \Longrightarrow \sigma_{11} = \mu_2 - \mu_1^2 \overset{\mu_1=0}{=} \mu_2 = \sigma^2$.

$$\sigma_{1k} = \mu_{k+1} - \mu_1\mu_k \overset{\mu_1=0}{=} \mu_{k+1}.$$

$$\sigma_{kk} = \mu_{2k} - \mu_k^2.$$

Οπότε, $[\nabla g(\boldsymbol{\mu})]^T \boldsymbol{\Sigma} [\nabla g(\boldsymbol{\mu})] = g_1^2\sigma_{11} + 2g_1g_k\sigma_{1k} + g_k^2\sigma_{kk}$

$$\begin{aligned} &= (-k\mu_{k-1})^2\sigma^2 + 2(-k\mu_{k-1})1\mu_{k+1} + 1^2(\mu_{2k} - \mu_k^2) \\ &= \mu_{2k} - \mu_k^2 + k\mu_{k-1}[k\sigma^2\mu_{k-1} - 2\mu_{k+1}] = v_k^2. \end{aligned}$$

Βρήκαμε οριακή διασπορά

$$v_k^2 = \mu_{2k} - \mu_k^2 + k\mu_{k-1}[k\sigma^2\mu_{k-1} - 2\mu_{k+1}]. \tag{4.2.2}$$

Για παράδειγμα, οι τρεις πρώτες τιμές της οριακής διασποράς είναι

$v_2^2 = \mu_4 - \mu_2^2 + 2\mu_1[2\sigma^2\mu_1 - 2\mu_3] \overset{\mu_1=0}{=} \mu_4 - \sigma^4$, όταν $E[X^4] < \infty$.

$v_3^2 = \mu_6 - \mu_3^2 + 3\mu_2[3\sigma^2\mu_2 - 2\mu_4] \overset{\mu_2=\sigma^2}{=} \mu_6 - \mu_3^2 + 9\sigma^6 - 6\sigma^2\mu_4$, όταν $E[X^6] < \infty$.

$v_4^2 = \mu_8 - \mu_4^2 + 4\mu_3[4\sigma^2\mu_3 - 2\mu_5] = \mu_8 - \mu_4^2 + 16\sigma^2\mu_3^2 - 8\mu_3\mu_5$, όταν $E[X^8] < \infty$.

Επομένως, με βάση τα παραπάνω καταλήξαμε ότι

$\sqrt{n}(g(\mathbf{m}_n) - g(\boldsymbol{\mu})) \xrightarrow{D} N(0, [\nabla g(\boldsymbol{\mu})]^T \boldsymbol{\Sigma} [\nabla g(\boldsymbol{\mu})])$ ή ισοδύναμα

$$\sqrt{n}(M_{k,n} - \mu_k) \xrightarrow{D} N(0, v_k^2), \tag{4.2.3}$$

δεδομένων των σχέσεων (4.1.6),(4.2.2) και επειδή με αντικατάσταση προκύπτει

$$g(\boldsymbol{\mu}) = g(\mu_1, \dots, \mu_k) = (-1)^{k-1}(k-1)\mu_1^k + \sum_{j=2}^{k-1} \binom{k}{j}(-1)^{k-j}\mu_j\mu_1^{k-j} + \mu_k \overset{\mu_1=0}{=} \mu_k. \tag{4.2.4}$$

Καταληκτικά, σύμφωνα με τον Ορισμό 1.5.4. παρατηρούμε ότι η δειγματική κεντρική ροπή $M_{k,n}$ είναι ασυμπτωτικά κανονική εκτιμήτρια της κεντρικής ροπής μ_k.

Οριακή κατανομή του διανύσματος δειγματικών κεντρικών ροπών $(M_{r,n}M_{k,n})$

Θεωρούμε το ζεύγος δειγματικών κεντρικών ροπών $(M_{r,n}M_{k,n})$, με $2 \le r < k$.
Από σχέση (4.1.6) ορίζουμε συνάρτηση $g : \mathbb{R}^k \to \mathbb{R}$ τέτοια ώστε
$g(\mathbf{x}) = g(x_1, \dots, x_k) = (-1)^{k-1}(k-1)x_1^k + \sum\limits_{j=2}^{k-1} \binom{k}{j}(-1)^{k-j}x_j x_1^{k-j} + x_k,$
για $\mathbf{x} = (x_1, \dots, x_k)' \in \mathbb{R}^k$. Οπότε ισχύει $M_{k,n} = g(\mathbf{m}_n)$.
Αντίστοιχα, ορίζουμε συνάρτηση $h : \mathbb{R}^k \to \mathbb{R}$ τέτοια ώστε
$h(\mathbf{x}) = h(x_1, \dots, x_k) = (-1)^{r-1}(r-1)x_1^r + \sum\limits_{j=2}^{r-1} \binom{r}{j}(-1)^{r-j}x_j x_1^{r-j} + x_r,$
για $\mathbf{x} = (x_1, \dots, x_k)^T \in \mathbb{R}^k$. Οπότε ισχύει $M_{r,n} = h(\mathbf{m}_n)$.
Αξίζει να σημειώσουμε ότι η συνάρτηση h εξαρτάται μόνο από τις πρώτες r μεταβλητές $x_1, \dots, x_r$ του $\mathbf{x} = (x_1, \dots, x_k)^T$ Από τη μονοδιάστατη περίπτωση βρήκαμε ότι

$$\nabla g(\boldsymbol{\mu}) = \begin{bmatrix} \frac{\partial g(x_1,\dots,x_k)}{\partial x_1}|_{\mathbf{x}=\boldsymbol{\mu}} \\ \frac{\partial g(x_1,\dots,x_k)}{\partial x_2}|_{\mathbf{x}=\boldsymbol{\mu}} \\ \dots \\ \frac{\partial g(x_1,\dots,x_k)}{\partial x_k}|_{\mathbf{x}=\boldsymbol{\mu}} \end{bmatrix} = \begin{bmatrix} -k\mu_{k-1} \\ 0 \\ \dots \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} g_1 \\ 0 \\ \dots \\ 0 \\ g_k \end{bmatrix}. \tag{4.2.5}$$

Για τη συνάρτηση h, με παραγώγιση προκύπτει:

$$\begin{aligned}
\frac{\partial h(x_1,\dots,x_k)}{\partial x_1} &= \frac{\partial}{\partial x_1}[(-1)^{r-1}(r-1)x_1^r + \sum_{j=2}^{r-1} \tbinom{r}{j}(-1)^{r-j}x_j x_1^{r-j} + x_r] \\
&= x_1 \sum_{j=1}^{r-2} \tbinom{r}{j}(r-j)x_j x_1^{r-j-2} - r x_{r-1}. \\
\frac{\partial h(x_1,\dots,x_k)}{\partial x_i} &= \frac{\partial}{\partial x_i}[(-1)^{r-1}(r-1)x_1^r + \sum_{j=2}^{r-1} \tbinom{r}{j}(-1)^{r-j}x_j x_1^{r-j} + x_r] \\
&= (-1)^{k-i}\tbinom{r}{i}x_1^{r-i}, \quad i \in \{2, 3, \dots, r-1\}. \\
\frac{\partial h(x_1,\dots,x_r)}{\partial x_r} &= \frac{\partial}{\partial x_r}[(-1)^{r-1}(r-1)x_1^r + \sum_{j=2}^{r-1} \tbinom{r}{j}(-1)^{r-j}x_j x_1^{r-j} + x_r]
\end{aligned}$$

$= x_r.$

$\frac{\partial h(x_1,\dots,x_k)}{\partial x_i} = \frac{\partial}{\partial x_i}[(-1)^{r-1}(r-1)x_1^r + \sum_{j=2}^{r-1}\binom{r}{j}(-1)^{r-j}x_j x_1^{r-j} + x_r]$

$= 0, \quad i \in \{r+1,\dots,k\}.$

Και επειδή $\boldsymbol{\mu} = (\mu_1,\mu_2,\dots,\mu_k) = (0,\sigma^2,\dots,\mu_k)$ προκύπτει

$\frac{\partial h(x_1,\dots,x_k)}{\partial x_1}|_{\mathbf{x}=\boldsymbol{\mu}} = \mu_1 \sum_{j=1}^{r-2}\binom{r}{j}(r-j)\mu_j\mu_1^{r-j-2} - r\mu_{r-1} \overset{\mu_1=0}{=} -r\mu_{r-1}.$

$\frac{\partial h(x_1,\dots,x_k)}{\partial x_i}|_{\mathbf{x}=\boldsymbol{\mu}} = (-1)^{r-i}\binom{r}{i}\mu_1^{r-i} \overset{\mu_1=0}{=} 0, \quad i \in \{2,3,\dots,k-1\}.$

$\frac{\partial h(x_1,\dots,x_k)}{\partial x_r}|_{\mathbf{x}=\boldsymbol{\mu}} = \mu_r.$

$\frac{\partial h(x_1,\dots,x_k)}{\partial x_i}|_{\mathbf{x}=\boldsymbol{\mu}} = 0, \quad i \in \{r+1,\dots,k\}.$

Οπότε,

$$\nabla h(\boldsymbol{\mu}) = \begin{bmatrix} \frac{\partial h(x_1,\dots,x_k)}{\partial x_1}|_{\mathbf{x}=\boldsymbol{\mu}} \\ \frac{\partial h(x_1,\dots,x_k)}{\partial x_2}|_{\mathbf{x}=\boldsymbol{\mu}} \\ \dots \\ \frac{\partial h(x_1,\dots,x_r)}{\partial x_r}|_{\mathbf{x}=\boldsymbol{\mu}} \\ \dots \\ \frac{\partial h(x_1,\dots,x_k)}{\partial x_k}|_{\mathbf{x}=\boldsymbol{\mu}} \end{bmatrix} = \begin{bmatrix} -r\mu_{r-1} \\ 0 \\ \dots \\ 0 \\ 1 \\ 0 \\ \dots \\ 0 \end{bmatrix} = \begin{bmatrix} h_1 \\ 0 \\ \dots \\ 0 \\ h_r \\ 0 \\ \dots \\ 0 \end{bmatrix}. \tag{4.2.6}$$

Όπως στη μονοδιάστατη περίπτωση, για $k \geq 2$ και $\mathbf{m} = \mathbf{m}_n := (m_{1,n}, m_{2,n},\dots,m_{k,n})'$ το πολυδιάστατο Κεντρικό Οριακό Θεώρημα (βλ. Θεώρημα 1.4.6.) συνεπάγεται καθώς $n \to \infty$, $\sqrt{n}(\mathbf{m}_n - \boldsymbol{\mu}) \xrightarrow{D} \mathbf{W} \sim N(\mathbf{0},\boldsymbol{\Sigma})$ και οι συναρτήσεις g,h έχουν συνεχείς παραγώγους, άρα μπορούμε να εφαρμόσουμε τη μέθοδο Δέλτα.

Από Πολυδιάστατη μέθοδο Δέλτα πρώτης τάξης από $\mathbb{R}^k$ στον $\mathbb{R}^2$ (βλ. Θεώρημα 3.1.2.) προκύπτει ότι

$$\sqrt{n}\begin{bmatrix} h(\mathbf{m}_n) - h(\boldsymbol{\mu}) \\ g(\mathbf{m}_n) - g(\boldsymbol{\mu}) \end{bmatrix} \xrightarrow{D} H(\boldsymbol{\mu})\mathbf{W} = \begin{bmatrix} \frac{\partial h(\mathbf{x})}{\partial x_1}|_{\mathbf{x}=\boldsymbol{\mu}} & \frac{\partial h(\mathbf{x})}{\partial x_2}|_{\mathbf{x}=\boldsymbol{\mu}} & \cdots & \frac{\partial h(\mathbf{x})}{\partial x_k}|_{\mathbf{x}=\boldsymbol{\mu}} \\ \frac{\partial g(\mathbf{x})}{\partial x_1}|_{\mathbf{x}=\boldsymbol{\mu}} & \frac{\partial g(\mathbf{x})}{\partial x_2}|_{\mathbf{x}=\boldsymbol{\mu}} & \cdots & \frac{\partial g(\mathbf{x})}{\partial x_k}|_{\mathbf{x}=\boldsymbol{\mu}} \end{bmatrix}\mathbf{W}$$

$$= \begin{bmatrix} [\nabla h(\boldsymbol{\mu})]^T \\ [\nabla g(\boldsymbol{\mu})]^T \end{bmatrix}\mathbf{W}.$$

Άρα, $\sqrt{n}\begin{bmatrix} h(\mathbf{m}_n) - h(\boldsymbol{\mu}) \\ g(\mathbf{m}_n) - g(\boldsymbol{\mu}) \end{bmatrix} \xrightarrow{D} N(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \mathbf{J}_\mu\boldsymbol{\Sigma}\mathbf{J}_\mu^T),$

όπου $\mathbf{J}_\mu = \begin{bmatrix} [\nabla h(\boldsymbol{\mu})]^T \\ [\nabla g(\boldsymbol{\mu})]^T \end{bmatrix} \in \mathbb{R}^{2xk}.$

Θα υπολογίσουμε την ποσότητα $\mathbf{J}_\mu\boldsymbol{\Sigma}\mathbf{J}_\mu^T$. Έχουμε

$$\mathbf{J}_\mu\boldsymbol{\Sigma}\mathbf{J}_\mu^T = \begin{bmatrix} [\nabla h(\boldsymbol{\mu})]^T \\ [\nabla g(\boldsymbol{\mu})]^T \end{bmatrix}\boldsymbol{\Sigma}\begin{bmatrix} \nabla h(\boldsymbol{\mu}) & \nabla g(\boldsymbol{\mu}) \end{bmatrix} = \begin{bmatrix} [\nabla h(\boldsymbol{\mu})]^T\boldsymbol{\Sigma}[\nabla h(\boldsymbol{\mu})] & [\nabla h(\boldsymbol{\mu})]^T\boldsymbol{\Sigma}[\nabla g(\boldsymbol{\mu})] \\ [\nabla g(\boldsymbol{\mu})]^T\boldsymbol{\Sigma}[\nabla h(\boldsymbol{\mu})] & [\nabla g(\boldsymbol{\mu})]^T\boldsymbol{\Sigma}[\nabla g(\boldsymbol{\mu})] \end{bmatrix}.$$

Από τη μονοδιάστατη περίπτωση καταλήξαμε ότι
$[\nabla g(\boldsymbol{\mu})]^T \boldsymbol{\Sigma} [\nabla g(\boldsymbol{\mu})] = \mu_{2k} - \mu_k^2 + k\mu_{k-1}[k\sigma^2\mu_{k-1} - 2\mu_{k+1}] = v_k^2$.
Αντίστοιχα προκύπτει ότι
$[\nabla h(\boldsymbol{\mu})]^T \boldsymbol{\Sigma} [\nabla h(\boldsymbol{\mu})] = \mu_{2r} - \mu_r^2 + r\mu_{r-1}[r\sigma^2\mu_{r-1} - 2\mu_{r+1}] = v_r^2$.
Επιπλέον,

$$[\nabla h(\boldsymbol{\mu})]^T \boldsymbol{\Sigma} [\nabla g(\boldsymbol{\mu})] = \begin{bmatrix} h_1 & 0 & ... & 0 & h_r & 0 & ... & 0 \end{bmatrix} \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1k} \\ \sigma_{12} & \sigma_{22} & \cdots & \sigma_{2k} \\ \cdots & \cdots & \cdots & \cdots \\ \sigma_{1r} & \sigma_{2r} & \cdots & \sigma_{rk} \\ \cdots & \cdots & \cdots & \cdots \\ \sigma_{1k} & \sigma_{2k} & \cdots & \sigma_{kk} \end{bmatrix} \begin{bmatrix} g_1 \\ 0 \\ ... \\ 0 \\ g_k \end{bmatrix}$$

$$= \begin{bmatrix} h_1\sigma_{11} + h_r\sigma_{1r} & h_1\sigma_{12} + h_r\sigma_{2r} & ... & h_1\sigma_{1r} + h_r\sigma_{rr} & ... & h_1\sigma_{1k} + h_r\sigma_{rk} \end{bmatrix} \begin{bmatrix} g_1 \\ 0 \\ ... \\ 0 \\ g_k \end{bmatrix}$$

