

ΕΘΝΙΚΟ ΚΑΙ ΚΑΠΟΔΙΣΤΡΙΑΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ
ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΤΜΗΜΑ ΜΑΘΗΜΑΤΙΚΩΝ
ΤΟΜΕΑΣ ΠΙΘΑΝΟΤΗΤΩΝ, ΣΤΑΤΙΣΤΙΚΗΣ
& ΕΠΙΧΕΙΡΗΣΙΑΚΗΣ ΕΡΕΥΝΑΣ

ΜΕΤΑΠΤΥΧΙΑΚΗ ΔΙΑΤΡΙΒΗ

ΠΕΡΙ ΑΣΥΜΠΤΩΤΙΚΗΣ ΣΥΜΠΕΡΙΦΟΡΑΣ ΤΩΝ ΔΕΙΓΜΑΤΙΚΩΝ ΚΕΝΤΡΙΚΩΝ ΡΟΠΩΝ

ΚΩΝΣΤΑΝΤΙΝΟΣ Ν. ΨΑΛΤΗΣ
ΑΘΗΝΑ 2013

Στους γονείς μου,

Νικόλαο και Δέσποινα

Στους πνευματικούς μου γονείς,

Κωνσταντίνο και Μαρία

Στην αδελφή μου,

Ειρήνη

Εις μνήμην των παππούδων μου,

Ιωάννη και Κωνσταντίνο

Στα «Πρόσωπα» που πριν μερικά χρόνια

τους υποσχέθηκα ότι.....

Η Μεταπτυχιακή Διατριβή εκπονήθηκε, το Ακαδημαϊκό Έτος 2012-2013, στα πλαίσια του Προγράμματος Μεταπτυχιακών Σπουδών του Τμήματος Μαθηματικών του Εθνικού & Καποδιστριακού Πανεπιστημίου Αθηνών για την απόκτηση του Μεταπτυχιακού Διπλώματος Ειδίκευσης στον κλάδο Στατιστική και Επιχειρησιακή Έρευνα υπό την εποπτεία του Αναπληρωτή Καθηγητή κ. **Νικόλαου Παπαδάτου**.

Εγκρίθηκε στις 2 Απριλίου 2013 από Τριμελή Εξεταστική Επιτροπή αποτελούμενη από τους:

Ονοματεπώνυμο	Βαθμίδα	Υπογραφή
1) Νικόλαος Παπαδάτος,	Αναπληρωτής Καθηγητής	
2) Ευτυχία Βαγγελάτου,	Επίκουρη Καθηγήτρια	
3) Δημήτριος Χελιώτης,	Επίκουρος Καθηγητής	

Περιεχόμενα

	Σελ.
Πρόλογος.....	1

Κεφάλαιο 1

Εισαγωγικά στοιχεία και βασικές έννοιες Μαθηματικής Ανάλυσης, Θεωρίας Μέτρου, Θεωρίας Πιθανοτήτων & Μαθηματικής Στατιστικής

1.1 Μαθηματική Ανάλυση.....	7
1.1.1 Εισαγωγή.....	7
1.1.2 Σειρές Taylor.....	8
1.1.3 Το πολυώνυμο Taylor μιας συνάρτησης.....	9
1.1.4 Μονοδιάστατο Θεώρημα Taylor.....	13
1.1.5 Πολυδιάστατο Θεώρημα Taylor.....	15
1.2 Είδη στοχαστικών συγκλίσεων.....	19
1.3 Θεωρία Μέτρου.....	22
1.3.1 Εισαγωγή.....	22
1.4 Θεωρία Πιθανοτήτων & Μαθηματική Στατιστική.....	30
1.4.1 Εισαγωγή.....	30

Κεφάλαιο 2

Μέθοδος Δέλτα πρώτης τάξης

2.1 Η οριακή κατανομή των δειγματικών κεντρικών ροπών.....	41
2.1.1 Εισαγωγή.....	41
2.2 Η οριακή κατανομή και ο χαρακτηρισμός της κανονικότητας.....	44
2.3 Οι ιδιάζουσες κατανομές.....	73
2.4 Μέθοδος Δέλτα πρώτης τάξης.....	89

Κεφάλαιο 3
Μέθοδος Δέλτα ανώτερης τάξης

3.1 Μέθοδος Δέλτα δεύτερης τάξης.....	93
3.2 Η οριακή συμπεριφορά της ιδιάζουσας κατανομής.....	96
3.3 Πολυδιάστατη μέθοδος Δέλτα ανώτερης τάξης.....	103
Βιβλιογραφία.....	107

Πρόλογος

Στην Στατιστική, η μέθοδος Δέλτα χρησιμοποιείται για την, κατά προσέγγιση, εξαγωγή κατανομής πιθανότητας για μια συνάρτηση ενός ασυμπτωτικά κανονικού στατιστικού εκτιμητή από τη γνώση της οριακής διασποράς του εν λόγω εκτιμητή. Είναι επιπλέον γνωστή, στην επιστημονική βιβλιογραφία, και ως μέθοδος Πολλαπλασιασμού των Σφαλμάτων και αναφέρεται σε εφαρμογές του αποτελέσματος ότι μια ομαλή συνάρτηση μιας ασυμπτωτικά κανονικής εκτιμήτριας ακολουθεί, επίσης, ασυμπτωτικά κανονική κατανομή.

Η μέθοδος αυτή παρουσιάστηκε από τον *Harald Cramér* (25 Σεπτεμβρίου 1893 - 5 Οκτωβρίου 1985) ο οποίος ήταν Σουηδός μαθηματικός, αναλογιστής και στατιστικός με ειδίκευση στην Μαθηματική Στατιστική και Πιθανολογική Θεωρία Αριθμών¹. Περιγράφεται από τον *John Kingman*² ως "*ένας από τους γίγαντες της στατιστικής θεωρίας*". Ο *Harald Cramér* γεννήθηκε στη *Στοκχόλμη* όπου και εισήλθε στο Πανεπιστήμιο της ομώνυμης πόλης για να σπουδάσει μαθηματικά και χημεία. Κατά τη διάρκεια αυτής της περιόδου, υπήρξε βοηθός έρευνας στο πλευρό του διακεκριμένου χημικού *Hans von Euler-Chelpin*³ με τον οποίο δημοσίευσε τα πρώτα του πέντε επιστημονικά άρθρα κατά τα έτη 1913 – 1914. Μετά από την εμπειρία του στο εργαστήριο χημείας, αποφάσισε να επικεντρώσει το ενδιαφέρον του αποκλειστικά και μόνο στα μαθηματικά. Ξεκίνησε τελικά την εργασία του για διδακτορικές σπουδές υπό την επίβλεψη του *Riesz*⁴ και επηρεαζόμενος από τον *Hardy*⁵ εκπόνησε την διδακτορική του

¹Η Πιθανολογική Θεωρία Αριθμών είναι μία υποκατηγορία της Θεωρίας Αριθμών που χρησιμοποιεί την πιθανότητα για να απαντήσει στις ερωτήσεις της Θεωρίας Αριθμών. Μία βασική ιδέα στην οποία βασίζεται είναι ότι οι διαφορετικοί πρώτοι αριθμοί θεωρούνται ως ανεξάρτητες τυχαίες μεταβλητές. Ιδρυτές της θεωρίας αυτής ήταν ο *Paul Erdős*, ο *Aurel Wintner* και ο *Mark Kac*.

²Βρετανός μαθηματικός με πλήρες όνομα *Sir John Frank Charles Kingman*, 28 Αυγούστου 1939 - Σήμερα. Διετέλεσε καθηγητής των μαθηματικών επιστημών και διευθυντής του Ινστιτούτου "*Isaac Newton*" στο Πανεπιστήμιο του *Cambridge* από το 2001 έως το 2006. Είναι διάσημος για την ανάπτυξη των μαθηματικών της συνένωσης, δηλαδή, ενός θεωρητικού μοντέλου της κληρονομιάς, το οποίο είναι θεμελιώδες για την σύγχρονη γενετική του πληθυσμού.

³Σουηδός βιοχημικός, γεννημένος στην *Γερμανία*, με πλήρες όνομα *Hans Karl August Simon von Euler-Chelpin*, 15 Φεβρουαρίου 1873 - 6 Νοεμβρίου 1964. Κέρδισε, το βραβείο Νόμπελ χημείας το 1929, από κοινού με τον Άγγλο συνάδελφό του, *Sir Arthur Harden* (12 Οκτωβρίου 1865 - 17 Ιουνίου 1940) για τις έρευνές τους περί ζύμωσης της ζάχαρης και ζύμωσης των ενζύμων. Διετέλεσε καθηγητής της γενικής και οργανικής χημείας στο Πανεπιστήμιο της *Στοκχόλμης* (1906-1941) και διευθυντής του Ινστιτούτου για την έρευνα της οργανικής χημείας κατά τα έτη 1938-1948.

⁴Ο *Marcel Riesz* ήταν Ουγγρικής καταγωγής μαθηματικός (16 Νοεμβρίου 1886 - 4 Σεπτεμβρίου 1969), γνωστός για τις εργασίες του σχετικά με τις μεθόδους άθροισης, δυναμικής θεωρίας, μερικών διαφορικών εξισώσεων, θεωρίας αριθμών και άλγεβρας *Clifford*.

διατριβή με τίτλο *"Περί της τάξης των σειρών Dirichlet"*⁶. Μετά το διδακτορικό του, υπηρέτησε ως Επίκουρος Καθηγητής Μαθηματικών στο Πανεπιστήμιο της Στοκχόλμης κατά τα έτη 1917 - 1929. Τα πρώτα χρόνια, ο Cramér είχε μεγάλη συμμετοχή στη Θεωρία Αριθμών. Ακολούθησαν, ορισμένες σημαντικές στατιστικές συνεισφορές του, για την κατανομή των πρώτων και δίδυμων πρώτων αριθμών⁷. Διακεκριμένο σύγγραμμα επ' αυτού, υπήρξε εκείνο με τίτλο *"Περί τάξης μεγέθους της διαφοράς των πρώτων αριθμών"*, το οποίο προσέφερε έναν αυστηρό απολογισμό του εποικοδομητικού ρόλου στον οποίο εφαρμόζεται η πιθανότητα στην Θεωρία Αριθμών. Στα τέλη της δεκαετίας του 1920, ο Cramér άρχισε να ενδιαφέρεται για την έννοια της πιθανότητας, που εκείνη την εποχή δεν ήταν αποδεκτός κλάδος των μαθηματικών. Ο Cramér, όμως, είχε την πεποίθηση ότι μια ριζική αλλαγή ήταν απαραίτητη σε αυτόν τον τομέα, χαρακτηριστικά δήλωσε σε ένα βιβλίο του το 1926 ότι *"Η έννοια της πιθανότητας θα πρέπει να εισαχθεί από έναν αυστηρά μαθηματικό ορισμό, από τον οποίο οι θεμελιώδεις ιδιότητές της και τα κλασικά θεωρήματα θα συμπεραίνονται από αμιγώς μαθηματικές πράξεις"*. Στις αρχές του 1930 έδωσε μια ενδιαφέρουσα αυστηρή μαθηματική διατύπωση της πιθανότητας στις εργασίες των συναδέλφων του μαθηματικών Kolmogorov⁸, Lévy⁹, Natanovich¹⁰, Khinchin¹¹. Αργότερα, ο Cramér έγραψε μια ενδελεχή μελέτη πάνω στο αντικείμενό του, η οποία δημοσιεύτηκε στο Πανεπιστήμιο του Cambridge το 1937, με τίτλο *"Τυχαίες μεταβλητές και κατανομές πιθανότητας"*, και μάλιστα, τούτο αποτέλεσε σημαντική συμβολή στην πρόοδο της Θεωρίας Πιθανοτήτων. Λίγο μετά το τέλος του Β΄ Παγκοσμίου Πολέμου (2 Σεπτεμβρίου 1945) ο Cramér έκανε γνωστά τα συμπεράσματα των ερευνών του, το 1946, στο άρθρο του με τίτλο *"Μαθηματικές Μέθοδοι Στατιστικής"*. Εν κατακλείδι, αυτό το κείμενο ήταν που *"έδειξε τον τρόπο με τον οποίο η*

⁵Ο Godfrey Harold Hardy ήταν Άγγλος μαθηματικός, γνωστός για τα επιτεύγματά του στη Θεωρία Αριθμών και στη Μαθηματική Ανάλυση, 7 Φεβρουαρίου 1877 - 1 Δεκεμβρίου 1947.