$= (h_1\sigma_{11} + h_r\sigma_{1r})g_1 + (h_1\sigma_{1k} + h_r\sigma_{rk})g_k$
$= h_1g_1\sigma_{11} + h_rg_1\sigma_{1r} + h_1g_k\sigma_{1k} + h_rg_k\sigma_{rk}$.
Και επειδή

$$\begin{array}{ll} h_1 = -r\mu_{r-1} & \sigma_{11} = \mu_2 - \mu_1^2 = \mu_2 = \sigma^2 \\ h_r = 1 & \sigma_{1r} = \mu_{r+1} - \mu_1\mu_r = \mu_{r+1} \\ g_1 = -k\mu_{k-1} & \sigma_{1k} = \mu_{k+1} - \mu_1\mu_k = \mu_{k+1} \\ g_k = 1 & \sigma_{rk} = \mu_{r+k} - \mu_k\mu_r \end{array}$$

με αντικατάσταση παίρνουμε

$$\begin{aligned} [\nabla h(\boldsymbol{\mu})]^T \boldsymbol{\Sigma} [\nabla g(\boldsymbol{\mu})] &= h_1g_1\sigma_{11} + h_rg_1\sigma_{1r} + h_1g_k\sigma_{1k} + h_rg_k\sigma_{rk} \\ &= (-r\mu_{r-1})(-k\mu_{k-1})\sigma^2 + 1(-k\mu_{k-1})\mu_{r+1} \\ &\quad +(-r\mu_{r-1}) \cdot 1\mu_{k+1} + 1 \cdot 1(\mu_{r+k} - \mu_k\mu_r) \\ &= \mu_{r+k} - \mu_k\mu_r - r\mu_{r-1}\mu_{k+1} - k\mu_{r+1}\mu_{k-1} + rk\sigma^2\mu_{r-1}\mu_{k-1} \\ &= v_{rk}. \end{aligned}$$

Τέλος έχουμε,

$$[\nabla g(\boldsymbol{\mu})]^T \boldsymbol{\Sigma} [\nabla h(\boldsymbol{\mu})] = \begin{bmatrix} g_1 & 0 & ... & 0 & g_k \end{bmatrix} \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1k} \\ \sigma_{12} & \sigma_{22} & \cdots & \sigma_{2k} \\ \cdots & \cdots & \cdots & \cdots \\ \sigma_{1r} & \sigma_{2r} & \cdots & \sigma_{rk} \\ \cdots & \cdots & \cdots & \cdots \\ \sigma_{1k} & \sigma_{2k} & \cdots & \sigma_{kk} \end{bmatrix} \begin{bmatrix} h_1 \\ 0 \\ ... \\ 0 \\ h_r \\ 0 \\ ... \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} g_1\sigma_{11} + g_k\sigma_{1k} & g_1\sigma_{12} + g_k\sigma_{2k} & \dots & g_1\sigma_{1r} + g_k\sigma_{rk} & \dots & g_1\sigma_{1k} + g_k\sigma_{kk} \end{bmatrix} \begin{bmatrix} h_1 \\ 0 \\ \dots \\ 0 \\ h_r \\ 0 \\ \dots \\ 0 \end{bmatrix}$$

$= (g_1\sigma_{11} + g_k\sigma_{1k})h_1 + (g_1\sigma_{1r} + g_k\sigma_{rk})h_r$
$= h_1 g_1 \sigma_{11} + h_r g_1 \sigma_{1r} + h_1 g_k \sigma_{1k} + h_r g_k \sigma_{rk}$
$= (-r\mu_{r-1})(-k\mu_{k-1})\sigma^2 + 1(-k\mu_{k-1})\mu_{r+1}$
$+(-r\mu_{r-1}) \cdot 1\mu_{k+1} + 1 \cdot 1(\mu_{r+k} - \mu_k\mu_r)$
$= \mu_{r+k} - \mu_k\mu_r - r\mu_{r-1}\mu_{k+1} - k\mu_{r+1}\mu_{k-1} + rk\sigma^2\mu_{r-1}\mu_{k-1}$
$= v_{kr}.$
Παρατηρούμε ότι $v_{rk} = [\nabla h(\boldsymbol{\mu})]^T \boldsymbol{\Sigma} [\nabla g(\boldsymbol{\mu})] = [\nabla g(\boldsymbol{\mu})]^T \boldsymbol{\Sigma} [\nabla h(\boldsymbol{\mu})] = v_{kr}$.

Επιλογικά, καταλήξαμε ότι $\sqrt{n} \begin{bmatrix} h(\mathbf{m}_n) - h(\boldsymbol{\mu}) \\ g(\mathbf{m}_n) - g(\boldsymbol{\mu}) \end{bmatrix} \xrightarrow{D} N(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} v_r^2 & v_{rk} \\ v_{rk} & v_k^2 \end{bmatrix})$.
Και εξ' ορισμού με αντικατάσταση στις συναρτήσεις h,g παίρνουμε $h(\mathbf{m}_n) = M_{r,n}, h(\boldsymbol{\mu}) = \mu_r, g(\mathbf{m}_n) = M_{k,n}, g(\boldsymbol{\mu}) = \mu_k$ Άρα, αποδείξαμε το παρακάτω θεώρημα:

Θεώρημα 4.2.1. *Εάν $E[|X|^{2k}] < \infty$, τότε για κάθε $2 \le r < k$ ισχύει ότι*

$$\sqrt{n} \begin{bmatrix} M_{r,n} - \mu_r \\ M_{k,n} - \mu_k \end{bmatrix} \xrightarrow{D} N(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} v_r^2 & v_{rk} \\ v_{rk} & v_k^2 \end{bmatrix}),$$

όπου

$$v_k^2 = \mu_{2k} - \mu_k^2 + k\mu_{k-1}[k\sigma^2\mu_{k-1} - 2\mu_{k+1}] \tag{4.2.7}$$

$$v_r^2 = \mu_{2r} - \mu_r^2 + r\mu_{r-1}[r\sigma^2\mu_{r-1} - 2\mu_{r+1}] \tag{4.2.8}$$

$$v_{rk} = v_{kr} = \mu_{r+k} - \mu_k\mu_r - r\mu_{r-1}\mu_{k+1} - k\mu_{r+1}\mu_{k-1} + rk\sigma^2\mu_{r-1}\mu_{k-1} \tag{4.2.9}$$

Παρατήρηση 4.2.1. Θέτοντας r=k στη σχέση (4.2.9), εύκολα μπορούμε να δούμε ότι ισχύει $v_{kk} = v_k^2$. Πράγματι, έχουμε
$v_{kk} = \mu_{k+k} - \mu_k\mu_k - r\mu_{k-1}\mu_{k+1} - k\mu_{k+1}\mu_{k-1} + k^2\sigma^2\mu_{k-1}\mu_{k-1}$
$= \mu_{2k} - \mu_k^2 - 2k\mu_{k-1}\mu_{k+1} + k^2\sigma^2\mu_{k-1}^2$
$= \mu_{2k} - \mu_k^2 + k\mu_{k-1}[k\sigma^2\mu_{k-1} - 2\mu_{k+1}]$
$= v_k^2.$

Βρήκαμε την οριακή κατανομή του ζεύγους $(M_{r,n}, M_{k,n})$, για $2 \le r < k$. Ακολουθώντας την ίδια διαδικασία, θα βρούμε την οριακή κατανομή της ποσότητας $(\overline{X}_n - \mu, M_{k,n})$.

Θεωρούμε συνάρτηση $h : \mathbb{R}^k \to \mathbb{R}$ η οποία δίνεται από τον τύπο
$h(x_1, \dots, x_k) = x_1$, για $\mathbf{x} = (x_1, \dots, x_k)^T \in \mathbb{R}^k$. Άμεσα παρατηρείται ότι

$\overline{X}_n - \mu = h(\mathbf{m}_n) = m_1$.
Όπως και στις παραπάνω περιπτώσεις ορίζουμε συνάρτηση $g : \mathbb{R}^k \to \mathbb{R}$ τέτοια ώστε
$g(\mathbf{x}) = g(x_1, \dots, x_k) = (-1)^{k-1}(k-1)x_1^k + \sum_{j=2}^{k-1} \binom{k}{j}(-1)^{k-j} x_j x_1^{k-j} + x_k,$
για $\mathbf{x} = (x_1, \dots, x_k)^T \in \mathbb{R}^k$. Οπότε ισχύει $M_{k,n} = g(\mathbf{m}_n)$.
Θεωρώντας όπως και παραπάνω ότι από Πολυδιάστατο Κεντρικό Οριακό Θεώρημα καθώς $n \to \infty$ ισχύει $\sqrt{n}(\mathbf{m}_n - \boldsymbol{\mu}) \xrightarrow{D} N(\mathbf{0}, \boldsymbol{\Sigma})$, εφαρμόζοντας την Πολυδιάστατη μέθοδο Δέλτα πρώτης τάξης (βλ. Θεώρημα 3.1.2.) παίρνουμε
$\sqrt{n}\begin{bmatrix} h(\mathbf{m}_n) - h(\boldsymbol{\mu}) \\ g(\mathbf{m}_n) - g(\boldsymbol{\mu}) \end{bmatrix} \xrightarrow{D} N(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \mathbf{J}_\mu \boldsymbol{\Sigma} \mathbf{J}_\mu^T)$, όπου $\mathbf{J}_\mu = \begin{bmatrix} [\nabla h(\boldsymbol{\mu})]^T \\ [\nabla g(\boldsymbol{\mu})]^T \end{bmatrix} \in \mathbb{R}^{2xk}$.

Βρήκαμε ότι $\mathbf{J}_\mu \boldsymbol{\Sigma} \mathbf{J}_\mu^T = \begin{bmatrix} [\nabla h(\boldsymbol{\mu})]^T \boldsymbol{\Sigma} [\nabla h(\boldsymbol{\mu})] & [\nabla h(\boldsymbol{\mu})]^T \boldsymbol{\Sigma} [\nabla g(\boldsymbol{\mu})] \\ [\nabla g(\boldsymbol{\mu})]^T \boldsymbol{\Sigma} [\nabla h(\boldsymbol{\mu})] & [\nabla g(\boldsymbol{\mu})]^T \boldsymbol{\Sigma} [\nabla g(\boldsymbol{\mu})] \end{bmatrix}$, όπου

$$\nabla g(\boldsymbol{\mu}) = \begin{bmatrix} \frac{\partial g(x_1,\dots,x_k)}{\partial x_1}|_{\mathbf{x}=\boldsymbol{\mu}} \\ \frac{\partial g(x_1,\dots,x_k)}{\partial x_2}|_{\mathbf{x}=\boldsymbol{\mu}} \\ \dots \\ \frac{\partial g(x_1,\dots,x_k)}{\partial x_k}|_{\mathbf{x}=\boldsymbol{\mu}} \end{bmatrix} = \begin{bmatrix} -k\mu_{k-1} \\ 0 \\ \dots \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} g_1 \\ 0 \\ \dots \\ 0 \\ g_k \end{bmatrix} \text{ και}$$

$$\nabla h(\boldsymbol{\mu}) = \begin{bmatrix} \frac{\partial h(x_1,\dots,x_k)}{\partial x_1}|_{\mathbf{x}=\boldsymbol{\mu}} \\ \frac{\partial h(x_1,\dots,x_k)}{\partial x_2}|_{\mathbf{x}=\boldsymbol{\mu}} \\ \dots \\ \frac{\partial h(x_1,\dots,x_k)}{\partial x_k}|_{\mathbf{x}=\boldsymbol{\mu}} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ \dots \\ 0 \end{bmatrix} = \begin{bmatrix} h_1 \\ 0 \\ \dots \\ 0 \end{bmatrix}.$$

Υπολογίζουμε
$[g(\boldsymbol{\mu})]^T \boldsymbol{\Sigma} [\nabla g(\boldsymbol{\mu})] = v_k^2$.

$$[\nabla h(\boldsymbol{\mu})]^T \boldsymbol{\Sigma} [\nabla h(\boldsymbol{\mu})] = \begin{bmatrix} h_1 & 0 & \dots & 0 \end{bmatrix} \begin{bmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1k} \\ \sigma_{12} & \sigma_{22} & \dots & \sigma_{2k} \\ \dots & \dots & \dots & \dots \\ \sigma_{1k} & \sigma_{2k} & \dots & \sigma_{kk} \end{bmatrix} \begin{bmatrix} h_1 \\ 0 \\ \dots \\ 0 \end{bmatrix}$$
$$= h_1^2 \sigma_{11}.$$

$$[\nabla h(\boldsymbol{\mu})]^T \boldsymbol{\Sigma} [\nabla g(\boldsymbol{\mu})] = \begin{bmatrix} h_1 & 0 & \dots & 0 \end{bmatrix} \begin{bmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1k} \\ \sigma_{12} & \sigma_{22} & \dots & \sigma_{2k} \\ \dots & \dots & \dots & \dots \\ \sigma_{1k} & \sigma_{2k} & \dots & \sigma_{kk} \end{bmatrix} \begin{bmatrix} g_1 \\ 0 \\ \dots \\ 0 \\ g_k \end{bmatrix}$$

$$= \begin{bmatrix} h_1\sigma_{11} & \dots & h_1\sigma_{1k} \end{bmatrix} \begin{bmatrix} g_1 \\ 0 \\ \dots \\ 0 \\ g_k \end{bmatrix} = h_1 g_1 \sigma_{11} + h_1 g_k \sigma_{1k}.$$

$$[\nabla g(\boldsymbol{\mu})]^T \boldsymbol{\Sigma} [\nabla h(\boldsymbol{\mu})] = \begin{bmatrix} g_1 & 0 & \dots & 0 & g_k \end{bmatrix} \begin{bmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1k} \\ \sigma_{12} & \sigma_{22} & \dots & \sigma_{2k} \\ \dots & \dots & \dots & \dots \\ \sigma_{1k} & \sigma_{2k} & \dots & \sigma_{kk} \end{bmatrix} \begin{bmatrix} h_1 \\ 0 \\ \dots \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} g_1\sigma_{11} + g_k\sigma_{1k} & \dots & g_1\sigma_{1k} + g_k\sigma_{kk} \end{bmatrix} \begin{bmatrix} h_1 \\ 0 \\ \dots \\ 0 \end{bmatrix}$$

$$= h_1 g_1 \sigma_{11} + h_1 g_k \sigma_{1k}.$$

Αφού ισχύουν $h_1 = 1$ $\quad \sigma_{11} = \mu_2 - \mu_1^2 = \mu_2 = \sigma^2$

$g_1 = -k\mu_{k-1}$ $\quad \sigma_{1k} = \mu_{\kappa+1} - \mu_1\mu_k = \mu_{\kappa+1}$

$g_k = 1,$

με αντικατάσταση προκύπτει

$[g(\boldsymbol{\mu})]^T \boldsymbol{\Sigma} [\nabla g(\boldsymbol{\mu})] = v_k^2.$

$[\nabla h(\boldsymbol{\mu})]^T \boldsymbol{\Sigma} [\nabla h(\boldsymbol{\mu})] = h_1^2 \sigma_{11} = \sigma^2.$

$$\begin{aligned}[\nabla h(\boldsymbol{\mu})]^T \boldsymbol{\Sigma} [\nabla g(\boldsymbol{\mu})] = [\nabla g(\boldsymbol{\mu})]^T \boldsymbol{\Sigma} [\nabla h(\boldsymbol{\mu})] &= h_1 g_1 \sigma_{11} + h_1 g_k \sigma_{1k} \\ &= 1(-k\mu_{k-1})\sigma^2 + 1 \cdot 1\mu_{k+1}. \\ &= \mu_{k+1} - k\sigma^2\mu_{k-1}.\end{aligned}$$

Συνεπώς, καταλήγουμε στο ακόλουθο θεώρημα:

Θεώρημα 4.2.2. *Εάν $E[|X|^{2k}] < \infty$, τότε ισχύει ότι*

$$\sqrt{n} \begin{bmatrix} \overline{X}_n - \mu \\ M_{k,n} - \mu_k \end{bmatrix} \xrightarrow{D} N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma^2 & \mu_{k+1} - k\sigma^2\mu_{k-1} \\ \mu_{k+1} - k\sigma^2\mu_{k-1} & v_\kappa^2 \end{bmatrix} \right),$$

όπου η v_k δίνεται από τη σχέση (4.2.7).