⁶Γερμανός μαθηματικός με πλήρες όνομα Johann Peter Gustav Lejeune Dirichlet, 13 Φεβρουαρίου 1805 - 5 Μαΐου 1859. Είχε μεγάλη συνεισφορά στην Θεωρία Αριθμών, στις σειρές Fourier και σε άλλα θέματα Μαθηματικής Ανάλυσης. Επίσης, σε αυτόν πιστώνεται ο ορισμός της συνάρτησης.

⁷Δίδυμοι πρώτοι αριθμοί λέγονται οι πρώτοι αριθμοί που έχουν διαφορά δύο. Για παράδειγμα, τα ζεύγη 5 & 7, 17 & 19, 29 & 31 αποτελούν τρεις τέτοιες χαρακτηριστικές περιπτώσεις.

⁸Ρώσος μαθηματικός με πλήρες όνομα Andrey Nikolaevich Kolmogorov, 25 Απριλίου 1903 - 20 Οκτωβρίου 1987. Διέπρεψε σε διάφορα επιστημονικά πεδία, μεταξύ των οποίων, θεωρία πιθανοτήτων, τοπολογία, ενορατική λογική, αναταράξεις, κλασική μηχανική και υπολογιστική πολυπλοκότητα.

⁹Ο Paul Pierre Lévy (15 Σεπτεμβρίου 1886 - 15 Δεκεμβρίου 1971) ήταν Γάλλος μαθηματικός, ο οποίος δραστηριοποιήθηκε κυρίως στην θεωρία πιθανοτήτων εισάγοντας τα martingales, την κατανομή Lévy, τα μέτρα Lévy, τις ανεξίτηλες Lévy, το νόμο Lévy περί τόξο ημιτόνου, την κλασματική καμπύλη C, την σταθερά Lévy, τις πτήσεις Lévy και την άλφα-ευσταθή κατανομή λοξότητας Lévy.

¹⁰Ρώσος μαθηματικός με πλήρες όνομα Sergei Bernstein Natanovich (5 Μαρτίου 1880 - 26 Οκτωβρίου 1968), συνεισέφερε στις μερικές διαφορικές εξισώσεις, στην θεωρία πιθανοτήτων, στη διαφορική γεωμετρία και στην θεωρία προσέγγισης.

¹¹Ρώσος μαθηματικός με πλήρες όνομα Aleksandr Yakovlevich Khinchin, 19 Ιουλίου 1894 - 18 Νοεμβρίου 1959. Ασχολήθηκε με την θεωρία αριθμών, την πραγματική ανάλυση και την θεωρία πιθανοτήτων στην οποία πέτυχε σημαντικά αποτελέσματα στα οριακά θεωρήματα. Επίσης, διατύπωσε τον ορισμό της στασιμότητας και δημοσίευσε πληθώρα έργων στην στατιστική φυσική.

στατιστική πρακτική εξαρτάται από ένα σύνολο αυστηρής μαθηματικής ανάλυσης, καθώς επίσης, και της Fisherian¹² διαίσθησης".

Η μέθοδος Δέλτα αποτελείται από δύο κύριες ομάδες εφαρμογών. Η πρώτη εστιάζεται στον υπολογισμό ενός μετασχηματισμού σταθεροποίησης διασποράς και η δεύτερη στον υπολογισμό της ασυμπτωτικής διακύμανσης και των διαστημάτων εμπιστοσύνης για μια μη-γραμμική συνάρτηση ενός συνόλου παραμέτρων. Είθισται, οι παράμετροι αυτοί να εκτιμώνται με τη μέθοδο Μέγιστης Πιθανοφάνειας, και να ακολουθούν ασυμπτωτικά κανονική κατανομή. Ευρύτερα, η μέθοδος αυτή μπορεί να θεωρηθεί ως ένα αρκετά γενικό Κεντρικό Οριακό Θεώρημα με τη βοήθεια του οποίου υπολογίζουμε, προσεγγιστικά πάντα, δοθέντος του μεγέθους της κρίσιμης περιοχής (επιπέδου στατιστικής σημαντικότητας) α , διαστήματα εμπιστοσύνης βαθμού $100(1-\alpha)\%$ για συναρτήσεις Εκτιμητών Μέγιστης Πιθανοφάνειας. Συγκεκριμένα, η μέθοδος Δέλτα παίρνει μια συνάρτηση η οποία όμως, στην πλειονότητα των περιπτώσεων, είναι υπερβολικά πολύπλοκη για αναλυτικό υπολογισμό της διακύμανσης και για το λόγο αυτό χρησιμοποιώντας το ανάπτυγμα σε σειρά *Taylor*, δημιουργεί μια γραμμική προσέγγιση της συνάρτησης αυτής, και εν συνεχεία, λογαριάζει τη διασπορά της απλούστερης γραμμικής συνάρτησης που μπορεί να χρησιμοποιηθεί στον έλεγχο υποθέσεων (μηδενική έναντι εναλλακτικής) για στατιστική συμπερασματολογία μεγάλου τυχαίου δείγματος από έναν πληθυσμό ως προς κάποιο χαρακτηριστικό (ηλικία, βάρος, ύψος, βαθμός σε διαγώνισμα, δείκτης νοημοσύνης κτλ.) των εκάστοτε πειραματικών μονάδων (άνθρωποι, ζώα, μηχανές, φυτά, υλικά αγαθά κτλ.).

Η μέθοδος Δέλτα εφαρμόζεται κυρίως στην κανονική κατανομή διότι είναι η πιο ευρέως χρησιμοποιούμενη κατανομή, καθώς αν υποθεθεί για έναν πληθυσμό, τότε δίνει οντότητα σε ένα σύνολο στατιστικών συμπερασμάτων. Επομένως, όσο σημαντική είναι αυτή η κατανομή εξίσου σημαντικό είναι να γνωρίζουμε πότε ξεφεύγουμε από αυτήν την ιδανική κατάσταση της κανονικότητας.

Σκοπός μας στη διατριβή αυτή, είναι να διερευνήσουμε την ασυμπτωτική συμπεριφορά περί δειγματικών κεντρικών ροπών η αξία των οποίων αναδεικνύεται στην αναλυτική περιγραφή της μεθόδου Δέλτα πρώτης, δεύτερης και ανώτερης τάξης τόσο στις μονοπαραγοντικές όσο και στις πολυπαραγοντικές περιπτώσεις. Προς αυτή την κατεύθυνση, στο πρώτο κεφάλαιο θα αναφερθούμε σε ορισμένα στοιχεία Μαθηματικής Ανάλυσης,

¹²Ο *Ronald Aylmer Fisher* υπήρξε Άγγλος στατιστικός, εξελικτικός βιολόγος, ευγονιστής και γενετιστής, 17 Φεβρουαρίου 1890 - 29 Ιουλίου 1962. Είναι πολύ γνωστός για την συμβολή του στη Στατιστική, με τη δημιουργία του πίνακα ANOVA (Ανάλυση Διακύμανσης), τεστ ακρίβειας του *Fisher* και εξίσωση του *Fisher*. Ο Δανός στατιστικός *Anders Hald* τον αποκάλεσε ως "μια ιδιοφυΐα που σχεδόν μόνος του δημιούργησε τα θεμέλια για την σύγχρονη Στατιστική" ενώ ο Βρετανός εξελικτικός βιολόγος *Clinton Richard Dawkins* τον χαρακτήρισε ως "τον μεγαλύτερο βιολόγο μετά το Δαρβίνο".

Θεωρίας Μέτρου, Θεωρίας Πιθανοτήτων και Μαθηματικής Στατιστικής τα οποία, συμβάλλουν στην κατανόηση των εννοιών που έπονται. Στο πλαίσιο αυτό, θα παραθέσουμε στο δεύτερο και τρίτο κεφάλαιο, Ορισμούς, Προτάσεις, Λήμματα, Πορίσματα και Παρατηρήσεις, σχετικά με την οριακή συμπεριφορά των ροπών, τα οποία, βοηθούν στις αποδείξεις των Θεωρημάτων των μεθόδων Δέλτα.

Η παρούσα διατριβή αποτελεί το τελευταίο μέρος των υποχρεώσεων μου για την απόκτηση του Μεταπτυχιακού Διπλώματος Ειδίκευσης στον κλάδο Πιθανοτήτων, Στατιστικής και Επιχειρησιακής Έρευνας που προσφέρει το Τμήμα Μαθηματικών του Εθνικού & Καποδιστριακού Πανεπιστημίου Αθηνών.

Εκμεταλλευόμενος τη θέση αυτή, επιθυμώ να εκφράσω την ευγνωμοσύνη και τις ευχαριστίες μου σε όσους συνέβαλαν στην αποπεράτωση αυτής της διατριβής.

Στον κ. Νικόλαο Παπαδάτο, Αναπληρωτή Καθηγητή του Τμήματος Μαθηματικών του Εθνικού & Καποδιστριακού Πανεπιστημίου Αθηνών, που με προθυμία δέχτηκε ν' αναλάβει την επίβλεψη και πρότεινε το θέμα της διατριβής. Τον ευχαριστώ θερμά για την επιστημονική βοήθεια που μου προσέφερε, την υπομονή και τον πολύτιμο χρόνο που μου αφιέρωσε ώστε να έρθει σε πέρας αυτή η διατριβή. Χωρίς τη συμβολή του κ. Παπαδάτου, η διατριβή αυτή θα ήταν αδύνατο να διεκπεραιωθεί. Τα ιδιαίτερα καλά του λόγια για το τελικό αποτέλεσμα ήταν η καλύτερη επιβράβευση της προσπάθειας μου.

Στην κ. Ευτυχία Βαγγελάτου, Επίκουρη Καθηγήτρια του Τμήματος Μαθηματικών του Εθνικού & Καποδιστριακού Πανεπιστημίου Αθηνών, για την συμμετοχή της στην Τριμελή Επιτροπή, καθώς και για τις ωφέλιμες συζητήσεις που είχα μαζί της σε θέματα ασυμπτωτικής στατιστικής.

Στον κ. Δημήτριο Χελιώτη, Επίκουρο Καθηγητή του Τμήματος Μαθηματικών του Εθνικού & Καποδιστριακού Πανεπιστημίου Αθηνών, για την συμμετοχή του στην Τριμελή Επιτροπή και ιδιαίτερα για τα εύστοχα σχόλια και τις χρήσιμες συμβουλές του.

Τέλος, εκφράζω τις θερμές μου ευχαριστίες προς τους γονείς μου Νικόλαο και Δέσποινα που με στήριξαν αποφασιστικά με το δικό τους τρόπο ο καθένας μέχρι τούτη τη στιγμή. Επιπλέον, ευχαριστώ ιδιαίτερα τους πνευματικούς μου γονείς Κωνσταντίνο και Μαρία που με βοήθησαν σε δύσκολες στιγμές της ζωής μου, καθώς επίσης, και τους Ιατρούς Παναγιώτη Σπυρίδη, Κωνσταντίνο Σκόνδρα, Βασίλειο Καφαράκη, Φρειδερίκη Καρβούνη, Άννα Γρίβα, Γεώργιο Πολίτη, Αναστάσιο Παππά, Ηλία Σταυρινίδη και Μάριο Κωνσταντίνου οι οποίοι με την επιστημονική τους γνώση συνεισέφεραν για να θεραπευτώ και να αναρρώσω γρήγορα. Τις κ. Μαρία Ελευθερίου, Αναστασία Τσίτση και Στέλλα Αλεξάκη, που ως προϊσταμένες με μεγάλη φροντίδα συνέβαλαν στη

νοσηλεία μου, τις ευχαριστώ θερμά. Βέβαια, πάνω απ' όλους ευχαριστώ Τον Ένα και Μοναδικό Τριαδικό Θεό που με αξίωσε να φτάσω σε αυτό το σημείο χαρίζοντας μου φώτιση.

Κωνσταντίνος Ν. Ψάλτης
Αθήνα, Μάρτιος 2013

Κεφάλαιο 1

Εισαγωγικά στοιχεία και βασικές έννοιες Μαθηματικής Ανάλυσης, Θεωρίας Μέτρου, Θεωρίας Πιθανοτήτων & Μαθηματικής Στατιστικής

1.1 Μαθηματική Ανάλυση

1.1.1 Εισαγωγή

Η Μαθηματική Ανάλυση είναι ένα από τα πιο ευρεία πεδία των μαθηματικών. Θεμελιωτές της ήταν ο *Gottfried Wilhelm Leibniz*¹³ και ο *Isaac Newton*¹⁴, οι οποίοι την ανακάλυψαν ανεξάρτητα στα τέλη του 17^{ου} αιώνα.

Κλάδοι της Μαθηματικής Ανάλυσης είναι ο Διαφορικός και Ολοκληρωτικός Λογισμός (οι οποίοι συλλήβδην καλούνται και "*Απειροστικός Λογισμός*"), η Τοπολογία, η Συναρτησιακή Ανάλυση και η Θεωρία Μέτρου.

¹³Ο *Gottfried Wilhelm von Leibniz* (1 Ιουλίου 1646 - 14 Νοεμβρίου 1716) υπήρξε Γερμανός μαθηματικός και φιλόσοφος. Κατέχει εξέχουσα θέση στην ιστορία των μαθηματικών και στην ιστορία της φιλοσοφίας. Μεταξύ άλλων, ανέπτυξε τον απειροστικό λογισμό και υποστήριξε το φιλοσοφικό δόγμα του ορθολογισμού.

¹⁴Ο *Sir Isaac Newton* (4 Ιανουαρίου 1643 – 31 Μαρτίου 1727) ήταν Άγγλος φυσικός, μαθηματικός, αστρονόμος, φιλόσοφος, αλχημιστής και θεολόγος. Θεωρείται πατέρας της Κλασικής Φυσικής, καθώς ξεκινώντας από τις παρατηρήσεις του *Γαλιλαίου* αλλά, και από τους νόμους του *Κέπλερ* για την κίνηση των πλανητών, διατύπωσε τους τρεις μνημειώδεις νόμους της κίνησης και τον περισπούδαστο "*νόμο της βαρύτητας*" (που η ιστορία αναφέρει πως αναζήτησε μετά από πτώση μήλου από μία μηλιά). Μεγάλης ιστορικής σημασίας υπήρξαν οι μελέτες του σχετικά με τη φύση του φωτός και είχε, επίσης, καθοριστική συμβολή στη θεμελίωση των σύγχρονων μαθηματικών και συγκεκριμένα του διαφορικού και ολοκληρωτικού λογισμού. Επιπλέον, γενίκευσε το διωνυμικό θεώρημα για μη-ακέραιους εκθέτες, ανέπτυξε την μέθοδο *Newton* για την προσέγγιση των ριζών μιας συνάρτησης και συνέβαλε στη μελέτη των δυναμοσειρών.

1.1.2 Σειρές *Taylor*

Στον πρόλογο ελέχθη ότι στην περίπτωση που συναντήσουμε συνάρτηση η οποία δεν επιδέχεται αναλυτικό υπολογισμό της διασποράς τότε, για να άρουμε αυτή τη δυσκολία, θα αναπτύξουμε την προκειμένη συνάρτηση σε σειρά *Taylor*. Στα μαθηματικά, σειρά *Taylor* είναι η αναπαράσταση μίας συνάρτησης ως άθροισμα απείρων όρων οι οποίοι υπολογίζονται από τις τιμές των παραγώγων της σε ένα συγκεκριμένο σημείο, το οποίο καλείται κέντρο της δυναμοσειράς. Η σειρά αυτή οφείλει το όνομά της στον Άγγλο μαθηματικό *Brook Taylor* (18 Αυγούστου 1685 - 29 Δεκεμβρίου 1731) ο οποίος γεννήθηκε στην κωμόπολη *Έντμοντον* στο *Μίντλεσεξ* της *Αγγλίας* και αργότερα σπούδασε μαθηματικά στο κολλέγιο του *Αγίου Ιωάννη* στο Πανεπιστήμιο του *Cambridge* υπό την εποπτεία του *Machin*¹⁵ και του *Keill*¹⁶. Το 1708 απέκτησε μια αξιόλογη λύση του προβλήματος "*του κέντρου της ταλάντωσης*" η οποία ως εκ τούτου, παρέμεινε αδημοσίευτη μέχρι το Μάιο του 1714, όταν το αίτημά του για προτεραιότητα αμφισβητήθηκε από τον *Johann Bernoulli*¹⁷. Ο *Taylor* το 1715 στο βιβλίο του με τίτλο "*Methodus Incrementorum Directa et Inversa*" (*Μέθοδος Άμεσων Προσαυξήσεων και Αντιστροφών*) προσθέτει έναν νέο κλάδο στα ανώτερα μαθηματικά γνωστό ως "*Λογισμός των πεπερασμένων διαφορών*". Μεταξύ άλλων ευφυών εφαρμογών του, χρησιμοποίησε αυτόν τον κλάδο για να καθορίσει τη μορφή κίνησης της παλλόμενης χορδής. Το ίδιο έργο περιλαμβάνει την περίφημη φόρμουλα γνωστή ως "*Θεώρημα Taylor*" η σημασία του οποίου παρέμεινε μη αναγνωρισμένη έως το 1772, όταν ο *Joseph-Louis Lagrange*¹⁸ συνειδητοποίησε την ισχύ του και το χαρακτήρισε ως "*le principal fondement du calcul différentiel*" τουτέστιν "*το κύριο θεμέλιο του Διαφορικού Λογισμού*".

Οι σειρές *Taylor* λοιπόν, εισήχθησαν από τον Άγγλο μαθηματικό *Brook Taylor* το 1715. Αν η σειρά έχει κέντρο το μηδέν, τότε καλείται επίσης σειρά *Maclaurin*, προς τιμήν του Σκωτσέζου μαθηματικού *Colin*

¹⁵Ο *John Machin* (1686 – 9 Ιουνίου 1751) ήταν Άγγλος μαθηματικός και καθηγητής αστρονομίας στο κολλέγιο *Gresham* του *Λονδίνου*. Είναι πολύ γνωστός διότι το 1706 ανέπτυξε μία, γρήγορα, συγκλίνουσα σειρά για την μαθηματική σταθερά π. Η φόρμουλα αυτή είναι $\frac{\pi}{4} = 4 \cot^{-1} 5 - \cot^{-1} 239$ και χρησιμοποιεί έως 100 δεκαδικά ψηφία για να προσεγγίσει την σταθερά π. Το πλεονέκτημα του τύπου αυτού, συγκριτικά με εκείνον που πρότεινε ο *Leibnitz* ($\frac{\pi}{4} = \arctan 1$), είναι ότι έχει σημαντικά αυξημένο ρυθμό σύγκλισης, γεγονός, που τον καθιστά μία πολύ πιο πρακτική μέθοδο υπολογισμού.

¹⁶Ο *John Keill* (1 Δεκεμβρίου 1671 - 31 Αυγούστου 1721) ήταν Σκωτσέζος μαθηματικός. Σπούδασε στο Πανεπιστήμιο του *Εδιμβούργου* και έλαβε το πτυχίο του το 1692 με διάκριση στη φυσική και στα μαθηματικά. Υπήρξε δε, μαθητής του *Νεύτωνα*.

¹⁷Διακεκριμένος Ελβετός μαθηματικός με συμβολή στον απειροστικό λογισμό. Είναι επίσης, γνωστός και ως *Jean Bernoulli* ή *John Bernoulli*, 27 Ιουλίου 1667 - 1 Ιανουαρίου 1748.

¹⁸Γάλλος μαθηματικός και αστρονόμος, γνωστός και ως *Giuseppe Luigi Lagrangia* στην επιστημονική βιβλιογραφία, 25 Ιανουαρίου 1736 – 10 Απριλίου 1813. Είχε σημαντική συνεισφορά σε όλους τους τομείς της ανάλυσης, στην θεωρία αριθμών, στην κλασική και ουράνια μηχανική.

Maclaurin (Φεβρουάριος 1698 - 14 Ιουνίου 1746) ο οποίος έκανε εκτεταμένη χρήση αυτής της ειδικής περίπτωσης των σειρών *Taylor* τον 18^ο αιώνα.