Οριακή κατανομή του διανύσματος δειγματικών κεντρικών ροπών $(\overline{X}_n - \mu, \dots, M_{k,n})^T$

Για την εύρεση της οριακής κατανομής του διανύσματος $(\overline{X}_n - \mu, M_{2,n}, \dots, M_{k,n})^T$ θα εργαστούμε με παρόμοιο τρόπο όπως παραπάνω.

Όπως στις προηγούμενες περιπτώσεις, εφαρμόζοντας το το Πολυδιάστατο Κεντρικό Οριακό Θεώρημα (βλ. Θεώρημα 1.4.6.) έχουμε καθώς $n \to \infty$,

$\sqrt{n}(\mathbf{m}_n - \boldsymbol{\mu}) \xrightarrow{D} \mathbf{W} \sim N(\mathbf{0}, \boldsymbol{\Sigma}).$

Επιπλέον, ορίζουμε συνάρτηση $\mathbf{g} = (g_1, g_2, \dots, g_k) : \mathbb{R}^k \to \mathbb{R}^k$ έτσι ώστε να ισχύει $g_r(\mathbf{m}_n) = M_{r,n}, \forall r \in \{1, 2, \dots, k\}$. Δηλαδή, η $\mathbf{g}$ ορίζεται ως

$$g_1(x_1, \dots, x_k) = x_1$$

$$g_r(x_1, \dots, x_k) = (-1)^{r-1}(r-1)x_1^r + \sum_{j=2}^{r-1} \tbinom{r}{j} (-1)^{r-j} x_j x_1^{r-j} + x_r,$$

$$\text{για } r \in \{2, \dots, k-1\}$$

$$g_k(x_1, \dots, x_k) = (-1)^{k-1}(k-1)x_1^k + \sum_{j=2}^{k-1} \tbinom{k}{j} (-1)^{k-j} x_j x_1^{k-j} + x_k$$

και $g_i : \mathbb{R}^k \to \mathbb{R}$, πολυωνυμικές συναρτήσεις, $\forall i \in \{1, 2, ..., k\}$.
Επομένως, από Πολυδιάστατη μέθοδο Δέλτα πρώτης τάξης από τον $\mathbb{R}^k$ στον $\mathbb{R}^k$ έχουμε ότι

$$\sqrt{n}\begin{bmatrix} g_1(\mathbf{m}_n) - g_1(\boldsymbol{\mu}) \\ g_2(\mathbf{m}_n) - g_2(\boldsymbol{\mu}) \\ ... \\ g_k(\mathbf{m}_n) - g_k(\boldsymbol{\mu}) \end{bmatrix} \xrightarrow{D} H(\boldsymbol{\mu})\mathbf{W} = \begin{bmatrix} \frac{\partial g_1(\mathbf{x})}{\partial x_1}\Big|_{\mathbf{x}=\boldsymbol{\mu}} & \frac{\partial g_1(\mathbf{x})}{\partial x_2}\Big|_{\mathbf{x}=\boldsymbol{\mu}} & \cdots & \frac{\partial g_1(\mathbf{x})}{\partial x_k}\Big|_{\mathbf{x}=\boldsymbol{\mu}} \\ \frac{\partial g_2(\mathbf{x})}{\partial x_1}\Big|_{\mathbf{x}=\boldsymbol{\mu}} & \frac{\partial g_2(\mathbf{x})}{\partial x_2}\Big|_{\mathbf{x}=\boldsymbol{\mu}} & \cdots & \frac{\partial g_2(\mathbf{x})}{\partial x_k}\Big|_{\mathbf{x}=\boldsymbol{\mu}} \\ ... & ... & ... & ... \\ \frac{\partial g_k(\mathbf{x})}{\partial x_1}\Big|_{\mathbf{x}=\boldsymbol{\mu}} & \frac{\partial g_k(\mathbf{x})}{\partial x_2}\Big|_{\mathbf{x}=\boldsymbol{\mu}} & \cdots & \frac{\partial g_k(\mathbf{x})}{\partial x_k}\Big|_{\mathbf{x}=\boldsymbol{\mu}} \end{bmatrix}\mathbf{W}$$

$$= \begin{bmatrix} [\nabla g_1(\boldsymbol{\mu})]^T \\ [\nabla g_2(\boldsymbol{\mu})]^T \\ ... \\ [\nabla g_k(\boldsymbol{\mu})]^T \end{bmatrix}\mathbf{W}.$$

Δηλαδή,

$$\sqrt{n}\begin{bmatrix} g_1(\mathbf{m}_n) - g_1(\boldsymbol{\mu}) \\ g_2(\mathbf{m}_n) - g_2(\boldsymbol{\mu}) \\ ... \\ g_k(\mathbf{m}_n) - g_k(\boldsymbol{\mu}) \end{bmatrix} = \sqrt{n}\begin{bmatrix} \overline{X}_n - \mu \\ M_{2,n} - \sigma^2 \\ ... \\ M_{k,n} - \mu_k \end{bmatrix} \xrightarrow{D} N(\begin{bmatrix} 0 \\ 0 \\ ... \\ 0 \end{bmatrix}, \mathbf{J}_\mu \boldsymbol{\Sigma} \mathbf{J}_\mu^T),$$

όπου $\mathbf{J}_\mu = \begin{bmatrix} [\nabla g_1(\boldsymbol{\mu})]^T \\ [\nabla g_2(\boldsymbol{\mu})]^T \\ ... \\ [\nabla g_k(\boldsymbol{\mu})]^T \end{bmatrix}$.

Θα υπολογίσουμε την ποσότητα $\mathbf{J}_\mu \boldsymbol{\Sigma} \mathbf{J}_\mu^T$. Έχουμε

$$\mathbf{J}_\mu \boldsymbol{\Sigma} \mathbf{J}_\mu^T = \begin{bmatrix} [\nabla g_1(\boldsymbol{\mu})]^T \\ [\nabla g_2(\boldsymbol{\mu})]^T \\ ... \\ [\nabla g_k(\boldsymbol{\mu})]^T \end{bmatrix} \boldsymbol{\Sigma} \left[[\nabla g_1(\boldsymbol{\mu})] \quad [\nabla g_2(\boldsymbol{\mu})] \quad ... \quad [\nabla g_k(\boldsymbol{\mu})] \right]$$

$$= \begin{bmatrix} [\nabla g_1(\boldsymbol{\mu})]^T \boldsymbol{\Sigma} [\nabla g_1(\boldsymbol{\mu})] & ... & [\nabla g_1(\boldsymbol{\mu})]^T \boldsymbol{\Sigma} [\nabla g_k(\boldsymbol{\mu})] \\ [\nabla g_2(\boldsymbol{\mu})]^T \boldsymbol{\Sigma} [\nabla g_1(\boldsymbol{\mu})] & ... & [\nabla g_2(\boldsymbol{\mu})]^T \boldsymbol{\Sigma} [\nabla g_k(\boldsymbol{\mu})] \\ ... & & \\ [\nabla g_k(\boldsymbol{\mu})]^T \boldsymbol{\Sigma} [\nabla g_1(\boldsymbol{\mu})] & ... & [\nabla g_k(\boldsymbol{\mu})]^T \boldsymbol{\Sigma} [\nabla g_k(\boldsymbol{\mu})] \end{bmatrix}.$$

Από τις προηγούμενες υποπεριπτώσεις βρήκαμε ότι
$[\nabla g_1(\boldsymbol{\mu})]^T \boldsymbol{\Sigma} [\nabla g_1(\boldsymbol{\mu})] = \sigma^2$.
$[\nabla g_1(\boldsymbol{\mu})]^T \boldsymbol{\Sigma} [\nabla g_k(\boldsymbol{\mu})] = \mu_{k+1} - k\sigma^2 \mu_{k-1}$.
$[\nabla g_1(\boldsymbol{\mu})]^T \boldsymbol{\Sigma} [\nabla g_r(\boldsymbol{\mu})] = \mu_{r+1} - r\sigma^2 \mu_{r-1}, r \in \{2, ..., k-1\}$.
$[\nabla g_r(\boldsymbol{\mu})]^T \boldsymbol{\Sigma} [\nabla g_r(\boldsymbol{\mu})] = v_r^2$.
$[\nabla g_k(\boldsymbol{\mu})]^T \boldsymbol{\Sigma} [\nabla g_k(\boldsymbol{\mu})] = v_k^2$.
$[\nabla g_r(\boldsymbol{\mu})]^T \boldsymbol{\Sigma} [\nabla g_k(\boldsymbol{\mu})] = [\nabla g_k(\boldsymbol{\mu})]^T \boldsymbol{\Sigma} [\nabla g_r(\boldsymbol{\mu})] = v_{rk} = v_{kr}$.
Καταλήξαμε λοιπόν στο παρακάτω:

Θεώρημα 4.2.3. *Με την προϋπόθεση ότι* $E[X^{2\kappa}] < \infty$ *έπεται ότι*

$$\sqrt{n}\begin{bmatrix}\overline{X}_n-\mu\\ M_{2,n}-\sigma^2\\ ...\\ M_{k,n}-\mu_k\end{bmatrix}\xrightarrow{D}N(\begin{bmatrix}0\\0\\...\\0\end{bmatrix},\begin{bmatrix}\sigma^2 & \mu_3 & \mu_4-3\sigma^2 & ... & \mu_{k+1}-k\sigma^2\mu_{k-1}\\ \mu_3 & v_2^2 & v_{23} & ... & v_{2k}\\ \mu_4-3\sigma^2 & v_{23} & v_3^2 & ... & v_{3k}\\ ... & ... & ... & ... & ...\\ \mu_{k+1}-k\sigma^2\mu_{k-1} & v_{2k} & v_{3k} & ... & v_k^2\end{bmatrix}),$$

όπου οι ποσότητες v_k^2, v_r^2, v_{rk}, για $r \in \{2,3,...,k-1\}$ δίνονται από τις σχέσεις (4.2.7), (4.2.8), (4,2,9) αντίστοιχα.

Παρατήρηση 4.2.2. Η αρχική υπόθεση ότι $E[X^{2\kappa}] < \infty$, η οποία δόθηκε στην αρχή του κεφαλαίου είχε στόχο να ορίζεται η ποσότητα $\mu_{2\kappa}$ και άρα η v_k^2, εφόσον $v_k^2 = \mu_{2k} - \mu_k^2 + k\mu_{k-1}[k\sigma^2\mu_{k-1} - 2\mu_{k+1}]$.Από παρατήρηση 1.4.1. όλες οι υπόλοιπες κεντρικές ροπές γύρω από το μέσο οι οποίες εμφανίζονται μπορούν να ορισθούν και το ίδιο ισχύει και για τις ποσότητες $v_{ij}, \forall i,j \in \{2,...,k\}$.

Πρόταση 4.2.1. *Εάν ισχύει $E[X^{2k}] < \infty$, ο παραπάνω πίνακας διασποράς της οριακής κατανομής των $[\overline{X}_n - \mu, M_{2,n},...,M_{k,n}]^T$, δηλαδή ο πίνακας*

$$\begin{bmatrix}\sigma^2 & \mu_3 & \mu_4-3\sigma^2 & ... & \mu_{k+1}-k\sigma^2\mu_{k-1}\\ \mu_3 & v_2^2 & v_{23} & ... & v_{2k}\\ \mu_4-3\sigma^2 & v_{23} & v_3^2 & ... & v_{3k}\\ ... & ... & ... & ... & ...\\ \mu_{k+1}-k\sigma^2\mu_{k-1} & v_{2k} & v_{3k} & ... & v_k^2\end{bmatrix}$$

ταυτίζεται με τον πίνακα διασποράς του τυχαίου διανύσματος

$$\begin{bmatrix}Y\\ Y^2\\ Y^3-3\sigma^2Y\\ Y^4-4\mu_3Y\\ ...\\ Y^k-k\mu_{k-1}Y\end{bmatrix}=\begin{bmatrix}X-\mu\\ (X-\mu)^2\\ (X-\mu)^3-3\sigma^2(X-\mu)\\ (X-\mu)^4-4\mu_3(X-\mu)\\ ...\\ (X-\mu)^k-k\mu_{k-1}(X-\mu)\end{bmatrix}.$$

Απόδειξη. Για την απόδειξη, αρκεί να δείξουμε ότι για $r,k \in \{2,3,...\}$ ισχύουν οι κάτωθι:

(i) $\sigma^2 = Var[X-\mu]$

(ii) $\mu_{k+1} - k\sigma^2\mu_{k-1} = Cov[X-\mu, (X-\mu)^k - k\mu_{k-1}(X-\mu)]$

(iii) $v_k^2 = Var[(X-\mu)^k - k\mu_{k-1}(X-\mu)]$

(iv) $v_{rk} = Cov[(X-\mu)^r - r\mu_{r-1}(X-\mu), (X-\mu)^k - k\mu_{k-1}(X-\mu)]$

Έχουμε,

(i) Η απόδειξη είναι προφανής από ορισμό διασποράς.

(ii) $$\begin{aligned}
\mu_{k+1} - k\sigma^2\mu_{k-1} &= \mu_{k+1} - k\mu_{k-1}\mu_2 \\
&= E[(X-\mu)^{k+1}] - k\mu_{k-1}E[(X-\mu)^2] \\
&= E[(X-\mu)^{k+1} - k\mu_{k-1}(X-\mu)^2] \\
&\overset{E[X-\mu]=0}{=} E[(X-\mu)^{k+1} - k\mu_{k-1}(X-\mu)^2] \\
&\quad -E[(X-\mu)]E[(X-\mu)^k - k\mu_{k-1}(X-\mu)] \\
&= E[(X-\mu)((X-\mu)^k - k\mu_{k-1}(X-\mu))] \\
&\quad -E[(X-\mu)]E[(X-\mu)^k - k\mu_{k-1}(X-\mu)] \\
&\overset{Cov[X,Y]=E[XY]-E[X]E[Y]}{=} Cov[(X-\mu),(X-\mu)^k - k\mu_{k-1}(X-\mu)].
\end{aligned}$$

(iii) $$\begin{aligned}
v_k^2 &= \mu_{2k} - \mu_k^2 + k\mu_{k-1}[k\sigma^2\mu_{k-1} - 2\mu_{k+1}] \\
&= \mu_{2k} - \mu_k^2 + k^2\mu_{k-1}^2\sigma^2 - 2k\mu_{k-1}\mu_{k+1} \\
&= E[(X-\mu)^{2k}] - E[(X-\mu)^k]^2 + k^2\mu_{k-1}^2 Var[X] \\
&\quad -2kE[(X-\mu)^{k+1}]\mu_{k-1} \\
&= Var[(X-\mu)^k] + k^2\mu_{k-1}^2 Var[X] - 2kE[(X-\mu)^k(X-\mu)]\mu_{k-1} \\
&\overset{E[X-\mu]=0}{=} Var[(X-\mu)^k] + k^2\mu_{k-1}^2 Var[X-\mu] \\
&\quad -2k[E[(X-\mu)^k(X-\mu)] - E[(X-\mu)^k]E[(X-\mu)]]\mu_{k-1} \\
&= Var[(X-\mu)^k] + Var[k\mu_{k-1}(X-\mu)] - 2k\mu_{k-1}Cov[(X-\mu)^k,(X-\mu)] \\
&= Var[(X-\mu)^k] + Var[k\mu_{k-1}(X-\mu)] - 2Cov[(X-\mu)^k, k\mu_{k-1}(X-\mu)] \\
&= Var[(X-\mu)^k - k\mu_{k-1}(X-\mu)].
\end{aligned}$$

(iv) $$\begin{aligned}
v_{rk} &= \mu_{r+k} - k\mu_{r+1}\mu_{k-1} - r\mu_{r-1}\mu_{k+1} + rk\sigma^2\mu_{r-1}\mu_{k-1} - \mu_k\mu_r \\
&= E[(X-\mu)^{r+k}] - k\mu_{k-1}E[(X-\mu)^{r+1}] - r\mu_{r-1}E[(X-\mu)^{k+1}] \\
&\quad +kr\mu_{r-1}\mu_{k-1}E[(X-\mu)^2] - E[(X-\mu)^r]E[(X-\mu)^k] \\
&= E[(X-\mu)^{r+k} - k\mu_{k-1}(X-\mu)^{r+1} - r\mu_{r-1}(X-\mu)^{k+1} \\
&\quad +kr\mu_{r-1}\mu_{k-1}(X-\mu)^2] - E[(X-\mu)^r]E[(X-\mu)^k] \\
&\overset{E[X-\mu]=0}{=} E[(X-\mu)^k[(X-\mu)^r - r\mu_{r-1}(X-\mu)] \\
&\quad -k\mu_{k-1}(X-\mu)[(X-\mu)^r - r\mu_{r-1}(X-\mu)]] \\
&\quad -[E[(X-\mu)^r - r\mu_{r-1}E[X-\mu]][E[(X-\mu)^k] - k\mu_{k-1}E[X-\mu]] \\
&= E[[(X-\mu)^r - r\mu_{r-1}(X-\mu)][(X-\mu)^k - k\mu_{k-1}(X-\mu)]] \\
&\quad -[E[(X-\mu)^r - r\mu_{r-1}E[X-\mu]][E[(X-\mu)^k] - k\mu_{k-1}E[X-\mu]] \\
&= Cov[(X-\mu)^r - r\mu_{r-1}(X-\mu),(X-\mu)^k - k\mu_{k-1}(X-\mu)]. \quad \square
\end{aligned}$$

Παρατήρηση 4.2.3. Αφού αποδείξαμε ότι $v_k^2 = Var[(X-\mu)^k - k\mu_{k-1}(X-\mu)]$, συνεπάγεται ότι $v_k^2 \geq 0$. Η Πρόταση 4.2.1. είναι ιδιαίτερα σημαντική, διότι επιτρέπει τη μελέτη των ιδιαζουσών κατανομών, δηλαδή των κατανομών για τις οποίες $v_k^2 = 0$. Συγκεκριμένα, $v_k = 0 \Leftrightarrow \exists c \in \mathbb{R} : (X-\mu)^k = k\mu_{k-1}(X-\mu) + c$, με πιθανότητα 1.

4.3 Σύγκλιση δειγματικών κεντρικών ροπών

Στην προηγούμενη ενότητα αποδείξαμε ότι η οριακή κατανομή του διανύσματος των δειγματικών κεντρικών ροπών $[M_{1,n}, ..., M_{k,n}]'$ είναι η Κανονική Κατανομή με τον πίνακα διασποράς που αναφέρθηκε παραπάνω. Δηλαδή,

$$\sqrt{n}\begin{bmatrix}\overline{X}_n-\mu\\ M_{2,n}-\sigma^2\\ ...\\ M_{k,n}-\mu_k\end{bmatrix} \xrightarrow{D} N(\begin{bmatrix}0\\0\\...\\0\end{bmatrix}, \begin{bmatrix}\sigma^2 & \mu_3 & \mu_4-3\sigma^2 & ... & \mu_{k+1}-k\sigma^2\mu_{k-1}\\ \mu_3 & v_2^2 & v_{23} & ... & v_{2k}\\ \mu_4-3\sigma^2 & v_{23} & v_3^2 & ... & v_{3k}\\ ... & ... & ... & ... & ...\\ \mu_{k+1}-k\sigma^2\mu_{k-1} & v_{2k} & v_{3k} & ... & v_k^2\end{bmatrix}).$$

Εντούτοις, η προαναφερθείσα ασθενής σύγκλιση δε συνεπάγεται τη σύγκλιση των επιμέρους δειγματικών κεντρικων ροπών στις αντίστοιχες ποσότητες, καθώς $n \to \infty$.

Για να αληθεύει η σύγκλιση των αντίστοιχων δειγματικών κεντρικών ροπών θα πρέπει να ισχύουν:

$E[\sqrt{n}(\overline{X}_n-\mu)] \to 0, E[\sqrt{n}(M_{i,n}-\mu_i)] \to 0, \forall i \in \{2,3,...,k\}$,
$Var[\sqrt{n}(\overline{X}_n-\mu)] \to \sigma^2, Var[\sqrt{n}(M_{i,n}-\mu_i)] \to v_k^2, \forall i \in \{2,3,...,k\}$,
$Cov[\overline{X}_n-\mu),(M_{i,n}-\mu_i)] \to \mu_{i+1}-i\sigma^2\mu_{i-1}, \forall i \in \{2,3,...,k\}$,
$Cov[(M_{r,n}-\mu_r)),(M_{i,n}-\mu_i)] \to v_{ri}, \forall r \leq i \leq k \in \{2,3,...,k\}$.

Για να δείξουμε ότι σχέσεις αυτές πράγματι αληθεύουν, αρκεί να αποδείξουμε τις παρακάτω τρεις προτάσεις.

Πρόταση 4.3.1. *Εάν $E[X^k] < \infty$, για κάποιο $k \in \{2,3,...\}$, τότε, καθώς $n \to \infty$*

$$E[M_{k,n}] = \mu_k + o(\frac{1}{\sqrt{n}}). \tag{4.3.1}$$

Απόδειξη. Θα δείξουμε ότι καθώς $n \to \infty$,

$$\begin{aligned} & E[M_{k,n}] = \mu_k + o(\tfrac{1}{\sqrt{n}})\\ \Leftrightarrow\ & E[M_{k,n}-\mu_k] = o(\tfrac{1}{\sqrt{n}})\\ \overset{O\rho.1.6.1.}{\Leftrightarrow}\ & \sqrt{n}E[M_{k,n}-\mu_k] \to 0\\ \Leftrightarrow\ & E[\sqrt{n}(M_{k,n}-\mu_k)] \to 0.\end{aligned}$$

Από τη σχέση (4.1.6) προκύπτει ότι

$$\begin{aligned} M_{k,n}-\mu_k &= m_{k,n} + (-1)^{k-1}(k-1)m_{1,n}^k + \sum_{j=2}^{k-1}\tbinom{k}{j}(-1)^{k-j}m_{j,n}m_{1,n}^{k-j} - \mu_k\\ &= (m_{k,n}-\mu_k) + (-1)^{k-1}(k-1)m_{1,n}^k + \sum_{j=2}^{k-1}\tbinom{k}{j}(-1)^{k-j}m_{j,n}m_{1,n}^{k-j}.\end{aligned}$$

Οπότε, θα αποδείξουμε ότι καθώς $n \to \infty$,

$$E[\sqrt{n}(M_{k,n}-\mu_k)] \to 0$$

$$\Leftrightarrow E[\sqrt{n}(m_{k,n}-\mu_k) + (-1)^{k-1}(k-1)\sqrt{n}m_{1,n}^k + \sum_{j=2}^{k-1}\tbinom{k}{j}(-1)^{k-j}\sqrt{n}m_{j,n}m_{1,n}^{k-j}]$$

$= \sqrt{n}E[(m_{k,n}-\mu_k)]+(-1)^{k-1}(k-1)\sqrt{n}E[m_{1,n}^k]+\sum_{j=2}^{k-1}\binom{k}{j}(-1)^{k-j}\sqrt{n}E[m_{1,n}^{k-j}m_{j,n}]]$
$\to 0$.
Επομένως, αρκεί να δείξουμε ότι για $n \to \infty$ αληθεύουν:
(i) $\sqrt{n}E[(m_{k,n} - \mu_k)] = 0$,
(ii) $\sqrt{n}E[m_{1,n}^k] \to 0$,
(iii) $\sqrt{n}E[m_{1,n}^{k-j}m_{j,n}] \to 0$,   όπου $j = 2, \dots, k-2$ και $k \geq 4$,
(iv) $\sqrt{n}E[m_{1,n}m_{k-1,n}] \to 0$,   όπου $k \geq 3$.

Πράγματι,
(i) $\sqrt{n}E[(m_{k,n} - \mu_k)] = \sqrt{n}E[(m_{k,n}] - \sqrt{n}E[\mu_k] = \sqrt{n}\mu_k - \sqrt{n}\mu_k = 0$,
δεδομένου ότι $E[(m_{k,n}] = E[\frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^k] = \frac{1}{n}\sum_{i=1}^{n}E[(X_i - \mu)^k] = \mu_k$.

$$\begin{aligned}
\text{(iv) } \sqrt{n}E[m_{1,n}m_{k-1,n}] &= E[\tfrac{1}{n}\sum_{i=1}^{n}(X_i - \mu)\tfrac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^{k-1}] \\
&= E[\tfrac{1}{n^2}\sum_{i=1}^{n}(X_i - \mu)\sum_{i=1}^{n}(X_i - \mu)^{k-1}] \\
&= \tfrac{1}{n^2}E[\sum_{i=1}^{n}(X_i - \mu)^k + \sum_{i=1}^{n}\sum_{j\neq i}(X_i - \mu)(X_j - \mu)^{k-1}] \\
&= \tfrac{1}{n^2}E[\sum_{i=1}^{n}(X_i-\mu)^k]+\tfrac{1}{n^2}E[\sum_{i=1}^{n}\sum_{j\neq i}(X_i-\mu)(X_j-\mu)^{k-1}] \\
&= \tfrac{1}{n^2}n\mu_k + E[\tfrac{1}{n}\sum_{i=1}^{n}(X_i - \mu)\tfrac{1}{n}\sum_{j\neq i}(X_j - \mu)^{k-1}] \\
&= \tfrac{1}{n}\mu_k + \tfrac{1}{n}\sum_{i=1}^{n}E[X_i - \mu]\tfrac{1}{n}\sum_{j\neq i}E[(X_j - \mu)^{k-1}]] \\
&\overset{E[X_i-\mu]=0}{=} \tfrac{\mu_k}{n} \to 0, \text{ καθώς } n \to \infty,
\end{aligned}$$

αφού $\mu_k = E[(X - \mu)^k] < \infty$.

(ii) Για την απόδειξη της (ii) σχέσης παραθέτουμε το εξής Λήμμα:

Λήμμα 4.3.1. *Έστω $X_1, \dots, X_n, X$ ανεξάρτητες και ισόνομες τυχαίες μεταβλητές με $E[X] = \mu, Var[X] = \sigma^2, E[|X|^\delta] < \infty$, για κάποιο $\delta \geq 2$. Τότε, $\forall \alpha \in (0, \delta]$ έπεται ότι $E[|\sqrt{n}(\overline{X}_n - \mu)|^\alpha] \to \sigma^\alpha E[|Z|^\alpha]$, καθώς $n \to \infty$, όπου $\overline{X}_n = \sum_{i=1}^{n} X_i$ και $Z \sim N(0, 1)$.*

Η εκτενής απόδειξη του παραπάνω λήμματος παρατίθεται στο [10], σελ. 18.
Έχουμε τώρα από το Λήμμα 4.3.1., για δ=α=1 ότι,

$$|n^{\frac{k}{2}}E[m_{1,n}^k]| \leq n^{\frac{k}{2}}E[|m_{1,n}^k|] = E[|\sqrt{n}m_{1,n}|^k]$$
$$= E[|\sqrt{n}(\overline{X}_n - \mu)|^k] \to \sigma^k E[|Z|^k].$$

Συνεπώς,
$E[|X|^k] < \infty \Rightarrow E[|\sqrt{n}(\overline{X}_n - \mu)|^k] \to \sigma^k E[|Z|^k]$.

Οπότε,

$$|\sqrt{n}E[m_{1,n}^k]| \leq \sqrt{n}E[|m_{1,n}|^k] = \sqrt{n}E[|\overline{X}_n - \mu|^k]$$
$$= \sqrt{n}n^{\frac{k}{2}}n^{-\frac{k}{2}}E[|\overline{X}_n - \mu|^k] = n^{\frac{-(k-1)}{2}}E[|\sqrt{n}(\overline{X}_n - \mu)|^k]$$
$$= \frac{E[|\sqrt{n}(\overline{X}_n-\mu)|^k]}{n^{\frac{(k-1)}{2}}} \to 0, \text{ καθώς } n \to \infty,$$

αφού $E[|\sqrt{n}(\overline{X}_n - \mu)|^k] = |n^{\frac{k}{2}}E[m_{1,n}^k]| \to \sigma^k E[|Z|^k] < \infty$.

(iii) Για την απόδειξη της (iii) θα κάνουμε χρήση του παρακάτω Λήμματος:

Λήμμα 4.3.2. *Έστω $X_1, \ldots, X_n, X$ ανεξάρτητες και ισόνομες τυχαίες μεταβλητές με $E[X] = \mu$ και $E[|X|^\nu] < \infty$, για κάποιο $\nu \in \{2, 3, \ldots\}$. Τότε, $\forall j \in \{2, \ldots, \nu\}$ έπεται ότι*

$$E[|m_{j,n}|^{\frac{\nu}{j}}] \leq E[|X - \mu|^\nu],$$

όπου $m_{j,n} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^j$.

Απόδειξη. Λήμματος.

•Για $j = \nu$ έχουμε ότι
$$|m_{j,n}| = |\frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^j| \leq \frac{1}{n}\sum_{i=1}^{n}|(X_i - \mu)|^j \overset{j=\nu}{=} \frac{1}{n}\sum_{i=1}^{n}|(X_i - \mu)|^\nu.$$

Άρα, $E[|m_{j,n}|^{\frac{\nu}{\nu}}] = E[|m_{j,n}|] \leq E[\frac{1}{n}\sum_{i=1}^{n}|(X_i - \mu)|^\nu] = \frac{1}{n}\sum_{i=1}^{n}E[|(X_i - \mu)|^\nu]$
$$= E[|(X - \mu)|^\nu].$$

•Για $j < \nu$ τότε,
αρχικά παραθέτουμε την ανισότητα Holder, η οποία είναι χρήσιμη στη συνέχεια.

Ανισότητα Holder : Για $(x_1, \ldots, x_n), (y_1, \ldots, y_n) \in \mathbb{R}^n$, για συζυγείς $p, q > 1$ δηλαδή $\frac{1}{p} + \frac{1}{q} = 1 (\Rightarrow q = \frac{p}{p-1})$ ισχύει η ανισότητα

$$\sum_{i=1}^{n}|x_i y_i| \leq (\sum_{i=1}^{n}|x_i|^p)^{\frac{1}{p}}(\sum_{i=1}^{n}|y_i|^q)^{\frac{1}{q}}, \tag{4.3.2}$$

ενώ μια δεύτερη έκφρασή της στο χώρο πιθανότητας (Ω, F, P), με τυχαίες μεταβλητές X, Υ στον Ω, είναι η ακόλουθη

$$E[|XY|] \leq (E[|X|^p])^{\frac{1}{p}}(E[|Y|^q])^{\frac{1}{q}}. \tag{4.3.3}$$

Από χρήση της (4.3.2) προκύπτει ότι,

$$\begin{aligned}
|\sum_{i=1}^{n} x_i|^p = |\sum_{i=1}^{n} 1x_i|^p \leq (\sum_{i=1}^{n} |1x_i|)^p &\leq [(\sum_{i=1}^{n} |x_i|^p)^{\frac{1}{p}}(\sum_{i=1}^{n} |1|^q)^{\frac{1}{q}}]^p \\
&= [(\sum_{i=1}^{n} |x_i|^p)^{\frac{p}{p}}][(\sum_{i=1}^{n} |1|)^{\frac{p}{q}}] \\
&= (\sum_{i=1}^{n} |x_i|^p) n^{\frac{p}{q}} \\
&= n^{\frac{p}{\frac{p}{p-1}}}(\sum_{i=1}^{n} |x_i|^p) \\
&= n^{p-1}(\sum_{i=1}^{n} |x_i|^p).
\end{aligned}$$

Εφαρμόζοντας την παραπάνω σχέση με $p = \frac{\nu}{j}$ και $x_i = (X_i - \mu)^j$ παίρνουμε,

$$\begin{aligned}
E[|m_{j,n}|^{\frac{\nu}{j}}] &= E[|\tfrac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^j|^{\frac{\nu}{j}}] = \frac{1}{n^{\frac{\nu}{j}}}E[|\sum_{i=1}^{n}(X_i-\mu)^j|^{\frac{\nu}{j}}] \\
&\leq \frac{1}{n^{\frac{\nu}{j}}}E[(\sum_{i=1}^{n}|(X_i-\mu)|^j)^{\frac{\nu}{j}}] \leq \frac{1}{n^{\frac{\nu}{j}}}E[n^{\frac{\nu}{j}-1}\sum_{i=1}^{n}|(X_i-\mu)|^{j\frac{\nu}{j}}] \\
&= \frac{1}{n^{\frac{\nu}{j}}}n^{\frac{\nu}{j}-1}E[\sum_{i=1}^{n}|(X_i-\mu)|^{\nu}] = \frac{1}{n}E[\sum_{i=1}^{n}|(X_i-\mu)|^{\nu}] \\
&= E[|(X-\mu)|^{\nu}].
\end{aligned}$$

□

Συνέχεια Απόδειξης.
Θέλουμε να δείξουμε ότι $\sqrt{n}E[m_{1,n}^{k-j}m_{j,n}] \to 0, \quad$ όπου $j = 2, \ldots, k-2$ και $k \geq 4$.