Είναι κοινή πρακτική να γίνεται χρήση πεπερασμένου πλήθους όρων μιας σειράς για την προσέγγιση μιας συνάρτησης. Το Θεώρημα *Taylor* δίνει ποσοτικές εκτιμήσεις για το σφάλμα της προσέγγισης. Οποιοσδήποτε πεπερασμένος αριθμός αρχικών όρων της σειράς καλείται πολυώνυμο *Taylor*. Η σειρά *Taylor* μίας συνάρτησης είναι το όριο του πολυωνύμου *Taylor* αυτής της συνάρτησης, υπό την προϋπόθεση ότι το όριο αυτό υπάρχει. Μία συνάρτηση ενδέχεται να μην ισούται με την οικεία σειρά *Taylor*, ακόμη και αν η σειρά *Taylor* συγκλίνει σε κάθε σημείο. Μία συνάρτηση η οποία είναι ίση με την οικεία σειρά *Taylor* σε ένα ανοιχτό διάστημα (ή δίσκο στο μιγαδικό επίπεδο) είναι γνωστή ως αναλυτική συνάρτηση. Μολαταύτα, το Θεώρημα *Taylor* είναι ένα αξιολογικό αποτέλεσμα του Διαφορικού Λογισμού, το οποίο αποβλέπει στην επέκταση μίας επαρκούς ομαλής συνάρτησης γύρω από ένα σημείο.

1.1.3 Το Πολυώνυμο *Taylor* μιας συνάρτησης

Έστω ότι δίνεται μία πραγματική ή μιγαδική συνάρτηση $f(x)$ που είναι απείρως παραγωγίσιμη, αντίστοιχα, σε μία γειτονιά ενός πραγματικού ή μιγαδικού αριθμού a , ο οποίος είναι (εσωτερικό) σημείο του πεδίου ορισμού της. Επιζητούμε εκείνο το πολυώνυμο βαθμού n που προσεγγίζει καλύτερα την συνάρτηση αυτή γύρω από το σημείο a . Δηλαδή, θέλουμε το πολυώνυμο να έχει όσο το δυνατόν περισσότερα κοινά χαρακτηριστικά με την συνάρτηση $f(x)$ σε μια γειτονιά του σημείου a . Γνωρίζουμε ότι η γεωμετρία μίας συνάρτησης συσχετίζεται με τις τιμές των παραγώγων της ($1^{\text{ης}}$, $2^{\text{ης}}$ κ.ο.κ. τάξης) στο συγκεκριμένο σημείο. Επιδιώκουμε λοιπόν, ένα πολυώνυμο $P_n^a(x)$ βαθμού n με την ιδιότητα $P_n^{a(k)}(a) = f^{(k)}(a)$, για όσο το δυνατόν περισσότερες τιμές του k . Είναι εύκολο να δούμε ότι για ένα πολυώνυμο $P_n^a(x)$ βαθμού n ισχύει πάντα ότι $P_n^{a(k)}(x) = 0$, για κάθε $k \geq n + 1$. Συνεπώς, το βέλτιστο που μπορούμε να ζητήσουμε είναι $P_n^{a(k)}(a) = f^{(k)}(a)$, για κάθε $k = 0, 1, \dots, n$.

Κατόπιν, γράφουμε το πολυώνυμο $P_n^a(x)$ σε δυνάμεις του $(x - a)$, δηλαδή στην μορφή $P_n^a(x) = \alpha_0 + \alpha_1(x - a) + \dots + \alpha_n(x - a)^n$. Αυτό επιτυγχάνεται με διαδοχικές διαιρέσεις του πολυωνύμου με το $(x - a)$. Έτσι, παρατηρούμε ότι $P_n^{a(k)}(x) = k! \alpha_k + (x - a)Q_{n-k}(x)$, για κάθε $k = 0, 1, \dots, n$, όπου $Q_{n-k}(x)$ κάποιο πολυώνυμο βαθμού $n - k$. Αυστηρή απόδειξη αυτού μπορεί να γίνει με επαγωγή ως προς το k .

Θα πρέπει επομένως,

$$f^{(k)}(a) = P_n^{a(k)}(a) = k! \alpha_k + (a - a) Q_{n-k}(a) = k! \alpha_k$$

και έτσι, $\alpha_k = \frac{f^{(k)}(a)}{k!}$. Άρα το ζητούμενο πολυώνυμο βαθμού n είναι το

$$P_n^a(x) = f(a) + \frac{f^{(1)}(a)}{1!}(x-a)^1 + \frac{f^{(2)}(a)}{2!}(x-a)^2 + \dots + \frac{f^{(n)}(a)}{n!}(x-a)^n,$$

όπου $n!$ υποδηλώνει το παραγοντικό του n και $f^{(n)}(a)$ την τιμή της n -οστής παραγώγου της f στο σημείο a .

Το παραπάνω πολυώνυμο καλείται πολυώνυμο *Taylor* βαθμού n για τη συνάρτηση f που αναπτύσσεται γύρω από το σημείο a . Η μηδενικής τάξης παράγωγος της f ορίζεται να είναι η ίδια η f και οι όροι $(x-a)^0$ και $0!$ είναι εξ' ορισμού ίσοι με 1. Στην περίπτωση που $a = 0$, η σειρά αποκαλείται και σειρά *Maclaurin*.

Θεώρημα 1.1: (Ολοκληρωτικός τύπος του Cauchy για το υπόλοιπο Taylor)

Έστω $h = h_1 + ih_2: I \rightarrow \mathbb{C}$ τυχούσα συνάρτηση με μιγαδικές τιμές, όπου το I είναι κάποιο διάστημα της μορφής $[0, \alpha)$, ή $(-\alpha, 0]$, ή $(-\alpha_1, \alpha_2)$, όπου $\alpha, \alpha_1, \alpha_2 \in (0, +\infty]$.

(i) Αν για κάποιο $m \geq 1$, οι (πραγματικές) συναρτήσεις h_1 και h_2 έχουν συνεχείς παραγώγους, m τάξεως, στο I , τότε

$$h(x) - \sum_{j=0}^{m-1} \frac{h^{(j)}(0)}{j!} x^j = \frac{1}{(m-1)!} \int_0^x (x-y)^{m-1} h^{(m)}(y) dy, \quad x \in I, \quad (1.1)$$

όπου $h^{(j)}(x) = h_1^{(j)}(x) + ih_2^{(j)}(x)$, και $h_n^{(j)}(x) \frac{d^j}{dx^j} h_n(x)$, $j = 0, 1, \dots, m$, $n = 1, 2$.
[Φυσικά, $h_n^{(0)} = h_n$, $n = 1, 2$ και $h^{(0)} = h$.]

(ii) Αν για κάποιο $m \geq 1$, οι (πραγματικές) συναρτήσεις h_1 και h_2 έχουν συνεχείς παραγώγους, m τάξεως, στο I , τότε

$$h(x) - \sum_{j=0}^m \frac{h^{(j)}(0)}{j!} x^j = \frac{1}{(m-1)!} \int_0^x (x-y)^{m-1} [h^{(m)}(y) - h^{(m)}(0)] dy, \quad x \in I, \quad (1.2)$$

όπου $h^{(j)}(x)$ ($j = 0, 1, \dots, m$) όπως στο (i).

Απόδειξη: Χωρίζοντας σε πραγματικά και φανταστικά μέρη, γίνεται φανερό ότι είναι αρκετό να αποδείξουμε τις ισότητες για πραγματικές συναρτήσεις $h: I \rightarrow \mathcal{R}$. Έστω ότι ισχύει η (1.2) ξεχωριστά για τις πραγματικές συναρτήσεις $h = h_1$ και $h = h_2$, τότε θα προκύψει ότι ισχύει και για την $h = h_1 + ih_2$, ως εξής:

$$\begin{aligned}
h(x) - \sum_{j=0}^m \frac{h^{(j)}(0)}{j!} x^j &= \left[h_1(x) - \sum_{j=0}^m \frac{h_1^{(j)}(0)}{j!} x^j \right] + i \left[h_2(x) - \sum_{j=0}^m \frac{h_2^{(j)}(0)}{j!} x^j \right] \\
&= \frac{1}{(m-1)!} \int_0^x (x-y)^{m-1} [h_1^{(m)}(y) - h_1^{(m)}(0)] dy \\
&\quad + i \frac{1}{(m-1)!} \int_0^x (x-y)^{m-1} [h_2^{(m)}(y) - h_2^{(m)}(0)] dy \\
&= \frac{1}{(m-1)!} \int_0^x (x-y)^{m-1} \{ [h_1^{(m)}(y) - h_1^{(m)}(0)] + i [h_2^{(m)}(y) - h_2^{(m)}(0)] \} dy \\
&= \frac{1}{(m-1)!} \int_0^x (x-y)^{m-1} [h^{(m)}(y) - h^{(m)}(0)] dy.
\end{aligned}$$

Για το (i) υποθέτουμε ότι η h είναι πραγματική συνάρτηση με συνεχή παράγωγο τάξεως m , και θεωρούμε την συνάρτηση

$$H_{m-1}(x) = H_{m-1}(x; h^{(m)}) = \frac{1}{(m-1)!} \int_0^x (x-y)^{m-1} h^{(m)}(y) dy, \quad m = 1, 2, \dots$$

Έτσι, προκύπτει ότι

$$H_{m-1}^{(m)}(x) = h^{(m)}(x) \quad \text{για κάθε } x \in I,$$

όπου H είναι μία παράγουσα (δηλ. μία αντιπαράγωγος) m τάξεως της h .

Θέτοντας

$$R_{m-1}(x) = h(x) - \sum_{j=0}^{m-1} \frac{h^{(j)}(0)}{j!} x^j$$

παρατηρούμε ότι και $R_{m-1}^{(m)}(x) = h^{(m)}(x)$, δηλαδή,

$$\frac{d^m}{dx^m} [R_{m-1}(x) - H_{m-1}(x)] = h^{(m)}(x) - h^{(m)}(x) = 0, \quad \text{για κάθε } x \in I,$$

οπότε η συνάρτηση $P_{m-1}(x) = R_{m-1}(x) - H_{m-1}(x)$ είναι πολυώνυμο, βαθμού το πολύ $m - 1$. Αφού για κάθε $n \in \{0, 1, \dots, m - 1\}$ είναι¹⁹ $R_{m-1}^{(n)}(0) = 0$, και ομοίως, $H_{m-1}^{(n)}(0) = H_{m-1-n}(0) = 0$, έπεται ότι $P_{m-1}^{(n)}(0) = 0$ για κάθε $n = 0, 1, \dots, m - 1$, και συνεπώς, $P_{m-1}(x) = 0$ για κάθε $x \in I$. Επομένως, $R_{m-1}(x) = H_{m-1}(x)$ για κάθε $x \in I$, που αποδεικνύει την (1.1).

Για το (ii) υποθέτουμε ότι η h είναι πραγματική συνάρτηση, με συνεχή παράγωγο τάξεως m στο I , και επιπλέον, θεωρούμε την συνάρτηση

$$\begin{aligned} H_{m-1}(x) &= H_{m-1}(x; h^{(m)} - c) \\ &= \frac{1}{(m-1)!} \int_0^x (x-y)^{m-1} [h^{(m)}(y) - c] dy, \quad m = 1, 2, \dots, \end{aligned}$$

όπου $c = h^{(m)}(0)$. Άρα, απορρέει ότι

$$H_{m-1}^{(n)}(x) = H_{m-1-n}(x), \quad n = 0, 1, \dots, m - 1,$$

και έτσι,

$$H_{m-1}(0) = H'_{m-1}(0) = H''_{m-1}(0) = \dots = H_{m-1}^{(m-1)}(0) = 0.$$

Επίσης,

$$H_{m-1}^{(m)}(x) = \left(H_{m-1}^{(m-1)}(x) \right)' = H'_0(x) = h^{(m)}(x) - h^{(m)}(0).$$

Όμως, και για το υπόλοιπο *Taylor*, R_m , ισχύουν ακριβώς τα ίδια:

$$R_m(0) = R'_m(0) = R''_m(0) = \dots = R_m^{(m-1)}(0) = 0,$$

$$R_m^{(m)}(x) = h^{(m)}(x) - h^{(m)}(0).$$

Οπότε, $(R_m - H_{m-1})^{(m)} \equiv 0$, που σημαίνει ότι το $R_m - H_{m-1} = P_{m-1}$ είναι πολυώνυμο, βαθμού το πολύ $m - 1$. Από τις συνθήκες

$$P_{m-1}^{(n)}(0) = R_m^{(n)}(0) - H_{m-1}^{(n)}(0) = 0, \quad n = 0, 1, \dots, m - 1,$$

έπεται ότι $P_{m-1}(x) \equiv 0$, δηλαδή, $R_m(x) = H_{m-1}(x)$ για κάθε $x \in I$, που αποδεικνύει την (1.2). ■

¹⁹Από τον ορισμό του, το προσεγγιστικό πολυώνυμο *Taylor*, βαθμού $m - 1$, της h , είναι το πολυώνυμο $T_{m-1}(x) = T_{m-1}(x; h)$, για το οποίο ισχύει $T_{m-1}^{(n)}(0) = h^{(n)}(0)$, $n = 0, 1, \dots, m - 1$, δηλ., το πολυώνυμο του οποίου η τιμή στο 0, καθώς και οι τιμές των πρώτων k παραγώγων του στο 0, συμπίπτουν με τις αντίστοιχες τιμές της h . Εύκολα αναγνωρίζεται ότι το πολυώνυμο *Taylor*, βαθμού $m - 1$, που αντιστοιχεί στην h , είναι το $T_{m-1}(x; h) = \sum_{j=0}^{m-1} h^{(j)}(0) x^j / j!$.

Ένα αποτέλεσμα όσον αφορά την επέκταση μίας επαρκούς ομαλής συνάρτησης, γύρω από ένα σημείο, είναι το Θεώρημα *Taylor* το οποίο παραθέτουμε διότι χρειάζεται σε ορισμένες αποδείξεις των μεθόδων Δέλτα.

1.1.4 Μονοδιάστατο Θεώρημα *Taylor*

Θεώρημα 1.2: (μονοδιάστατη περίπτωση *Taylor*)

Έστω πραγματική συνάρτηση $h: (-\alpha, \alpha) \rightarrow \mathcal{R}$ η οποία έχει m συνεχείς παραγώγους $h', h'', \dots, h^{(m)}$ τότε για κάθε $x \in (-\alpha, \alpha)$ υπάρχει $\xi = \xi(x) \in [-|x|, |x|]$ τέτοιο ώστε

$$\begin{aligned} h(x) &= P_{m-1}(x) + \frac{h^{(m)}(\xi)}{m!} x^m \\ &= h(0) + \frac{h'(0)}{1!} x^1 + \frac{h''(0)}{2!} x^2 + \dots + \frac{h^{(m-1)}(0)}{(m-1)!} x^{m-1} + \frac{h^{(m)}(\xi)}{m!} x^m. \end{aligned} \quad (1.3)$$

Επίσης, ισχύει

$$\lim_{x \rightarrow 0} \frac{h(x) - h(0) - \frac{h'(0)}{1!} x^1 - \dots - \frac{h^{(m)}(0)}{m!} x^m}{x^m} = 0. \quad (1.4)$$

Ο όρος $R_{m-1}(x) = \frac{h^{(m)}(\xi)}{m!} x^m$ λέγεται υπόλοιπο *Taylor* τάξεως $m-1$, κέντρου 0 για την συνάρτηση h . Άρα, το Θεώρημα *Taylor* γράφεται ως ακολούθως,

$$h(x) = P_{m-1}(x) + R_{m-1}(x),$$

όπου $P_{m-1}(x)$ είναι το πολυώνυμο *Taylor* τάξεως $m-1$, κέντρου 0 για την συνάρτηση h και $R_{m-1}(x)$ είναι το αντίστοιχο υπόλοιπο *Taylor*. Σημειώνουμε ότι ο όρος $h^{(m)}(\xi)$ που εμφανίζεται στο υπόλοιπο *Taylor* είναι συνάρτηση του x και όχι μία σταθερά, διαφορετικά κάθε συνάρτηση θα γραφόταν τοπικά ως πολυώνυμο τάξεως m , για οποιοδήποτε $m = 1, 2, \dots$.

Το παραπάνω Θεώρημα δίνει μια έκφραση της διαφοράς $h(x) - P_{m-1}(x)$. Στην περίπτωση που έχουμε μια πολύπλοκη συνάρτηση $h(x)$ μπορούμε να την προσεγγίσουμε με το πολυώνυμο *Taylor* $P_{m-1}(x)$ που είναι μία απλούστερη συνάρτηση και επομένως, το Θεώρημα *Taylor* δίνει μια έκφραση για το σφάλμα προσέγγισης.

Απόδειξη: Υποθέτουμε, χωρίς βλάβη της γενικότητας, ότι $I = [0, \alpha)$ και $x > 0$, έτσι, $x \in [0, \alpha)$. Στο προηγούμενο Θεώρημα αποδείξαμε τις σχέσεις

$$h(x) - \sum_{j=0}^{m-1} \frac{h^{(j)}(0)}{j!} x^j = \frac{1}{(m-1)!} \int_0^x (x-y)^{m-1} h^{(m)}(y) dy, \quad x \in I,$$

$$h(x) - \sum_{j=0}^m \frac{h^{(j)}(0)}{j!} x^j = \frac{1}{(m-1)!} \int_0^x (x-y)^{m-1} [h^{(m)}(y) - h^{(m)}(0)] dy, \quad x \in I.$$

Η $h^{(m)}(y)$ αναπαριστά την τιμή της m -οστής παραγώγου της h στο σημείο y , και είναι μία συνεχής συνάρτηση, οπότε

$$h^{(m)}(\xi_1) \leq h^{(m)}(y) \leq h^{(m)}(\xi_2)$$

για κάποια $\xi_1, \xi_2 \in [0, x]$.

Επιπλέον, θα έχουμε

$$\frac{h^{(m)}(\xi_1)}{m!} x^m \leq \frac{1}{(m-1)!} \int_0^x (x-y)^{m-1} h^{(m)}(y) dy \leq \frac{h^{(m)}(\xi_2)}{m!} x^m$$

συνεπώς, υπάρχει $\xi \in [\xi_1, \xi_2]$ έτσι ώστε

$$\frac{1}{(m-1)!} \int_0^x (x-y)^{m-1} h^{(m)}(y) dy = \frac{h^{(m)}(\xi)}{m!} x^m.$$

Τώρα, η σχέση (1.3) γράφεται και ως

$$\begin{aligned} h(x) &= \sum_{j=0}^{m-1} \frac{h^{(j)}(0)}{j!} x^j + \frac{h^{(m)}(\xi)}{m!} x^m = \sum_{j=0}^m \frac{h^{(j)}(0)}{j!} x^j + \frac{h^{(m)}(\xi) - h^{(m)}(0)}{m!} x^m \Rightarrow \\ &\Rightarrow h(x) - \sum_{j=0}^m \frac{h^{(j)}(0)}{j!} x^j = \frac{h^{(m)}(\xi) - h^{(m)}(0)}{m!} x^m, \quad \xi \in [0, x] \end{aligned}$$

άρα, καθώς $x \rightarrow 0$, ισχύει ότι

$$\frac{h(x) - \sum_{j=0}^m \frac{h^{(j)}(0)}{j!} x^j}{x^m} = \frac{1}{m!} [h^{(m)}(\xi) - h^{(m)}(0)] \rightarrow 0,$$

διότι

$$|\xi(x)| \leq |x| \rightarrow 0 \text{ και } h^{(m)}(\xi) - h^{(m)}(0) \rightarrow 0$$

λόγω συνεχείας της $h^{(m)}(x)$.

Επομένως,

$$\lim_{x \rightarrow 0} \frac{h(x) - h(0) - \frac{h'(0)}{1!}x^1 - \dots - \frac{h^{(m)}(0)}{m!}x^m}{x^m} = 0.$$

■

1.1.5 Πολυδιάστατο Θεώρημα Taylor

Θεώρημα 1.3: (πολυδιάστατη περίπτωση Taylor)

Έστω $D \subseteq \mathcal{R}^k$ είναι ένα ανοικτό σύνολο που περιλαμβάνει τον δίσκο $D(\mu, r)$, κέντρου $\mu = (\mu_1, \mu_2, \dots, \mu_k)' \in \mathcal{R}^k$ και ακτίνας $r > 0$. Ας είναι j_i το πλήθος των φορών που παραγωγίζουμε ως προς την μεταβλητή x_i για κάθε $i = 1, 2, \dots, k$. Αν $g: D \rightarrow \mathcal{R}$ είναι μία συνάρτηση με συνεχείς μερικές παραγώγους μέχρι m τάξης, δηλαδή, οι συναρτήσεις

$$g_{j_1, j_2, \dots, j_k}(x_1, x_2, \dots, x_k) = \frac{\partial^{j_1+j_2+\dots+j_k} g(\mathbf{x})}{\partial x_1^{j_1} \partial x_2^{j_2} \dots \partial x_k^{j_k}}, \quad \mathbf{x} = (x_1, x_2, \dots, x_k)' \in D$$

είναι συνεχείς για κάθε $j_1, j_2, \dots, j_k \in \{0, 1, \dots\}$ με $j_1 + j_2 + \dots + j_k \leq m$, τότε

$$\lim_{\mathbf{x} \rightarrow \mu} \frac{g(\mathbf{x}) - g(\mu) - \frac{D^1 g(\mu)}{1!}(\mathbf{x} - \mu)^1 - \dots - \frac{D^m g(\mu)}{m!}(\mathbf{x} - \mu)^m}{\sum_{i=1}^k |x_i - \mu_i|^m} = 0,$$

όπου

$$\begin{aligned} g^{(j)}(\mu)(\mathbf{x} - \mu)^j &\equiv D^j g(\mu)(\mathbf{x} - \mu)^j \\ &= \sum_{j_1 + \dots + j_k = j} g_{j_1, \dots, j_k}(\mu)(x_1 - \mu_1)^{j_1} \dots (x_k - \mu_k)^{j_k}, \quad j = 1, 2, \dots, m. \end{aligned}$$

Πριν προχωρήσουμε στην απόδειξη αυτού του Θεωρήματος και για να γίνει σαφές η ποσότητα $D^j g(\mu)(\mathbf{x} - \mu)^j$, αναλύουμε τις δύο πρώτες περιπτώσεις.