Με εφαρμογή της ανισότητας Holder δηλαδή της έκφρασης (4.3.3), για $p = \frac{k}{k-j}$ και άρα $q = \frac{p}{p-1} = \frac{k}{j}$, με $X = |m_{1,n}|^{k-j}, Y = |m_{j,n}|$ προκύπτει

$$\begin{aligned}
\sqrt{n}E[m_{1,n}^{k-j}m_{j,n}] &\leq \sqrt{n}(E[|m_{1,n}|^{k-j\frac{k}{k-j}}])^{\frac{k-j}{k}}(E[|m_{j,n}|^{\frac{k}{j}}])^{\frac{j}{k}} \\
&= \sqrt{n}(E[|m_{1,n}|^k])^{\frac{k-j}{k}}(E[|m_{j,n}|^{\frac{k}{j}}])^{\frac{j}{k}} \\
&= \sqrt{n}(E[n^{\frac{k}{2}}n^{-\frac{k}{2}}|m_{1,n}|^k])^{\frac{k-j}{k}}(E[|m_{j,n}|^{\frac{k}{j}}])^{\frac{j}{k}} \\
&= \sqrt{n}(E[n^{-\frac{k}{2}}|\sqrt{n}m_{1,n}|^k])^{\frac{k-j}{k}}(E[|m_{j,n}|^{\frac{k}{j}}])^{\frac{j}{k}} \\
&= \sqrt{n}(E[n^{-\frac{k}{2}}|\sqrt{n}(\overline{X}_n-\mu)|^k])^{\frac{k-j}{k}}(E[|\tfrac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^j|^{\frac{k}{j}}])^{\frac{j}{k}} \\
&= n^{\frac{1}{2}}n^{-\frac{k}{2}}(E[|\sqrt{n}(\overline{X}_n-\mu)|^k])^{\frac{k-j}{k}}\frac{1}{n^{k\frac{j}{k}}}(E[|\sum_{i=1}^{n}(X_i-\mu)^j|^{\frac{k}{j}}])^{\frac{j}{k}}
\end{aligned}$$

$$= n^{\frac{1}{2}} n^{-\frac{k}{2}} (E[|\sqrt{n}(\overline{X}_n - \mu)|^k])^{\frac{k-j}{k}} n^{-j} n (E[|X-\mu|^k])^{\frac{j}{k}}$$
$$= n^{-\frac{k-1-j}{2}} (E[|\sqrt{n}(\overline{X}_n - \mu)|^k])^{\frac{k-j}{k}} (E[|X-\mu|^k])^{\frac{j}{k}}$$
$$\to 0,$$

διότι από Λήμμα 4.3.1. για δ=α=k ισχύει
$E[|\sqrt{n}(\overline{X}_n - \mu)|^k] \to \sigma^k E[|Z|^k] < \infty$.
Η απόδειξη ολοκληρώθηκε. □

Πρόταση 4.3.2. *Εάν* $E[|X|^{k+1}] < \infty$ *, για κάποιο* $k \in \{2, 3, ...\}$*, τότε καθώς* $n \to \infty$,

$$Cov[\overline{X}_n, M_{k,n}] = \frac{1}{n}(\mu_{k+1} - k\sigma^2\mu_{k-1}) + o(\frac{1}{n}). \quad (4.3.4)$$

Απόδειξη. Θα δείξουμε ότι, καθώς $n \to \infty$,

$$Cov[\overline{X}_n, M_{k,n}] = \tfrac{1}{n}(\mu_{k+1} - k\sigma^2\mu_{k-1}) + o(\tfrac{1}{n})$$
$$\Leftrightarrow nCov[\overline{X}_n - \mu, M_{k,n} - \mu_k] = \mu_{k+1} - k\sigma^2\mu_{k-1} + no(\tfrac{1}{n})$$
$$\Leftrightarrow Cov[\sqrt{n}(\overline{X}_n - \mu), \sqrt{n}(M_{k,n} - \mu_k)] = \mu_{k+1} - k\sigma^2\mu_{k-1} + o(\tfrac{n}{n})$$
$$\Leftrightarrow Cov[\sqrt{n}(\overline{X}_n - \mu), \sqrt{n}(M_{k,n} - \mu_k)] - \mu_{k+1} + k\sigma^2\mu_{k-1} = o(1)$$
$$\overset{\Lambda\eta\mu\mu\alpha 1.6.1.}{\Leftrightarrow} Cov[\sqrt{n}(\overline{X}_n - \mu), \sqrt{n}(M_{k,n} - \mu_k)] - \mu_{k+1} + k\sigma^2\mu_{k-1} \to 0$$
$$\Leftrightarrow Cov[\sqrt{n}(\overline{X}_n - \mu), \sqrt{n}(M_{k,n} - \mu_k)] \to \mu_{k+1} - k\sigma^2\mu_{k-1}$$
$$\Leftrightarrow E[n(\overline{X}_n - \mu)(M_{k,n} - \mu_k)] - E[\sqrt{n}(\overline{X}_n - \mu)]E[\sqrt{n}(M_{k,n} - \mu_k]$$
$$\to \mu_{k+1} - k\sigma^2\mu_{k-1}$$
$$\overset{E[\sqrt{n}(\overline{X}_n - \mu)]=0}{\Leftrightarrow} E[n(\overline{X}_n - \mu)(M_{k,n} - \mu_k)] \to \mu_{k+1} - k\sigma^2\mu_{k-1}$$
$$\Leftrightarrow nE[(\overline{X}_n - \mu)(M_{k,n} - \mu_k)] = nE[m_{1,n}(M_{k,n} - \mu_k)] \to \mu_{k+1} - k\sigma^2\mu_{k-1}.$$

•Για $k = 2$,
από (4.1.6) με αντικατάσταση παίρνουμε $M_{2,n} = m_{2,n} - m_{1,n}^2$.
Άρα, $nE[m_{1,n}(M_{2,n} - \mu_2)] = nE[(\overline{X}_n - \mu)(M_{2,n} - \mu_2)]$

$$= nE[(\overline{X}_n - \mu)(m_{2,n} - m_{1,n}^2 - \mu_2)]$$
$$= nE[(\overline{X}_n - \mu)(m_{2,n} - \mu_2 - (\overline{X}_n - \mu)^2)]$$
$$= nE[(\overline{X}_n - \mu)(m_{2,n} - \mu_2) - (\overline{X}_n - \mu)^3)]$$
$$= nE[(\overline{X}_n - \mu)(m_{2,n} - \mu_2)] - nE[(\overline{X}_n - \mu)^3)]$$
$$= \mu_3 - nE[(\overline{X}_n - \mu)^3)],$$

γιατί $nE[(\overline{X}_n - \mu)(m_{2,n} - \mu_2)] = nE[(\overline{X}_n - \mu)m_{2,n}] - n\mu_2 E[\overline{X}_n - \mu]$

$$\overset{E[\overline{X}_n - \mu]=0}{=} nE[(\overline{X}_n - \mu)m_{2,n}]$$
$$= nE[\tfrac{1}{n}\sum_{i=1}^{n}(X_i - \mu)\tfrac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2]$$

$$= \tfrac{1}{n}E[\sum_{i=1}^{n}(X_i-\mu)^3 + \sum_{i=1}^{n}\sum_{j\neq i}(X_i-\mu)(X_j-\mu)^2]$$
$$= \tfrac{1}{n}\sum_{i=1}^{n} E[(X_i-\mu)^3] + E[\sum_{i=1}^{n}(X_i-\mu)\sum_{j\neq i}(X_j-\mu)^2]$$
$$= \tfrac{1}{n}n\mu_3 + \sum_{i=1}^{n} E[X_i-\mu]\sum_{j\neq i}E[(X_j-\mu)^2]$$
$$\overset{E[X_i-\mu]=0}{=} \mu_3.$$

Οπότε $nE[m_{1,n}(M_{2,n}-\mu_2)] = \mu_3 - nE[(\overline{X}_n-\mu)^3)]$
και από Λήμμα 4.3.1. για δ=α=3 έχουμε ότι

$$|nE[(\overline{X}_n-\mu)^3)]| = |nn^{\frac{1}{2}}n^{-\frac{1}{2}}E[(\overline{X}_n-\mu)^3)]| = |n^{\frac{3}{2}}n^{-\frac{1}{2}}E[(\overline{X}_n-\mu)^3)]|$$
$$\leq n^{-\frac{1}{2}}E[|\sqrt{n}(\overline{X}_n-\mu)|^3]$$
$$\to n^{-\frac{1}{2}}\sigma^3E[|Z|^3] \to 0, \text{ καθώς } n\to\infty,$$

αφού $\sigma^3E[|Z|^3] < \infty$.

Επομένως, καταλήξαμε ότι καθώς $n\to\infty$,
$nE[m_{1,n}(M_{2,n}-\mu_2)] \to \mu_3 \overset{\mu_1=0}{=} \mu_3 - 3\sigma^2\mu_1$.

•Για $k=3$,
από (4.1.6) με αντικατάσταση παίρνουμε

$$M_{3,n} = (-1)^{3-1}(3-1)m_{1,n}^3 + \sum_{j=2}^{3-1}(-1)^{3-j}m_{1,n}^{3-j}m_{j,n} + m_{3,n}$$
$$= 2m_{1,n}^3 + \tbinom{3}{2}(-1)^{3-2}m_{1,n}^{3-2}m_{2,n} + m_{3,n}$$
$$= 2m_{1,n}^3 - 3m_{1,n}m_{2,n} + m_{3,n}$$

Άρα,

$$nE[m_{1,n}(M_{2,n}-\mu_2)] = nE[m_{1,n}(2m_{1,n}^3 - 3m_{1,n}m_{2,n} + m_{3,n} - \mu_3)]$$
$$= 2nE[m_{1,n}^4] - 3nE[m_{1,n}^2m_{2,n}] + nE[m_{1,n}(m_{3,n}-\mu_3)]$$
$$= nE[(\overline{X}_n-\mu)(m_{3,n}-\mu_3)] + 2nE[(\overline{X}_n-\mu)^4]$$
$$-3nE[(\overline{X}_n-\mu)^2m_{2,n}].$$

Και $nE[(\overline{X}_n-\mu)(m_{3,n}-\mu_3)] = nE[(\overline{X}_n-\mu)m_{3,n}] - n\mu_3E[\overline{X}_n-\mu]$

$$\overset{E[\overline{X}_n-\mu]=0}{=} nE[(\overline{X}_n-\mu)m_{3,n}]$$
$$= nE[\tfrac{1}{n}\sum_{i=1}^{n}(X_i-\mu)\tfrac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^3]$$
$$= \tfrac{1}{n}E[\sum_{i=1}^{n}(X_i-\mu)^4 + \sum_{i=1}^{n}\sum_{j\neq i}(X_i-\mu)(X_j-\mu)^3]$$
$$= \tfrac{1}{n}\sum_{i=1}^{n} E[(X_i-\mu)^4] + E[\sum_{i=1}^{n}(X_i-\mu)\sum_{j\neq i}(X_j-\mu)^3]$$
$$= \tfrac{1}{n}n\mu_4 + \sum_{i=1}^{n} E[X_i-\mu]\sum_{j\neq i}E[(X_j-\mu)^3]$$

$$\overset{E[X_i-\mu]=0}{=} \mu_4.$$

Επιπλέον, από Λήμμα 4.3.1. για δ=α=4, καθώς $n \to \infty$,
$|2nE[(\overline{X}_n - \mu)^4]| = 2||n^2 n^{-1}E[(\overline{X}_n - \mu)^4] \leq n^{-1}E[|\sqrt{n}(\overline{X}_n - \mu)^4|] \to 0$,
επειδή $n^{-1}E[|\sqrt{n}(\overline{X}_n - \mu)^4|] \to \sigma^4 E[|Z|^4] < \infty$.
Τέλος,

$$\begin{aligned}
-3nE[m_{2,n}(\overline{X}_n - \mu)^2] &= -3nE[\tfrac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2(\overline{X}_n - \mu)^2] \\
&= -3nE[\tfrac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2(\tfrac{1}{n}\sum_{i=1}^{n}(X_i - \mu))^2] \\
&= -3\tfrac{1}{n^2}E[(X_1-\mu)^2(\sum_{i=1}^{n}(X_i-\mu))^2+...+(X_n-\mu)^2(\sum_{i=1}^{n}(X_i-\mu))^2] \\
&= -3\tfrac{1}{n^2}E[(X_1 - \mu)^2(\sum_{i=1}^{n}(X_i - \mu)^2 + 2\sum_{i=1}^{n}\sum_{j\neq i}(X_i - \mu)(X_j - \mu) \\
&\quad + ... + (X_n - \mu)^2(\sum_{i=1}^{n}(X_i - \mu)^2 + 2\sum_{i=1}^{n}\sum_{j\neq i}(X_i - \mu)(X_j - \mu))] \\
&\overset{E[X-\mu]=0}{=} -3\tfrac{1}{n^2}n[E[(X_1 - \mu)^2\sum_{i=1}^{n}(X_i - \mu)^2] \\
&\qquad + ... + E[(X_n - \mu)^2\sum_{i=1}^{n}(X_i - \mu)^2]] \\
&= -3\tfrac{1}{n^2}[E[(X_1 - \mu)^4 + (X_1 - \mu)^2\sum_{i\neq 1}(X_i - \mu)^2] \\
&\quad + ... + E[(X_n - \mu)^4 + (X_n - \mu)^2\sum_{i\neq n}(X_i - \mu)^2]] \\
&= -3\tfrac{1}{n^2}n\mu_4 - 3\tfrac{1}{n^2}E[\sum_{i=1}^{n}(X_i - \mu)^2\sum_{j\neq i}(X_j - \mu)^2] \\
&= -3\tfrac{1}{n}\mu_4 - 3\tfrac{1}{n^2}n\mu_2(n-1)\mu_2 \\
&= -3\tfrac{\mu_4+(n-1)\mu_2^2}{n} \\
&\to -3\mu_2^2, \quad \text{καθώς } n \to \infty.
\end{aligned}$$

Επομένως,

$$nE[m_{1,n}(M_{2,n} - \mu_2)] \to \mu_4 - 3\mu_2^2 \overset{\mu_2=\sigma^2}{=} \mu_{3+1} - 3\sigma^2\mu_{3-1}.$$

•Για $k \geq 4$,
από (4.1.6) με αντικατάσταση παίρνουμε

$$M_{k,n} - \mu_k = (m_{k,n} - \mu_k) + (-1)^{k-1}(k-1)m_{1,n}^k + \sum_{j=2}^{k-1}(-1)^{k-j}m_{1,n}^{k-j}m_{j,n}.$$

Άρα, άμεσα βλέπουμε ότι για να δείξουμε ότι

$nE[m_{1,n}(M_{k,n}-\mu_k)] \to \mu_k - k\sigma^2\mu_2^2$, αρκεί να αποδείξουμε τις κάτωθι σχέσεις, καθώς $n \to \infty$:

(i) $nE[m_{1,n}(m_{k,n}-\mu_k)] = \mu_{k+1}$,

(ii) $nE[m_{1,n}^{k+1}] \to 0$,

(iii) $nE[m_{1,n}^{k+1-j}m_{j,n}] \to 0, \quad$ για $j \in \{2,3,\dots,k-2\}$,

(iv) $nE[m_{1,n}^2 m_{k-1,n}] \to \sigma^2\mu_{k-1}$.