Έχουμε λοιπόν, για $j = 1$:

$$\begin{aligned} g^{(1)}(\mu)(\mathbf{x} - \mu)^1 &\equiv D^1 g(\mu)(\mathbf{x} - \mu)^1 \\ &= \sum_{j_1 + \dots + j_k = 1} g_{j_1, \dots, j_k}(\mu)(x_1 - \mu_1)^{j_1} \dots (x_k - \mu_k)^{j_k} \end{aligned}$$

$$= g_{(1)}(\boldsymbol{\mu})(x_1 - \mu_1) + \cdots + g_{(k)}(\boldsymbol{\mu})(x_k - \mu_k) = \sum_{i=1}^k g_{(i)}(\boldsymbol{\mu})(x_i - \mu_i)$$

$$= [\nabla g(\boldsymbol{\mu})]^T(\mathbf{x} - \boldsymbol{\mu}), \text{ όπου } g_{(i)} \text{ η μερική παράγωγος ως προς } x_i,$$

δηλαδή, $g_{(i)} = g_{0,0,\dots,0,1,0,\dots,0}$ όπου η μονάδα βρίσκεται στην i -οστή θέση.

Ομοίως, για $j = 2$ ισχύει

$$g^{(2)}(\boldsymbol{\mu})(\mathbf{x} - \boldsymbol{\mu})^2 \equiv D^2 g(\boldsymbol{\mu})(\mathbf{x} - \boldsymbol{\mu})^2$$

$$= \sum_{j_1 + \dots + j_k = 2} g_{j_1, \dots, j_k}(\boldsymbol{\mu})(x_1 - \mu_1)^{j_1} \cdots (x_k - \mu_k)^{j_k}$$

$$= \sum_{i=1}^k g_{(ii)}(\boldsymbol{\mu})(x_i - \mu_i)^2 + 2 \sum_{i < j} g_{(ij)}(\boldsymbol{\mu})(x_i - \mu_i)(x_j - \mu_j)$$

$$= (\mathbf{x} - \boldsymbol{\mu})^T H(\boldsymbol{\mu})(\mathbf{x} - \boldsymbol{\mu}), \text{ όπου } H(\boldsymbol{\mu}) = (g_{(ij)}(\boldsymbol{\mu}))_{k \times k} \text{ με } g_{(ij)} \text{ να αναπαριστά την δεύτερη μερική παράγωγο ως προς } x_i, x_j \text{ κ.ο.κ.}$$

Απόδειξη: Θέτουμε,

$$h(t) = g(\boldsymbol{\mu} + t(\mathbf{x} - \boldsymbol{\mu})) = g(t\mathbf{x} + (1-t)\boldsymbol{\mu}), \quad 0 \leq t \leq 1.$$

Παρατηρούμε ότι

$$h(0) = g(\boldsymbol{\mu}) \text{ και } h(1) = g(\mathbf{x}).$$

Συνεπώς, $\exists \xi \in [0,1]$ τέτοιο ώστε

$$h(1) = h(0) + \frac{h'(0)}{1!} + \frac{h''(0)}{2!} + \cdots + \frac{h^{(m-1)}(0)}{(m-1)!} + \frac{h^{(m)}(\xi)}{m!}. \quad (1.5)$$

Επίσης, από τον κανόνα της αλυσίδας,

$$h'(t) \equiv h^{(1)}(t) = \sum_{i=1}^k g_{(i)}(\boldsymbol{\mu} + t(\mathbf{x} - \boldsymbol{\mu}))(x_i - \mu_i) \quad [h'(0) = [\nabla g(\boldsymbol{\mu})]^T(\mathbf{x} - \boldsymbol{\mu})].$$

Στη συνέχεια, θα δείξουμε με επαγωγή ότι

$$\begin{aligned} h^{(j)}(t) &= [D^j g(\boldsymbol{\mu} + t(\mathbf{x} - \boldsymbol{\mu}))](\mathbf{x} - \boldsymbol{\mu})^j \\ &= \sum_{j_1 + \dots + j_k = j} g_{j_1, \dots, j_k}(\boldsymbol{\mu} + t(\mathbf{x} - \boldsymbol{\mu}))(x_1 - \mu_1)^{j_1} \cdots (x_k - \mu_k)^{j_k}, \quad j = 1, 2, \dots, m. \end{aligned}$$

Έστω ότι

$$h^{(j)}(t) = \sum_{j_1+\dots+j_k=j} g_{j_1,\dots,j_k}(\boldsymbol{\mu} + t(\mathbf{x} - \boldsymbol{\mu}))(x_1 - \mu_1)^{j_1} \dots (x_k - \mu_k)^{j_k}$$

οπότε,

$$h^{(j)}(0) = \sum_{j_1+\dots+j_k=j} g_{j_1,\dots,j_k}(\boldsymbol{\mu})(x_1 - \mu_1)^{j_1} \dots (x_k - \mu_k)^{j_k} = D^j g(\boldsymbol{\mu})(\mathbf{x} - \boldsymbol{\mu})^j.$$

Τότε,

$$\begin{aligned} h^{(j+1)}(t) &= \sum_{j_1+\dots+j_k=j} [g_{j_1,\dots,j_k}(\boldsymbol{\mu} + t(\mathbf{x} - \boldsymbol{\mu}))]'(x_1 - \mu_1)^{j_1} \dots (x_k - \mu_k)^{j_k} \\ &= \sum_{s=1}^k (x_s - \mu_s) \sum_{j_1+\dots+j_k=j} g_{j_1,\dots,j_s+1,\dots,j_k}(\boldsymbol{\mu} + t(\mathbf{x} - \boldsymbol{\mu}))(x_1 - \mu_1)^{j_1} \dots (x_k - \mu_k)^{j_k} \\ &= D^{j+1} g(\boldsymbol{\mu} + t(\mathbf{x} - \boldsymbol{\mu}))(\mathbf{x} - \boldsymbol{\mu})^{j+1}. \end{aligned}$$

Έτσι, από την σχέση (1.5) προκύπτει ότι

$$\begin{aligned} g(\mathbf{x}) &= g(\boldsymbol{\mu}) + \frac{D^1 g(\boldsymbol{\mu})}{1!}(\mathbf{x} - \boldsymbol{\mu})^1 + \dots + \frac{D^{m-1} g(\boldsymbol{\mu})}{(m-1)!}(\mathbf{x} - \boldsymbol{\mu})^{m-1} \\ &\quad + \frac{D^m g(\boldsymbol{\mu} + \xi(\mathbf{x} - \boldsymbol{\mu}))}{m!}(\mathbf{x} - \boldsymbol{\mu})^m \\ &= \sum_{j=0}^{m-1} \frac{D^j g(\boldsymbol{\mu})}{j!}(\mathbf{x} - \boldsymbol{\mu})^j + \frac{D^m g(\boldsymbol{\mu} + \xi(\mathbf{x} - \boldsymbol{\mu}))}{m!}(\mathbf{x} - \boldsymbol{\mu})^m \\ &= \sum_{j=0}^m \frac{D^j g(\boldsymbol{\mu})}{j!}(\mathbf{x} - \boldsymbol{\mu})^j + \frac{1}{m!} \{D^m g(\boldsymbol{\mu} + \xi(\mathbf{x} - \boldsymbol{\mu})) - D^m g(\boldsymbol{\mu})\}(\mathbf{x} - \boldsymbol{\mu})^m \\ &\Rightarrow g(\mathbf{x}) - \sum_{j=0}^m \frac{1}{j!} [D^j g(\boldsymbol{\mu})(\mathbf{x} - \boldsymbol{\mu})^j] \\ &= \frac{1}{m!} \sum_{j_1+\dots+j_k=m} \{g_{j_1,\dots,j_k}(\boldsymbol{\mu} + \xi(\mathbf{x} - \boldsymbol{\mu})) - g_{j_1,\dots,j_k}(\boldsymbol{\mu})\}(x_1 - \mu_1)^{j_1} \dots (x_k - \mu_k)^{j_k} \\ &\Rightarrow m! \cdot \frac{\left| g(\mathbf{x}) - \sum_{j=0}^m \frac{1}{j!} [D^j g(\boldsymbol{\mu})(\mathbf{x} - \boldsymbol{\mu})^j] \right|}{\sum_{i=1}^k |x_i - \mu_i|^m} \end{aligned}$$

$$\leq \frac{\sum_{j_1+\dots+j_k=m} |x_1 - \mu_1|^{j_1} \dots |x_k - \mu_k|^{j_k}}{\sum_{i=1}^k |x_i - \mu_i|^m} \cdot \max_{j_1, j_2, \dots, j_k, \xi} |g_{j_1, \dots, j_k}(\mu + \xi(x - \mu)) - g_{j_1, \dots, j_k}(\mu)|.$$

Όμως, $\xi \in [0,1]$, οπότε

$$\max_{j_1, j_2, \dots, j_k, \xi} |g_{j_1, \dots, j_k}(\mu + \xi(x - \mu)) - g_{j_1, \dots, j_k}(\mu)| \leq \text{πεπερασμένο άθροισμα από } |g_{j_1, \dots, j_k}(\mu + \xi(x - \mu)) - g_{j_1, \dots, j_k}(\mu)| \rightarrow 0.$$

Τέλος, το πηλίκο

$$\frac{\sum_{j_1+\dots+j_k=m} |x_1 - \mu_1|^{j_1} \dots |x_k - \mu_k|^{j_k}}{\sum_{i=1}^k |x_i - \mu_i|^m}$$

είναι φραγμένο διότι ο αριθμητής περιέχει $\begin{bmatrix} k \\ m \end{bmatrix}$ προσθετέους της μορφής $|x_1 - \mu_1|^{j_1} \dots |x_k - \mu_k|^{j_k}$.

Συγκεκριμένα, για τυχόντα προσθετέο έχουμε

$$\begin{aligned} |x_1 - \mu_1|^{j_1} \dots |x_k - \mu_k|^{j_k} &\leq (\max\{|x_1 - \mu_1|, \dots, |x_k - \mu_k|\})^{j_1} \\ &\quad \cdot (\max\{|x_1 - \mu_1|, \dots, |x_k - \mu_k|\})^{j_2} \dots (\max\{|x_1 - \mu_1|, \dots, |x_k - \mu_k|\})^{j_k} \\ &= (\max\{|x_1 - \mu_1|, \dots, |x_k - \mu_k|\})^{j_1+j_2+\dots+j_k} \\ &= (\max\{|x_1 - \mu_1|, \dots, |x_k - \mu_k|\})^m \\ &= \max\{|x_1 - \mu_1|^m, \dots, |x_k - \mu_k|^m\} \\ &\leq \sum_{i=1}^k |x_i - \mu_i|^m. \end{aligned}$$

Δηλαδή, για κάθε επιλογή δεικτών $j_1, j_2, \dots, j_k \geq 0$ με $j_1 + \dots + j_k = m$,

$$\prod_{s=1}^k |x_s - \mu_s|^{j_s} \leq \sum_{i=1}^k |x_i - \mu_i|^m.$$

Οπότε,

$$\frac{\sum_{j_1+\dots+j_k=m} |x_1 - \mu_1|^{j_1} \dots |x_k - \mu_k|^{j_k}}{\sum_{i=1}^k |x_i - \mu_i|^m} \leq \begin{bmatrix} k \\ m \end{bmatrix} = \binom{k+m-1}{m},$$

και έτσι, εφόσον, το όριο μηδενικής συνάρτησης επί φραγμένης συνάρτησης ισούται με μηδέν, συμπεραίνουμε ότι

$$\lim_{x \rightarrow \mu} \frac{g(x) - g(\mu) - \frac{D^1 g(\mu)}{1!}(x - \mu)^1 - \dots - \frac{D^m g(\mu)}{m!}(x - \mu)^m}{\sum_{i=1}^k |x_i - \mu_i|^m} = 0.$$

■

1.2 Είδη στοχαστικών συγκλίσεων

Εν συνεχεία, θα αναφερθούμε στα τέσσερα είδη στοχαστικών συγκλίσεων που άπτονται των πιθανοτήτων και είναι ιδιαιτέρως χρήσιμα σε ότι ακολουθεί.