Πράγματι, βρίσκουμε

(i) $$\begin{aligned} nE[m_{1,n}(m_{k,n}-\mu_k)] &= nE[(\overline{X}_n-\mu)(m_{k,n}-\mu_k)] \\ &= nE[(\overline{X}_n-\mu)m_{k,n}] - n\mu_k E[(\overline{X}_n-\mu)] \\ &\overset{E[\overline{X}_n-\mu]=0}{=} nE[(\overline{X}_n-\mu)m_{k,n}] \\ &= nE[\tfrac{1}{n}\sum_{i=1}^n (X_i-\mu)\tfrac{1}{n}\sum_{i=1}^n (X_i-\mu)^k] \\ &= \tfrac{1}{n}E[\sum_{i=1}^n (X_i-\mu)^{k+1} + \sum_{i=1}^n\sum_{j\neq i}(X_i-\mu)(X_j-\mu)^k] \\ &= \tfrac{1}{n}E[\sum_{i=1}^n (X_i-\mu)^{k+1}] + \sum_{i=1}^n E[(X_i-\mu)]\sum_{j\neq i}E[(X_j-\mu)^k] \\ &\overset{E[\overline{X}_n-\mu]=0}{=} \tfrac{1}{n}E[\sum_{i=1}^n (X_i-\mu)^{k+1}] \\ &= \mu_{k+1}. \end{aligned}$$

(ii) $$\begin{aligned} |nE[m_{1,n}^{k+1}]| \le nE[|m_{1,n}|^{k+1}] &= nn^{\frac{k+1}{2}}n^{-\frac{k+1}{2}}E[|\overline{X}_n-\mu|^{k+1}] \\ &= n^{-\frac{k-1}{2}}E[|\sqrt{n}(\overline{X}_n-\mu)|^{k+1}] \to 0, \quad \text{καθώς } n\to\infty, \end{aligned}$$
δεδομένου ότι από Λήμμα 4.3.1. για δ=α=k+1 παίρνουμε
$E[|\sqrt{n}(\overline{X}_n-\mu)|^{k+1}] \le \sigma^{k+1}E[|Z|^{k+1}] < \infty$.

(iv) $$\begin{aligned} nE[m_{1,n}^2 m_{k-1,n}] &= nE[(\tfrac{1}{n}\sum_{i=1}^n (X_i-\mu))^2\tfrac{1}{n}\sum_{i=1}^n (X_i-\mu)^{k-1}] \\ &= \tfrac{1}{n^2}E[(\sum_{i=1}^n (X_i-\mu)^2 - 2\sum_{i=1}^n\sum_{j\neq i}(X_i-\mu)(X_j-\mu))\sum_{i=1}^n (X_i-\mu)^{k-1}] \\ &= \tfrac{1}{n^2}E[\sum_{i=1}^n (X_i-\mu)^2\sum_{i=1}^n (X_i-\mu)^{k-1}] \\ &= \tfrac{1}{n^2}E[\sum_{i=1}^n (X_i-\mu)^{k+1} + \sum_{i=1}^n\sum_{j\neq i}(X_i-\mu)^2(X_j-\mu)^{k-1}] \end{aligned}$$

$$= \tfrac{1}{n^2}\big(\sum_{i=1}^{n} E[(X_i-\mu)^{k+1}] + \sum_{i=1}^{n} E[(X_i-\mu)^2]\sum_{j\neq i} E[(X_j-\mu)^{k-1}]\big)$$
$$= \tfrac{1}{n^2}(n\mu_{k+1} + n\sigma^2(n-1)\mu_{k-1}) \to \sigma^2\mu_{k-1}, \text{ καθώς } n\to\infty.$$

(iii) Για την παρακάτω απόδειξη θα χρησιμοποιήσουμε την ανισότητα Holder, δηλαδή τη σχέση (4.4.3) για $p = \frac{k+1}{k+1-j} (\to \frac{1}{q} = \frac{p-1}{p} = \frac{\frac{k+1-k-1+j}{k+1-j}}{\frac{k+1}{k+1-j}} = \frac{j}{k+1})$. Έχουμε,

$$|nE[m_{1,n}^{k+1-j} m_{j,n}]| \le nE[|m_{1,n}|^{k+1-j}|m_{j,n}|]$$
$$\le n(E[|m_{1,n}|^{k+1-j\frac{k+1}{k+1-j}}])^{\frac{k+1-j}{k+1}}(E[|m_{j,n}|^{\frac{k+1}{j}}])^{\frac{j}{k+1}}$$
$$= n(E[|m_{1,n}|^{k+1}])^{\frac{k+1-j}{k+1}}(E[|m_{j,n}|^{\frac{k+1}{j}}])^{\frac{j}{k+1}}$$

και με εφαρμογή του Λήμματος 4.3.2. για $\frac{\nu}{j} = \frac{k+1}{j}$ παίρνουμε

$$\le n(E[|m_{1,n}|^{k+1}])^{\frac{k+1-j}{k+1}}(E[|X-\mu|^{k+1}])^{\frac{j}{k+1}}$$
$$= n(E[n^{\frac{k+1}{2}}n^{-\frac{k+1}{2}}|\overline{X}_n-\mu|^{k+1}])^{\frac{k+1-j}{k+1}}(E[|X-\mu|^{k+1}])^{\frac{j}{k+1}}$$
$$= n^{\frac{-(k-1-j)}{2}}(E[|\sqrt{n}(\overline{X}_n-\mu)|^{k+1}])^{\frac{k+1-j}{k+1}}(E[|X-\mu|^{k+1}])^{\frac{j}{k+1}},$$
$$\to 0, \qquad \text{καθώς } n\to\infty,$$

αφού από Λήμμα 4.3.1. για α=δ=k+1,
$E[|\sqrt{n}(\overline{X}_n-\mu)|^{k+1}] \le \sigma^{k+1}E[|Z|^{k+1}] < \infty, \mu_{k+1} < \infty$ και $n^{\frac{-(k-1-j)}{2}} \to 0.$

□

Πρόταση 4.3.3. *Εάν $E[|X|^{k+r}] < \infty$, για $r,k \in \{2,3,...\}$, τότε καθώς $n\to\infty$,*

$$Cov[M_{r,n}, M_{k,n}] = \frac{1}{n}v_{rk} + o(\frac{1}{n}), \tag{4.3.5}$$

ενώ αν $E[|X|^{2k}] < \infty$, τότε καθώς καθώς $n\to\infty$,

$$Var[\sqrt{n}(M_{k,n}-\mu_k)] \to v_k^2. \tag{4.3.6}$$

Απόδειξη. Θα δείξουμε ότι ,καθώς $n\to\infty$,

$$Cov[M_{r,n}, M_{k,n}] = \tfrac{1}{n}v_{rk} + o(\tfrac{1}{n})$$
$$\Leftrightarrow nCov[M_{r,n}, M_{k,n}-\mu_k] = v_{rk} + no(\tfrac{1}{n})$$
$$\Leftrightarrow Cov[\sqrt{n}(M_{r,n}-\mu_r), \sqrt{n}(M_{k,n}-\mu_k)] = v_{rk} + o(\tfrac{n}{n})$$
$$\Leftrightarrow Cov[\sqrt{n}(M_{r,n}-\mu_r), \sqrt{n}(M_{k,n}-\mu_k)] - v_{rk} = o(1)$$
$$\overset{\Lambda\eta\mu\mu\alpha 1.6.1.}{\Leftrightarrow} Cov[\sqrt{n}(M_{r,n}-\mu_r), \sqrt{n}(M_{k,n}-\mu_k)] - v_{rk} \to 0$$
$$\Leftrightarrow Cov[\sqrt{n}(\overline{X}_n-\mu), \sqrt{n}(M_{k,n}-\mu_k)] \to v_{rk}$$
$$\Leftrightarrow E[n(M_{r,n}-\mu_r)(M_{k,n}-\mu_k)] - E[\sqrt{n}(M_{r,n}-\mu_r)]E[\sqrt{n}(M_{k,n}-\mu_k)]$$
$$\to v_{rk}$$

$\Leftrightarrow nE[(M_{r,n} - \mu_r)(M_{k,n} - \mu_k)] \to v_{rk}$,
αφού από Πρόταση 4.3.1. αποδείξαμε ότι, καθώς $n \to \infty$, $E[M_{k,n}] = \mu_k + o(\frac{1}{\sqrt{n}})$ ή ισοδύναμα $E[\sqrt{n}(M_{k,n} - \mu_k)] \to 0$.

Άρα, μένει να δείξουμε ότι $nE[(M_{r,n} - \mu_r)(M_{k,n} - \mu_k)] \to v_{rk}$

$$\overset{(4.2.9)}{\Leftrightarrow} nE[(M_{r,n} - \mu_r)(M_{k,n} - \mu_k)] \to \mu_{r+k} - k\mu_{r+1}\mu_{k-1} - r\mu_{r-1}\mu_{k+1} + rk\sigma^2\mu_{r-1}\mu_{k-1} - \mu_k\mu_r.$$

Η παραπάνω απόδειξη χωρίζεται στις εξής υποπεριπτώσεις:

•$r = k = 2$,

•$r = 2, k = 3$,

•$r = k = 3$,

•$r = 2, k \geq 4$,

•$r = 3, k \geq 4$,

•$4 \leq r \leq k$.

Οι υποπεριπτώσεις αυτές αποδεικνύονται με παρόμοιο τρόπο. Ενδεικτικά, για την τελευταία, δηλαδή για $4 \leq r \leq k$ εργαζόμαστε ως εξής:
Από (4.1.6) παίρνουμε

$$M_{k,n} - \mu_k = (m_k - \mu_k) + (-1)^{k-1}(k-1)m_1^k + \sum_{j=2}^{k-1} \tbinom{k}{j}(-1)^{k-j} m_j m_1^{k-j}.$$

$$M_{r,n} - \mu_r = (m_r \mu_r) + (-1)^{r-1}(r-1)m_1^r + \sum_{j=2}^{r-1} \tbinom{r}{j}(-1)^{r-j} m_j m_1^{r-j}.$$

Οπότε, άμεσα με αντικατάσταση προκύπτει ότι για να ισχύει $nE[(M_{r,n} - \mu_r)(M_{k,n} - \mu_k)] \to v_{rk}$ αρκεί να ισχύουν οι παρακάτω:

(i) $nE[(m_{r,n} - \mu_r)(m_{k,n} - \mu_k)] = \mu_{r+k} - \mu_r\mu_k$,

(ii) $nE[m_{1,n} m_{k-1,n}(m_{r,n} - \mu_r)] \to \mu_{r+1}\mu_{k-1}$,

(iii) $nE[m_{1,n} m_{r-1,n}(m_{k,n} - \mu_k)] \to \mu_{k+1}\mu_{r-1}$,

(iv) $nE[m_{1,n}^2 m_{r-1,n} m_{k-1,n}] \to \sigma^2\mu_{r-1}\mu_{k+1}$,

(v) $nE[m_{1,n}^{k-j_2} m_{j_2,n}(m_{r,n} - \mu_r)] \to 0, \quad j_2 \in \{2, 3, \dots, k-2\}$,

(vi) $nE[m_{1,n}^{k}(m_{r,n}-\mu_r)] \to 0,$

(vii) $nE[m_{1,n}^{k+1-j_2} m_{j_2,n} m_{r-1,n}] \to 0, \quad j_2 \in \{2,3,\dots,k-2\},$

(viii) $nE[m_{1,n}^{k+1} m_{r-1,n}] \to 0,$

(ix) $nE[m_{1,n}^{r-j_1} m_{j_1,n}(m_{k,n}-\mu_k)] \to 0, \quad j_1 \in \{2,3,\dots,r-2\},$

(x) $nE[m_{1,n}^{r+1-j_1} m_{j_1,n} m_{k-1,n}] \to 0, \quad j_1 \in \{2,3,\dots,r-2\},$

(xi) $nE[m_{1,n}^{r+k-j_1-j_2} m_{j_1,n} m_{j_2,n}] \to 0, j_1 = 2,3,\dots,r-2, j_2 = 2,3,\dots,k-2,$

(xii) $nE[m_{1,n}^{r+k-j_1} m_{j_1,n}] \to 0, \quad j_1 = 2,3,\dots,r-2,$

(xiii) $nE[m_{1,n}^{r}(m_{k,n}-\mu_k)] \to 0,$

(xiv) $nE[m_{1,n}^{r+1} m_{k-1,n}] \to 0,$

(xv) $nE[m_{1,n}^{r+k-j_2} m_{j_2,n}] \to 0, \quad j_2 = 2,3,\dots,k-2,$

(xvi) $nE[m_{1,n}^{r+k}] \to 0.$

Οι σχέσεις αυτές αποδεικνύονται με τη χρήση των ίδιων τεχνασμάτων και Λημμάτων που αναφέρθηκαν στις αποδείξεις των προτάσεων 4.3.1. και 4.3.2. ενώ λεπτομερής απόδειξή τους δύναται να βρεθεί στο [16].

□

4.4 Ανεξαρτησία δειγματικών κεντρικών ροπών από το δειγματικό μέσο

Ορισμός 4.4.1. Λέμε ότι ο δειγματικός μέσος $\overline{X}_n$ είναι ασυμπτωτικά ανεξάρτητος της δειγματικής κεντρικής ροπής $M_{k,n}$, για $k \in \{2,3,\dots\}$, εάν υπάρχουν τυχαίες μεταβλητές W_1 και W_k, οι οποίες είναι ανεξάρτητες, τέτοιες ώστε, καθώς $n \to \infty$, να αληθεύει ότι

$$\sqrt{n}\begin{bmatrix} \overline{X}_n - \mu \\ M_{k,n} - \mu_k \end{bmatrix} \xrightarrow{D} \begin{bmatrix} W_1 \\ W_k \end{bmatrix}.$$

Ορισμός 4.4.2. Λέμε ότι ο δειγματικός μέσος $\overline{X}_n$ είναι ασυμπτωτικά ανεξάρτητος του διανύσματος δειγματικών κεντρικών ροπών $(M_{2,n}, M_{3,n}, \dots, M_{k,n})$, για $k \in \{2, 3, ...\}$, εάν υπάρχουν τυχαίες μεταβλητές $W_1, W_2 \ldots, W_k$ τέτοιες ώστε οι W_1 και $(W_2 \ldots, W_k)'$ να είναι ανεξάρτητες και καθώς $n \to \infty$, να αληθεύει ότι

$$\sqrt{n} \begin{bmatrix} \overline{X}_n - \mu \\ M_{2,n} - \mu_k \\ M_{3,n} - \mu_k \\ \ldots \\ M_{k,n} - \mu_k \end{bmatrix} \xrightarrow{D} \begin{bmatrix} W_1 \\ W_2 \\ W_3 \\ \ldots \\ W_k \end{bmatrix}.$$

Ορισμός 4.4.3. Λέμε ότι ο δειγματικός μέσος $\overline{X}_n$ και η δειγματική κεντρική ροπή $M_{k,n}$ είναι ασυμπτωτικά ασυσχέτιστες ποσότητες εάν ισχύει

$$Cov[\sqrt{n}\overline{X}_n, \sqrt{n}M_{k,n}] \to 0, \quad \text{καθώς } n \to \infty.$$

Πρόταση 4.4.1. *Έστω $E[X^{2k}] < \infty$, για κάποιο $k \in \{2, 3, ...\}$. Οι ποσότητες $\overline{X}_n, M_{k,n}$ είναι ασυμπτωτικά ανεξάρτητες αν και μόνο αν είναι ασυμπτωτικά ασυσχέτιστες.*

Απόδειξη. Για το ευθύ, υποθέτουμε ότι οι $\overline{X}_n, M_{k,n}$ είναι ασυμπτωτικά ανεξάρτητες. Τότε, υπάρχουν ανεξάρτητες τυχαίες μεταβλητές W_1 και W_k τέτοιες ώστε, καθώς $n \to \infty$, να αληθεύει ότι

$$\sqrt{n} \begin{bmatrix} \overline{X}_n - \mu \\ M_{k,n} - \mu_k \end{bmatrix} \xrightarrow{D} \begin{bmatrix} W_1 \\ W_k \end{bmatrix}.$$

Παραπάνω, βρήκαμε ότι η οριακή κατανομή του ζεύγους $(M_{1,n}, M_{k,n})'$ δίνεται :

$$\sqrt{n} \begin{bmatrix} \overline{X}_n - \mu \\ M_{k,n} - \mu_k \end{bmatrix} = \sqrt{n} \begin{bmatrix} \overline{X}_n - \mu \\ M_{k,n} - \mu_k \end{bmatrix} \xrightarrow{D} N(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma^2 & \mu_{k+1} - k\sigma^2\mu_{k-1} \\ \mu_{k+1} - k\sigma^2\mu_{k-1} & v_\kappa^2 \end{bmatrix}).$$

Άμεσα βλέπουμε ότι για να καλύπτεται η υπόθεση της ανεξαρτησίας των W_1, W_k, πρέπει $Cov[\sqrt{n}(\overline{X}_n - \mu), \sqrt{n}(M_{k,n} - \mu_k)] \to 0$

$\Rightarrow Cov[\sqrt{n}\overline{X}_n, \sqrt{n}M_{k,n}] \to 0 \overset{O\rho.4.4.3.}{\Rightarrow} \overline{X}_n, M_{k,n}$ είναι ασυμπτωτικά ασυσχέτιστες.

Για το αντίστροφο, υποθέτουμε ότι οι $\overline{X}_n, M_{k,n}$ είναι ασυμπτωτικά ασυσχέτιστες. Τότε, καθώς $n \to \infty$, αληθεύει ότι
$Cov[\sqrt{n}\overline{X}_n, \sqrt{n}M_{k,n}] \to 0 \Rightarrow Cov[\sqrt{n}(\overline{X}_n - \mu), \sqrt{n}(M_{k,n} - \mu_k)] \to 0.$
Επειδή από Πρόταση 4.3.2 ισχύει
$Cov[\sqrt{n}(\overline{X}_n - \mu), \sqrt{n}(M_{k,n} - \mu_k)] \to \mu_{k+1} - k\sigma^2\mu_{k-1}$, έπεται ότι
$\mu_{k+1} - k\sigma^2\mu_{k-1} = 0$. Άρα, γενικά αποδείξαμε παραπάνω πως ισχύει

$$\sqrt{n} \begin{bmatrix} \overline{X}_n - \mu \\ M_{k,n} - \mu_k \end{bmatrix} \xrightarrow{D} \begin{bmatrix} W_1 \\ W_k \end{bmatrix} \sim N(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma^2 & \mu_{k+1} - k\sigma^2\mu_{k-1} \\ \mu_{k+1} - k\sigma^2\mu_{k-1} & v_\kappa^2 \end{bmatrix})$$

$$\overset{\mu_{k+1} - k\sigma^2\mu_{k-1} = 0}{\Rightarrow} \sqrt{n} \begin{bmatrix} \overline{X}_n - \mu \\ M_{k,n} - \mu_k \end{bmatrix} \xrightarrow{D} \begin{bmatrix} W_1 \\ W_k \end{bmatrix} \sim N(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma^2 & 0 \\ 0 & v_\kappa^2 \end{bmatrix}).$$

Επομένως, υπάρχουν ανεξάρτητες τυχαίες μεταβλητές W_1 και W_k τέτοιες ώστε, καθώς $n \to \infty$, να ισχύει

$$\sqrt{n}\begin{bmatrix}\overline{X}_n-\mu\\ M_{k,n}-\mu_k\end{bmatrix}\xrightarrow{D}\begin{bmatrix}W_1\\ W_k\end{bmatrix}.$$

Οπότε, από Ορισμό 4.4.1. οι $\overline{X}_n, M_{k,n}$ είναι ασυμπτωτικά ανεξάρτητες. □

Πρόταση 4.4.2. *Έστω $E[X^{2k}]<\infty$, για κάποιο $k\in\{2,3,...\}$. Οι ποσότητες $\overline{X}_n$ και $(M_{2,n},\dots,M_{k,n})'$ είναι ασυμπτωτικά ανεξάρτητες αν και μόνο αν οι ποσότητες $\overline{X}_n, M_{r,n}$ είναι ασυμπτωτικά ασυσχέτιστες, για κάθε $r\in\{2,3,\dots,k\}$.*

Απόδειξη. Για το ευθύ, υποθέτουμε ότι οι $\overline{X}_n$ και $(M_{2,n},\dots,M_{k,n})'$ είναι ασυμπτωτικά ανεξάρτητες. Τότε, υπάρχουν ανεξάρτητες τυχαίες μεταβλητές $W_1, W_2\dots,W_k$ τέτοιες ώστε οι W_1 και $(W_2\dots,W_k)'$ να είναι ανεξάρτητες και καθώς $n\to\infty$, να αληθεύει ότι

$$\sqrt{n}\begin{bmatrix}\overline{X}_n-\mu\\ M_{2,n}-\mu_k\\ M_{3,n}-\mu_k\\ \dots\\ M_{k,n}-\mu_k\end{bmatrix}\xrightarrow{D}\begin{bmatrix}W_1\\ W_2\\ W_3\\ \dots\\ W_k\end{bmatrix}.$$

Παραπάνω, βρήκαμε ότι η οριακή κατανομή του διανύσματος δειγματικών κεντρικών ροπών $(M_{1,n},\dots,M_{k,n})'$ δίνεται :

$$\sqrt{n}\begin{bmatrix}\overline{X}_n-\mu\\ M_{2,n}-\sigma^2\\ \dots\\ M_{k,n}-\mu_k\end{bmatrix}\xrightarrow{D}N(\begin{bmatrix}0\\0\\ \dots\\0\end{bmatrix},\begin{bmatrix}\sigma^2 & \mu_3 & \mu_4-3\sigma^2 & \dots & \mu_{k+1}-k\sigma^2\mu_{k-1}\\ \mu_3 & v_2^2 & v_{23} & \dots & v_{2k}\\ \mu_4-3\sigma^2 & v_{23} & v_3^2 & \dots & v_{3k}\\ \dots & \dots & \dots & \dots & \dots\\ \mu_{k+1}-k\sigma^2\mu_{k-1} & v_{2k} & v_{3k} & \dots & v_k^2\end{bmatrix}).$$

Άμεσα βλέπουμε ότι για να είναι οι W_1 και$(W_2\dots,W_k)'$ ανεξάρτητες, πρέπει

$$Cov[\sqrt{n}(\overline{X}_n-\mu),\sqrt{n}(M_{r,n}-\mu_r)]\to 0,\quad \text{για } r\in\{2,3,\dots,k\}$$
$$\Rightarrow Cov[\sqrt{n}\overline{X}_n,\sqrt{n}M_{r,n}]\to 0,\quad \text{για } r\in\{2,3,\dots,k\}$$
$$\overset{Op.4.4.3.}{\Rightarrow}\ \overline{X}_n, M_{r,n} \text{ είναι ασυμπτωτικά ασυσχέτιστες.}$$

Αντίστροφα, αν $\overline{X}_n, M_{r,n}$ είναι ασυμπτωτικά ασυσχέτιστες, για $r\in\{2,3,\dots,k\}$, τότε εξ' ορισμού παίρνουμε για $r\in\{2,3,\dots,k\}$
$Cov[\sqrt{n}\overline{X}_n,\sqrt{n}M_{r,n}]\to 0\Rightarrow Cov[\sqrt{n}(\overline{X}_n-\mu),\sqrt{n}(M_{r,n}-\mu_r)]\to 0.$
Και λόγω της Πρότασης 4.3.2. προκύπτει $\mu_{k+1}-k\sigma^2\mu_{k-1}=0$. Άρα, από την τελευταία σχέση και την οριακή κατανομή του διανύσματος δειγματικων κεντρικών ροπών καταλήγουμε ότι

$$\sqrt{n}\begin{bmatrix}\overline{X}_n-\mu\\ M_{2,n}-\sigma^2\\ \dots\\ M_{k,n}-\mu_k\end{bmatrix}\xrightarrow{D}N(\begin{bmatrix}0\\0\\ \dots\\0\end{bmatrix},\begin{bmatrix}\sigma^2 & 0 & 0 & \dots & 0\\ 0 & v_2^2 & v_{23} & \dots & v_{2k}\\ 0 & v_{23} & v_3^2 & \dots & v_{3k}\\ \dots & \dots & \dots & \dots & \dots\\ 0 & v_{2k} & v_{3k} & \dots & v_k^2\end{bmatrix}).$$

Καταληκτικά, υπάρχουν W_1 και$(W_2\dots,W_k)'$ ανεξάρτητες ώστε

$$\sqrt{n}\begin{bmatrix}\overline{X}_n-\mu\\ M_{2,n}-\mu_k\\ \dots\\ M_{k,n}-\mu_k\end{bmatrix}\xrightarrow{D}\begin{bmatrix}W_1\\ W_2\\ \dots\\ W_k\end{bmatrix}$$

και άρα από Ορισμό 4.4.2. καταλήγουμε στο ζητούμενο.

$\square$

Παρατήρηση 4.4.1. Έστω $E[|X|^{k+1}] < \infty$. Τότε, ακόμα και για τις περιπτώσεις όπου η οριακή κατανομή της ποσότητας $\sqrt{n}(M_{k,n}-\mu_k)$ δεν υπάρχει, η Πρόταση 4.3.2. σύμφωνα με την οποία,
αν $E[|X|^{k+1}] < \infty, \Rightarrow Cov[\sqrt{n}(\overline{X}_n-\mu), \sqrt{n}(M_{k,n}-\mu_k)] \to \mu_{k+1}-k\sigma^2\mu_{k-1}$ ή ισοδύναμα $Cov[\sqrt{n}\overline{X}_n, \sqrt{n}M_{k,n}] \to \mu_{k+1}-k\sigma^2\mu_{k-1}$, μας επιτρέπει να αποφανθούμε αν οι $\overline{X}_n, M_{k,n}$ είναι ασυμπτωτικά ασυσχέτιστες. Δηλαδή σύμφωνα με τον ορισμό 4.3.3. ελέγχουμε κατά πόσο ισχύει ότι $\mu_{k+1} = k\sigma^2\mu_{k-1}$.

Θεώρημα 4.4.1. *Έστω X μια μη εκφυλισμένη τυχαία μεταβλητή με πεπερασμένες ροπές οποιασδήποτε τάξης. Οι $\overline{X}_n, M_{k,n}$ είναι ασυμπτωτικά ανεξάρτητες (ή ισοδύναμα ασυμπτωτικά ασυσχέτιστες) $\forall k \in \{2,3,...\}$, αν και μόνο αν η X ακολουθεί Κανονική Κατανομή.*

Απόδειξη. Έστω $E[X]=\mu, Var[X]=\sigma^2<\infty, \mu_k = E[(X-\mu)^k]$.

Για το ευθύ, υποθέτουμε ότι οι $\overline{X}_n, M_{k,n}$ είναι ασυμπτωτικά ανεξάρτητες. Ισοδύναμα από Πρόταση 4.4.1. οι $\overline{X}_n, M_{k,n}$ είναι ασυμπτωτικά ασυσχέτιστες. Τότε, από ορισμό 4.4.3. προκύπτει ότι
$Cov[\sqrt{n}\overline{X}_n, \sqrt{n}M_{k,n}] \to 0 \Rightarrow Cov[\sqrt{n}(\overline{X}_n-\mu), \sqrt{n}(M_{k,n}-\mu_k)] \to 0$
Άρα, με χρήση της Πρότασης 4.3.2. συμπεραίνουμε ότι
$\mu_{k+1} = k\sigma^2\mu_{k-1}, \forall k \in \{2,3,...\}$. Υπολογίζουμε
$\mu_1 = 1$,
$\mu_2 = \sigma^2$,
$\mu_3 = 2\sigma^2\mu_1 = 0$,
$\mu_4 = 3\sigma^2\mu_2 = 3\sigma^4$,
$\mu_5 = 4\sigma^2\mu_3 = 0$,
$\mu_6 = 5\sigma^2\mu_4 = 15\sigma^6$. Με επαγωγή προκύπτει ότι, $\mu_k = \begin{cases}0, & \text{εάν } k \text{ είναι περιττός,}\\ \sigma^k(k-1)!\,!, & \text{εάν } k \text{ είναι άρτιος.}\end{cases}$
όπου για $r=1,2,...,\quad (k-1)!\,! = (2r-1)!\,! = \frac{(2r)!}{2^r r!} = (k-1)(k-3)\dots 3\cdot 1$.
Όμως, αποδεικνύεται ότι οι ροπές της Κανονικής κατανομής είναι
$\mu_k = \begin{cases}0, & \text{εάν } k \text{ είναι περιττός,}\\ \sigma^k(k-1)!\,!, & \text{εάν } k \text{ είναι άρτιος.}\end{cases}$
Επιπλέον, η Κανονική κατανομή είναι η μοναδική κατανομή με τις συγκεκριμένες

ροπές, διότι έχει πεπερασμένη ροπογεννήτρια σε μια περιοχή του μηδενός[1] (βλ. [11]). Οπότε χαρακτηρίζεται από αυτές. Άρα άμεσα βλέπουμε ότι
$X - \mu \sim N(0, \sigma^2)$, δηλαδή $X \sim N(\mu, \sigma^2)$.

Αντιστρόφως, αν $X \sim N(\mu, \sigma^2)$ τότε οι ροπές δίνονται κατά τα γνωστά ως $\mu_1 = \mu_3 = \mu_5 = ... = 0$ και $\mu_{2r} = \sigma^2(2r-1)!!$, για $r = 1, 2, ...$ Οι σχέσεις αυτές επαληθεύουν ότι $\mu_{k+1} = k\sigma^2\mu_{k-1} \overset{\Pi\rho.4.3.2.}{\Rightarrow} Cov[\sqrt{n}\overline{X}_n, \sqrt{n}M_{k,n}] \to 0$. Και επομένως, οι $\overline{X}_n, M_{k,n}$ είναι ασυμπτωτικά ασυσχέτιστες (ή ισοδύναμα ανεξάρτητες). $\square$

Παρατήρηση 4.4.2. Το σημαντικό πόρισμα του παραπάνω θεωρήματος είναι το ευθύ, καθώς το αντίστροφο είναι ήδη γνωστό χωρίς να εργαστούμε ασυμπτωτικά. Επεξηγηματικά, ο δειγματικός μέσος $\overline{X}$ είναι ανεξάρτητος του διανύσματος $\mathbf{Y} = (X_1 - \overline{X}, ..., X_n - \overline{X})'$ εφόσον το διάνυσμα $(\overline{X}, X_1 - \overline{X}, ..., X_n - \overline{X})'$ ακολουθεί πολυδιάστατη Κανονική κατανομή και $Cov[\overline{X}, X_i - \overline{X}] = 0$. Επομένως, ο δειγματικός μέσος $\overline{X}$ είναι ανεξάρτητος (ή ισοδύναμα ασυσχέτιστος) οποιασδήποτε ακολουθίας της μορφής $\{f_k(\mathbf{Y}), k = 2, 3, ...\}$, με $f_k : \mathbb{R}^n \to \mathbb{R}$ μετρήσιμες Borel συναρτήσεις. Τέλος, αφού εξ' ορισμού $M_{k,n} = \frac{1}{n}\sum\limits_{i=1}^{n}(X_i - \overline{X})^k = f_k(\mathbf{Y})$, άρα οι ποσότητες $\overline{X}$ και $(M_{2,n}, ..., M_{k,n})'$ είναι ανεξάρτητες (άρα $\overline{X}$ και $M_{k,n}$ ασυσχέτιστες) για όλα τα k και n. Οπότε, ισχύει το ίδιο και για την οριακή τους συμπεριφορά.