1. Σύγκλιση κατά πιθανότητα: Έστω οι τυχαίες μεταβλητές $X, X_1, X_2, \dots$ που είναι ορισμένες στον ίδιο χώρο πιθανότητας $(\Omega, \mathcal{A}, \mathcal{P})$. Η ακολουθία $\{X_n\}_{n \geq 1}$ συγκλίνει κατά πιθανότητα στην τυχαία μεταβλητή X ($X_n \xrightarrow{p} X$) αν και μόνο αν

$$\forall \varepsilon > 0, P(|X_n - X| \geq \varepsilon) \rightarrow 0, \text{ καθώς } n \rightarrow \infty$$

ή ισοδύναμα $X_n \xrightarrow{p} X$ όταν και μόνο όταν

$$\forall \varepsilon > 0, P(|X_n - X| < \varepsilon) \rightarrow 1, \text{ καθώς } n \rightarrow \infty.$$

2. Σύγκλιση με πιθανότητα 1 (ή σχεδόν βέβαιη σύγκλιση ή σχεδόν παντού σύγκλιση ή ισχυρή σύγκλιση): Έστω $X, X_1, X_2, \dots$ τυχαίες μεταβλητές ορισμένες στον ίδιο χώρο πιθανότητας $(\Omega, \mathcal{A}, \mathcal{P})$. Η ακολουθία $\{X_n\}_{n \geq 1}$ συγκλίνει με πιθανότητα 1 ή σχεδόν παντού ή ισχυρά ή σχεδόν βέβαια στην τυχαία μεταβλητή X ($X_n \xrightarrow{a.s.} X$) όταν και μόνο όταν

$$P\left(\lim_{n \rightarrow \infty} X_n = X\right) = 1$$

ή ισοδύναμα $X_n \xrightarrow{a.s.} X$ αν και μόνο αν

$$\lim_{n \rightarrow \infty} X_n(\omega) = X(\omega), \quad \forall \omega \in A \in \mathcal{A} \text{ με } P(A^c) = 0.$$

Η σχεδόν παντού σύγκλιση είναι κατά σημείο σύγκλιση εκτός συνόλου μέτρου μηδέν και εναλλακτικά, μπορεί ο ορισμός της να δοθεί και από το κάτωθι Λήμμα.

Λήμμα 1.1: $X_n \xrightarrow{a.s.} X \Leftrightarrow \forall \varepsilon > 0, \lim_{n \rightarrow \infty} P(\sup_{k \geq n} |X_k - X| > \varepsilon) = 0$

δηλαδή, εάν και μόνο εάν

$$\sup_{k \geq n} |X_k - X| \xrightarrow{p} 0.$$

Απόδειξη: Καταρχάς σημειώνουμε ότι η τυχαία μεταβλητή X είναι μη εκτεταμένη, δηλ., $-\infty < X(\omega) < +\infty$ για κάθε $\omega \in \Omega$. Γίνεται αντιληπτό ότι $X_n(\omega) \rightarrow X(\omega)$ αν και μόνο αν για κάθε $m = 1, 2, \dots$, $|X_n(\omega) - X(\omega)| \leq \frac{1}{m}$ τελικά για κάθε n . Στην "γλώσσα" των ενδεχομένων, αυτό γράφεται ως εξής:

$$\begin{aligned} \{\omega: X_n(\omega) \rightarrow X(\omega)\} &= \bigcap_{m=1}^{\infty} \lim_{n \rightarrow +\infty} \inf \left\{ \omega: |X_n - X| \leq \frac{1}{m} \right\} \\ &= \bigcap_{m=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} \left\{ \omega: |X_k - X| \leq \frac{1}{m} \right\}. \end{aligned}$$

Έτσι,

$$\{\omega: X_n(\omega) \nrightarrow X(\omega)\} = \{\omega: X_n(\omega) \nrightarrow X(\omega)\}^c = \bigcup_{m=1}^{\infty} \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} \left\{ \omega: |X_k - X| > \frac{1}{m} \right\}$$

Παρατηρούμε ότι η ακολουθία ενδεχομένων

$$B_n^m = \bigcup_{k=n}^{\infty} \left\{ \omega: |X_k - X| > \frac{1}{m} \right\} = \left\{ \sup_{k \geq n} |X_k - X| > \frac{1}{m} \right\}$$

είναι φθίνουσα ως προς n , με

$$\lim_{n \rightarrow +\infty} B_n^m = B^m = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} \left\{ |X_k - X| > \frac{1}{m} \right\}.$$

[Σημειώνουμε ότι $B^m = \lim_{n \rightarrow +\infty} \sup_n \{|X_n - X| > 1/m\}$.]

Επομένως,

$$\{\omega: X_n(\omega) \nrightarrow X(\omega)\} = \bigcup_{m=1}^{\infty} B^m.$$

Τώρα, αν $X_n \xrightarrow{a.s.} X$, τότε θα έχουμε ότι $P(X_n \nrightarrow X) = 0$ και συνεπώς,

$$P(B^m) \leq P\left(\bigcup_{m=1}^{\infty} B^m\right) = P(X_n \not\rightarrow X) = 0,$$

δηλ., $P(B^m) = 0$ για κάθε m . Οπότε,

$$\lim_{n \rightarrow +\infty} P\left(\sup_{k \geq n} |X_k - X| > \frac{1}{m}\right) = \lim_{n \rightarrow +\infty} P(B_n^m) = P\left(\lim_{n \rightarrow +\infty} B_n^m\right) = P(B^m) = 0$$

που σημαίνει ότι $P(\sup_{k \geq n} |X_k - X| > \varepsilon) \rightarrow 0$ (καθώς $n \rightarrow \infty$) για κάθε $\varepsilon > 0$, διότι για τυχόν $\varepsilon > 0$, μπορούμε να εκλέξουμε $m = \left\lceil \frac{1}{\varepsilon} \right\rceil + 1 > \frac{1}{\varepsilon}$, έτσι ώστε

$$\{\sup_{k \geq n} |X_k - X| > \varepsilon\} \subset \left\{\sup_{k \geq n} |X_k - X| > \frac{1}{m}\right\} = B_n^m.$$

Αντίστροφα, ας υποθέσουμε ότι $P(\sup_{k \geq n} |X_k - X| > \varepsilon) \rightarrow 0$ για κάθε $\varepsilon > 0$. Θέτοντας, $\varepsilon = \frac{1}{m}$, προκύπτει ότι $P(\sup_{k \geq n} |X_k - X| > \varepsilon) \rightarrow 0$, δηλ. $P(B_n^m) \xrightarrow{n \rightarrow \infty} 0$. Αυτό σημαίνει ότι $P(B^m) = P(\lim_{n \rightarrow +\infty} B_n^m) = \lim_{n \rightarrow +\infty} P(B_n^m) = 0$, και άρα,

$$P\left(\bigcup_{m=1}^{\infty} B^m\right) = 0$$

ήτοι, $P(X_n \not\rightarrow X) = 0$, τουτέστιν, $X_n \xrightarrow{a.s.} X$. ■

3. Σύγκλιση στον L^p : Έστω οι τυχαίες μεταβλητές $X, X_1, X_2, \dots$ οι οποίες ορίζονται στον ίδιο χώρο πιθανότητας $(\Omega, \mathcal{A}, \mathcal{P})$. Η ακολουθία $\{X_n\}_{n \geq 1}$ συγκλίνει στην τυχαία μεταβλητή X στον L^p , $\forall p \geq 1$ ($X_n \xrightarrow{L^p} X$) όταν $E|X|^p < \infty$, $E|X_n|^p < \infty$ από κάποιο n και πάνω (δηλαδή τελικά για κάθε n) και ισχύει

$$E|X_n - X|^p \rightarrow 0, \text{ καθώς } n \rightarrow \infty.$$

Οι πιο ενδιαφέρουσες περιπτώσεις προκύπτουν για $p = 1$ και $p = 2$:

- $p = 1$: $X_n \xrightarrow{L^1} X \Leftrightarrow \overset{o.p.}{E|X_n - X|} \rightarrow 0$. (σύγκλιση στον L^1)
- $p = 2$: $X_n \xrightarrow{L^2} X \Leftrightarrow \overset{o.p.}{E(X_n - X)^2} \rightarrow 0$. (σύγκλιση στον L^2 ή σύγκλιση κατά μέσο τετράγωνο)

4. Σύγκλιση κατά κατανομή (ή κατά νόμο ή ασθενής σύγκλιση): Έστω X τυχαία μεταβλητή ορισμένη σε αυθαίρετο χώρο πιθανότητας $(\Omega, \mathcal{A}, \mathcal{P})$ και $\{X_n\}_{n \geq 1}$ ακολουθία τυχαίων μεταβλητών ορισμένη σε αυθαίρετη ακολουθία χώρων πιθανότητας $(\Omega_n, \mathcal{A}_n, \mathcal{P}_n)$. Η $\{X_n\}$ συγκλίνει κατά κατανομή ή κατά νόμο ή ασθενώς στην τυχαία μεταβλητή X ($X_n \xrightarrow{d} X$) αν και μόνο αν

$$\lim_{n \rightarrow \infty} P_n(X_n \leq x) = P(X \leq x), \quad \forall x: P(X = x) = 0,$$

δηλαδή, για κάθε σημείο συνεχείας της αθροιστικής συνάρτησης κατανομής $F(x) = P(X \leq x)$.

1.3 Θεωρία Μέτρου

1.3.1 Εισαγωγή

Η Θεωρία Μέτρου στα μαθηματικά περιλαμβάνει την αυστηρή αξιωματική θεμελίωση, καθώς επίσης, και τη γενίκευση των εννοιών του "μήκους", του "εμβαδού" και του "όγκου". Οι εφαρμογές επεκτείνονται και σε άλλες, πέρα από τις γεωμετρικές, έννοιες του "μέτρου". Σε απλουστευμένη αλλά αδόκιμη ορολογία, για δεδομένο σύνολο S , ονομάζουμε "μέτρο" οποιαδήποτε διαδικασία ή κανόνα αποδίδει ένα "μέγεθος" σε κάθε ένα υποσύνολο του S κατά συνεπή (δηλ., χωρίς αντιφάσεις) τρόπο.