Παρατήρηση 4.4.3. Από το Θεώρημα 4.4.1. καταλήξαμε ότι μπορούμε να συμπεράνουμε την Κανονικότητα μιας τυχαίας μεταβλητής μέσω της ασυμπτωτικής ανεξαρτησίας του δειγματικού μέσου από όλες τις κεντρικές δειγματικές ροπές. Ισοδύναμα, θα πρέπει $\mu_{k+1} = k\sigma^2\mu_{k-1}$, δηλαδή
$$\mu_k = \begin{cases} 0, & \text{εάν } k \text{ είναι περιττός,} \\ \sigma^k(k-1)!!, & \text{εάν } k \text{ είναι άρτιος.} \end{cases}$$
για όλα τα $k \in 2, 3, ...$, ώστε να αποφανθούμε την κανονικότητα της αντίστοιχης τυχαίας μεταβλητής. Για να είναι ασυμπτωτικά ανεξάρτητες οι ποσότητες $\overline{X}$ και $S^2 = \frac{n}{n-1}M_2$ θα πρέπει $\mu_3 = 2\sigma^2\mu_1 = 0$, εφ' όσον $E[|X|^3] < \infty$. Όμως, $\mu_3 = E[(X-\mu)^3] = 0$ για όλες τις συμμετρικές κατανομές (με πεπερασμένη τρίτη ροπή). Άρα, η ασυμπτωτική ανεξαρτησία των $\overline{X}$ και S^2 δε συνεπάγεται την κανονικότητα. Το άμεσο ερώτημα που ανακύπτει είναι εάν δύναται να βρεθεί κάποιος πεπερασμένος αριθμός από $k \in 2, 3, ...$, ώστε αν αυτά επαληθεύουν τη σχέση

[1]Γενικότερα, μια κατανομή δε μπορεί να χαρακτηριστεί από τις ροπές της παρά μόνο στην περίπτωση που έχει πεπερασμένη ροπογεννήτρια σε μια περιοχή του μηδενός. Για παράδειγμα, η Λογαριθμο-Κανονική κατανομή δε δύναται να χαρακτηριστεί από τις ροπές της, αφού οι τελευταίες ταυτίζονται με τις ροπές μιας οικογένειας κατανομών διάφορων της Λογαριθμο-Καανονικής.

$\mu_{k+1} = k\sigma^2\mu_{k-1}$ να συνεπάγονται την κανονικότητα. Η απάντηση είναι αρνητική, γεγονός που αποδεικνύεται από το παρακάτω θεώρημα.

Θεώρημα 4.4.2. *Για κάθε πεπερασμένο $k \geq 2$, υπάρχουν άπειρες μη εκφυλισμένες τυχαίες μεταβλητές X με πεπερασμένες ροπές οποιασδήποτε τάξης, οι οποίες δεν ακολουθούν Κανονική κατανομή, τέτοιες ώστε οι ποσότητες $\overline{X}_n$ και $(M_{2,n}, \dots, M_{k,n})^T$ να είναι ασυμπτωτικά ανεξάρτητες.*

Απόδειξη. Θα σταθεροποιήσουμε ένα k για το οποίο θέλουμε να φτιάξουμε μια συνάρτηση πυκνότητας πιθανότητας f, η οποία να είναι μη Κανονική, για την οποία οι $\overline{X}_n$ και $(M_{2,n}, \dots, M_{k,n})^T$ να είναι ασυμπτωτικά ανεξάρτητες (ή ισοδύναμα ασυμπτωτικά ασυσχέτιστες) τυχαίες μεταβλητές, καθώς $n \to \infty$. Επομένως, σύμφωνα με τον Ορισμό 4.4.3. και την Πρόταση 4.3.2., για να είναι οι παραπάνω ποσότητες ασυμπτωτικά ανεξάρτητες θα πρέπει να ισχύει
$\mu_{i+1} = i\sigma^2\mu_{i-1}, \forall i \in \{2, 3, \dots, k\}$. Παρατηρούμε ότι για $\sigma^2 = 1$, η παραπάνω σχέση γίνεται $\mu_{i+1} = i\mu_{i-1}, \forall i \in \{2, 3, \dots, k\}$, και αυτή είναι η σχέση που ικανοποιούν οι πρώτες k+1 ροπές της Κανονικής κατανομής N(0,1).

Επομένως, αρκεί να κατασκευάσουμε μια μη κανονική τυχαία μεταβλητή, με ίδιες ροπές, τάξεως μέχρι k+1, με την Κανονική κατανομή N(0,1). Αυτό επιτυγχάνεται με τη χρήση των πολυωνύμων Legendre[2].

Συγκεκριμένα, θεωρώντας τη Hamel βάση $\{1, t, t^2, \dots, t^n\}$, το οποίο αποδεικνύεται άμεσα αφού το παραπάνω σύνολο είναι γραμμικά ανεξάρτητο από ισότητα πολυωνύμων και παράγει το σύνολο των πολυωνύμων και εφαρμόζοντας στη βάση αυτή τη μέθοδο Gram-Schmidt με εσωτερικό γινόμενο
$< p(t), q(t) >= \int_0^1 p(t)q(t)dt$, λαμβάνονται τα ορθογώνια πολυώνυμα
$P_0(t) = 1, P_1(t) = 2t - 1, P_2(t) = 6t^2 - 6t + 1, \dots$. Τα πολυώνυμα είναι ορθογώνια μεταξύ τους αφού κατασκευάστηκαν έτσι ώστε να ισχύει η συνθήκη ορθογωνιότητας $< P_i, P_j >= 0, \forall i \in \{0, \dots, n\}, j = 0, 1, \dots, i-1$.

Έστω η συνάρτηση πυκνότητας πιθανότητας της τυποποιημένης Κανονικής κατανομής $\varphi(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$. Επιπλέον, θεωρούμε το πολυώνυμο Legendre
$P_m(x) = \frac{d^m}{dx^m}(x^m(1-x)^m)$ βαθμού m.

Το πολυώνυμο P_m είναι συνεχής συνάρτηση στο διάστημα $[0, 1]$ οπότε ισχύει $0 < \max_{x\in[0,1]} |P_m(x)| = \alpha_m < \infty$. Ακόμη, επειδή η συνάρτηση φ είναι φθίνουσα για $x \in [0, \infty)$ έπεται ότι $\min_{x\in[0,1]} \varphi(x) = \varphi(1) = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}} = \frac{1}{\sqrt{2\pi e}} > 0$. Επομένως, για $0 < \varepsilon_m < \frac{1}{\alpha_m\sqrt{2\pi e}}$, άμεσα βλέπουμε ότι $\varphi(x) + \varepsilon_m P_m(x) > 0, \forall x \in [0, 1]$. Επιπλέον, $\int_0^1 \varphi(x) + \varepsilon_m P_m(x)dx = \int_0^1 \varphi(x)dx + \varepsilon_m \int_0^1 1 \cdot P_m(x)dx$

[2]Adrien-Marie Legendre (1752-1833): Γάλλος μαθηματικός. Εισήγαγε τα πολυώνυμα Legendre το 1785 προκειμένου να λύσει ένα πρόβλημα δυναμικού. Η γενική τους μορφή δόθηκε από τον μαθηματικο Rodrigues το 1814 και ειναι $P_n(t) = \alpha_n \frac{d^n}{dt^n}[t^n(1-t)^n], n = 0, 1, \dots, \alpha_n \in \mathbb{R}^+$.

$$= \int_0^1 \varphi(x)dx + \varepsilon_m \int_0^1 P_0(x)P_m(x)dx$$
$$= \int_0^1 \varphi(x)dx,$$

λόγω ορθογωνιότητας των πολυωνύμων Legendre .

Συνεπώς, μπορούμε να ορίσουμε την ακολουθία συναρτήσεων πυκνότητας πιθανότητας $\{f_m, m \geq k+2\}$ με

$$f_m = \begin{cases} \varphi(x), & \text{εάν } x < 0 \text{ ή } x > 1, \\ \varphi(x) + \varepsilon_m P_m(x), & \text{εάν } x \in [0,1]. \end{cases}$$

Θεωρούμε μια τυχαία μεταβλητή X, τέτοια ώστε $X \sim f_m, m \geq k+2$.
Τότε, για $j = 1, \ldots, k+1$ προκύπτει

$$\begin{aligned} E[X^j] &= \int_{-\infty}^{\infty} x^j f_n(x)dx \\ &= \int_{-\infty}^{0} x^j \varphi(x)dx + \int_0^1 x^j[\varphi(x) + \varepsilon_m P_m(x)]dx + \int_1^{\infty} x^j \varphi(x)dx \\ &= \int_{-\infty}^{0} x^j \varphi(x)dx + \int_0^1 x^j \varphi(x) + \int_1^{\infty} x^j \varphi(x) + \varepsilon_m \int_0^1 x^j P_m(x)dx \\ &= \int_{-\infty}^{\infty} x^j \varphi(x)dx + \varepsilon_m \int_0^1 x^j P_m(x)dx \\ &= E[Z^j], \quad \text{όπου } Z \sim N(0,1), \end{aligned}$$

αφού εξ' ορισμού το πολυώνυμο Legendre είναι ορθογώνιο ως προς τη βάση $\{1, x, x^2, \ldots, x^{k+1}\}$, στο διάστημα [0,1] δεδομένου ότι $m \geq k+2$, που σημαίνει ότι ισχύει $\int_0^1 x^j P_m(x)dx = 0, \quad j = 0, 1, \ldots, k+1$.

Επομένως, η τυχαία μεταβλητή X έχει πεπερασμένες ροπές οποιασδήποτε τάξης, είναι μη εκφυλισμένη και μη κανονική ($X \sim f_m$) και $E[X^j] = E[Z^j]$, $j = 0, 1, \ldots, k+1$. Επιπλέον, από τον Ορισμό 4.4.2. και το Θεώρημα 4.2.3. προκύπτει ότι οι ποσότητες $\overline{X}$ και $(M_{2,n}, \ldots, M_{k,n})$ είναι ασυμπτωτικά ανεξάρτητες. □

Παρατήρηση 4.4.4. Γενικότερα, σύμφωνα με το Θεώρημα Geary[3] εάν ο δειγματικός μέσος $\overline{X}$ και η δειγματική διασπορά S^2 είναι ανεξάρτητες ποσότητες για κάποιο $n_0 \geq 2$, τότε η τυχαία μεταβλητή Χ ακολουθεί Κανονική κατανομή. Το συγκεκριμένο θεώρημα αναπτύσσεται ενδελεχώς στο [9]. Στα Θεωρήματα 4.4.1. και 4.4.2 που αναφέραμε παραπάνω, αποδεικνύεται το ασυμπτωτικό ανάλογο. Στην ασυμπτωτική περίπτωση, αν οι $\overline{X}$ και S^2 είναι ασυμπτωτικά ανεξάρτητες, δηλαδή $\mu_3 = 0$, δεν έπεται την Κανονικότητα της τυχαίας μεταβλητής Χ, όπως προαναφέρθηκε στην Παρατήρηση 4.4.3. Δηλαδή, δεν αρκεί μόνο ένα k με $\mu_{k+1} = k\sigma^2\mu_{k-1}$ όπως στο Θεώρημα Geary. Όμως, αν οι ποσότητες $\overline{X}_n$ και $(S_n^2, M_{3,n}, \ldots, M_{k,n})$ είναι ασυμπτωτικά ανεξάρτητες για όλα τα $k \geq 2$, όπου S_n^2 πολλαπλάσιο της δειγματικής κεντρικής ροπής $M_{2,n}$, τότε $X \sim N(\mu, \sigma^2)$.

[3]Dr. Robert (Roy) Charles Geary (1896–1983): Ιρλανδός στατιστικός με σπουδές στα Πανεπιστήμια του Δουβλίνου και της Σορβόνης. Δίδαξε μαθηματικά στο Πανεπιστήμιο του Southampton και εφαρμοσμένα οικονομικά στο Πανεπιστήμιο του Cambridge. Το θεώρημα Geary και η συνάρτηση χρησιμότητας Stone-Geary φέρουν το όνομά του.

Βιβλιογραφία

[1] G. Afendras, *Moment-based inference for Pearson's quadratic q subfamily of distributions*, arXiv: 61104.0040v2.

[2] T. M. Apostol, *Calculus, Volume 1*, Jon Wiley and Sons, 1967.

[3] P. Billingsley, *Probability and Measure*, 2^{nd} edition, John Willey, 1986.

[4] K.L.Chung, *A Course in Probability Theory,* 3^{rd} edition, Academic Press, 2001.

[5] N. Etemadi, *An Elementary Proof of the Strong Law of Large Numbers,* Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete, Vol. 55, 1981, pp. 119-122.

[6] L. Feignbaum, *Brook Taylor and the Methods of Increments*, Springer-Verlang, 1921.

[7] T. S. Ferguson, *A Course in Large Sample Theory*, Chapman and Hall , 1996.

[8] G. A. Gibson, *Taylor's Theorem and Bernoulli's Theorem: A Historical Note*, Proceedings of the Edinburgh Mathematical Society, 1921, pp. 25-33.

[9] R. C. Geary, *The distributions of "Students" ratio for non-normal samples*, J. Roy. Statist. Soc. Ser. B, pp.178-184, 1936.

[10] A. Gut, *Stopped Random Walks: Limit Theorems and Applications*, Springer-Verlag, 1988.

[11] A. M. Kagan, Y. V. Linnik, C. R. Rao *Characterization Problems in Mathematical Statistics*, John Wiley, N.Y., 1973.

[12] Γ. Κοκολάκης, Ι. Σπηλιώτης *Εισαγωγή στις Πιθανότητες*, Εκδόσεις Συμεών, 2002.

[13] Γ. Κοκολάκης, Δ. Φουσκάκης *Στατιστική, Θεωρία και Εφαρμογές*, Εκδόσεις Συμεών, 2009.

[14] R. G. Laha, E. Lukacs, M. Newman *On the independence of a sample central moment and the sample mean*, Ann. Math. Statist., **31**, 1960, pp. 1028-1033.

[15] E. L. Lehmann, *Elements of Large-Sample Theory*, Springer, 1999.

[16] N. Παπαδάτος, *On the limiting distribution of sample central moments*, (http://users.uoa.gr/ npapadat/).

[17] Θ. M. Ρασσιάς, *Μαθηματική Ανάλυση Ι, Τεύχος Β'*, Εκδόσεις Σαββάλας, 2005.

[18] Γ. Σαραντόπουλος, *Σημειώσεις Θεωρίας Μέτρου και Ολοκλήρωσης*, Εκδόσεις Ε.Μ.Π, 2010.

[19] R. J. Serfling, *Approximation Theorems of Mathematical Statistics*, John Wiley, 1980.

[20] Ι. Σπηλιώτης, *Σημειώσεις του μαθήματος Θεωρία Πιθανοτητων*, ΕΜΠ, (http://www.math.ntua.gr/ jspil/).

[21] A. W. Van der Vaart, *Asymptotic Statistics*, Cambridge University Press, 1998.

[22] Χ. Χαραλαμπίδης, *Θεωρία Πιθανοτήτων και Εφαρμογές,* Εκδόσεις Συμμετρία, 2000.

[23] Wikipedia, Ιστορικά στοιχεία, Βιογραφικά Σημειώματα Μαθηματικων.

[24] A. A. Zinger *Independence of quasi polynomial statistics and analytical properties of distributions,* Teor. Veroy. i Primenen, **3**, 1958, pp.265-284.