Ο ακριβής μαθηματικός ορισμός του όρου "μέτρο" μοιάζει δυσνόητος (προϋποθέτοντας έννοιες όπως "δ-δακτύλιος" και "σ-άλγεβρα") αλλά είναι εξαιρετικά σημαντικός για τη θεμελίωση, μεταξύ άλλων, βασικών εννοιών της ανάλυσης και του απειροστικού λογισμού. Σημαντικά παραδείγματα είναι το "μέτρο Jordan²⁰", το "μέτρο Lebesgue²¹", το "μέτρο πιθανότητας" και άλλα.

Ο "χώρος μέτρου" ("measure space") δεν θα πρέπει να συγχέεται με το "μετρικό χώρο" ("metric space"). Οι δύο έννοιες είναι διαφορετικές. Για παράδειγμα, στον "ευκλείδειο χώρο" τριών διαστάσεων το "μέτρο Lebesgue" ορίζει το "χώρο μέτρου" κατά τον οποίο αποδίδεται "όγκος" στα υποσύνολα του $\mathcal{R}^3$, ενώ η "ευκλείδεια μετρική" ορίζει το "μετρικό χώρο" που "μετράει" την "απόσταση" μεταξύ δύο σημείων του $\mathcal{R}^3$.

Ορισμός 1.1: (εξωτερικό μέτρο)

Έστω τυχόν, μη κενό, σύνολο Ω και μία μη-αρνητική συνολοσυνάρτηση $\varphi: \mathcal{P}(\Omega) \rightarrow [0, +\infty]$ (όπου $\mathcal{P}(\Omega)$ είναι το δυναμοσύνολο του Ω το οποίο περιλαμβάνει το $\emptyset$, το Ω καθώς και όλα τα υποσύνολα του Ω). Η φ καλείται εξωτερικό μέτρο εάν πληροί τις επόμενες ιδιότητες:

- (i) $\varphi(\emptyset) = 0$.

²⁰Ο Marie Ennemond Camille Jordan (5 Ιανουαρίου 1838 – 22 Ιανουαρίου 1922) ήταν Γάλλος μαθηματικός, γνωστός τόσο για το θεμελιώδες έργο του στην θεωρία ομάδων όσο και για την επιρροή του στην ανάλυση. Επίσης, ασχολήθηκε με την θεωρία Galois και τις ομάδες Mathieu.

²¹Ο Henri Léon Lebesgue (28 Ιουνίου 1875 - 26 Ιουλίου 1941) υπήρξε Γάλλος μαθηματικός πολύ γνωστός για την ενασχόλησή του με τη Θεωρία Ολοκλήρωσης και Θεμελιωτής του Ολοκληρώματος, όπως αυτό χρησιμοποιείται σήμερα.

- (ii) Για κάθε $A, B \in \mathcal{P}(\Omega)$ με $A \subseteq B \Rightarrow \varphi(A) \leq \varphi(B)$ (δηλαδή η φ είναι μονότονη).
- (iii) Για τυχούσα ακολουθία $\{A_n, n \geq 1\} \in \mathcal{P}(\Omega) \Rightarrow \varphi(\bigcup_{n=1}^{\infty} A_n) \leq \sum_{n=1}^{\infty} \varphi(A_n)$
(δηλαδή η φ είναι αριθμήσιμα υποπροσθετική ή σ -υποπροσθετική).

Ορισμός 1.2: (εξωτερικό μέτρο του Lebesgue στον $(0,1)$)

Το εξωτερικό μέτρο του Lebesgue, στο διάστημα $(0,1)$, είναι η συνολοσυνάρτηση $\lambda^*: \mathcal{P}(\Omega) \rightarrow [0,1]$ που ορίζεται ως

$$\lambda^*(A) = \inf \left\{ \sum_{n=1}^{\infty} (b_n - a_n), A \subset \bigcup_{n=1}^{\infty} (a_n, b_n), 0 \leq a_n \leq b_n \leq 1, n = 1, 2, \dots \right\}$$

για κάθε $A \subseteq (0,1)$.

Ορισμός 1.3: (σ -άλγεβρα)

Έστω Ω τυχόν, μη κενό, σύνολο. Μία οικογένεια $\mathcal{A}$, υποσυνόλων του Ω (δηλ. $\mathcal{A} \subset \mathcal{P}(\Omega)$), καλείται σ -άλγεβρα στον Ω (ή σ -άλγεβρα επί του Ω ή, απλά, σ -άλγεβρα, όταν το Ω εννοείται), όταν η $\mathcal{A}$ περιέχει το $\emptyset$, και είναι κλειστή ως προς τα συμπληρώματα και τις αριθμήσιμες ενώσεις, δηλ. όταν:

- (i) $\emptyset \in \mathcal{A}$.
- (ii) Αν $A \in \mathcal{A}$ τότε $A^c \in \mathcal{A}$. (κλειστότητα ως προς τα συμπληρώματα)
- (iii) Αν η $\{A_n, n \geq 1\}$ είναι τυχούσα ακολουθία στην $\mathcal{A}$, τότε $A = \bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$.
(κλειστότητα ως προς τις αριθμήσιμες ενώσεις, ή σ -κλειστότητα)

Είναι φανερό ότι η απαίτηση "η $\mathcal{A}$ περιέχει το $\emptyset$ " ισοδυναμεί με την απαίτηση "η $\mathcal{A}$ περιέχει τον Ω ", αφού $\emptyset^c = \Omega$. Επιπλέον, η κλειστότητα ως προς τις αριθμήσιμες ενώσεις μπορεί να αντικατασταθεί με την κλειστότητα ως προς τις αριθμήσιμες τομές, διότι

$$\bigcap_{n=1}^{\infty} A_n = \left(\bigcup_{n=1}^{\infty} A_n^c \right)^c.$$

Ορισμός 1.4: (χώρος πιθανότητας, Kolmogorov 1930)

Έστω Ω τυχόν, μη κενό, σύνολο, και $\mathcal{A}$ μία σ -άλγεβρα στον Ω . Οποιαδήποτε μη-αρνητική συνολοσυνάρτηση $\mathcal{P}$ θα ονομάζεται πιθανότητα εάν πληροί τις εξής ιδιότητες:

- (i) $\mathcal{P}: \mathcal{A} \rightarrow [0,1]$.
- (ii) $\mathcal{P}(\Omega) = 1$.
- (iii) Για τυχούσα ακολουθία $\{A_n, n \geq 1\} \in \mathcal{A}$ με $A_i \cap A_j = \emptyset$ για $i \neq j$ έπεται

$$\mathcal{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mathcal{P}(A_n)$$

(δηλαδή η $\mathcal{P}$ είναι αριθμήσιμα προσθετική ή σ -προσθετική).

Τα στοιχεία της $\mathcal{A}$ λέγονται ενδεχόμενα και η τριάδα $(\Omega, \mathcal{A}, \mathcal{P})$ καλείται χώρος πιθανότητας.

Ορισμός 1.5: (μετρήσιμος χώρος-μέτρο)

Έστω Ω τυχόν, μη κενό, σύνολο, και $\mathcal{A}$ τυχούσα σ -άλγεβρα επί του Ω . Το ζεύγος $(\Omega, \mathcal{A})$ καλείται μετρήσιμος χώρος. Αν η μη-αρνητική συνολοσυνάρτηση $\mu: \mathcal{A} \rightarrow (0, +\infty)$ ικανοποιεί τις ακόλουθες ιδιότητες:

- (i) $\mu(\emptyset) = 0$, και
- (ii) Για κάθε ακολουθία $\{A_n, n \geq 1\}$, ξένων ανά δύο συνόλων στην $\mathcal{A}$
(δηλαδή $A_i \cap A_j = \emptyset$ για $i \neq j$),

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n),$$

(δηλαδή η συνολοσυνάρτηση μ είναι σ -προσθετική)

τότε το μ καλείται μέτρο, η τριάδα $(\Omega, \mathcal{A}, \mu)$ καλείται χώρος μέτρου και τα σύνολα της $\mathcal{A}$ καλούνται μετρήσιμα (ή $\mathcal{A}$ -μετρήσιμα) σύνολα.

Ορισμός 1.6: (παραγόμενη σ -άλγεβρα)

Έστω Ω τυχόν, μη κενό, σύνολο, και $\mathfrak{D}$ τυχούσα, μη κενή, συλλογή (κλάση) υποσυνόλων του Ω . Τότε οποιαδήποτε σ -άλγεβρα $\mathcal{A}$, για την οποία ικανοποιούνται οι εξής δύο ιδιότητες:

(i) $\mathfrak{D} \subset \mathcal{A}$, και

(ii) Αν η $\mathcal{A}'$ είναι σ -άλγεβρα στον Ω και $\mathfrak{D} \subseteq \mathcal{A}'$ τότε $\mathcal{A} \subseteq \mathcal{A}'$,

ονομάζεται σ -άλγεβρα παραγόμενη από την $\mathfrak{D}$, και συμβολίζεται με $\sigma(\mathfrak{D})$.

Ορισμός 1.7: (σ -άλγεβρα του Borel σε τυχόντα μετρικό χώρο Ω)

Έστω (Ω, ρ) τυχόν μετρικός χώρος (ρ η μετρική του). Η σ -άλγεβρα που παράγεται από τα ανοικτά υποσύνολα του Ω ονομάζεται σ -άλγεβρα του Borel²² (στον Ω), και συμβολίζεται ως $\mathcal{B}(\Omega)$. Τα σύνολα που περιέχει η $\mathcal{B}(\Omega)$ ονομάζονται σύνολα Borel, ή Borel-μετρήσιμα σύνολα (του Ω).

Πρόταση 1.1: Εάν η $\{A_n, n \geq 1\}$ είναι μονότονη ακολουθία ενδεχομένων (σε χώρο πιθανότητας $(\Omega, \mathcal{A}, \mathcal{P})$), τότε

$$P\left(\lim_{n \rightarrow +\infty} A_n\right) = \lim_{n \rightarrow +\infty} P(A_n).$$

Απόδειξη: Γνωρίζουμε ότι το $\lim_{n \rightarrow +\infty} A_n$ υπάρχει, και δη,

$$\lim_{n \rightarrow +\infty} A_n = \begin{cases} \bigcup_{n=1}^{\infty} A_n, & \text{αν η } A_n \text{ είναι αύξουσας,} \\ \bigcap_{n=1}^{\infty} A_n, & \text{αν η } A_n \text{ είναι φθίνουσας,} \end{cases}$$

άρα, $\lim_{n \rightarrow +\infty} A_n \in \mathcal{A}$. Επίσης, επειδή η $\{P(A_n), n \geq 1\}$ είναι μονότονη (λόγω της μονοτονίας της P), και φραγμένη, έπεται ότι το $\lim_{n \rightarrow +\infty} P(A_n)$ υπάρχει, και μάλιστα,

$$0 \leq \lim_{n \rightarrow +\infty} P(A_n) \leq 1.$$

Συνεπώς, οι $P(\lim_{n \rightarrow +\infty} A_n)$ και $\lim_{n \rightarrow +\infty} P(A_n)$ είναι καλά ορισμένοι πραγματικοί αριθμοί στο $[0, 1]$ (όταν η $\{A_n, n \geq 1\}$ είναι μονότονη), και αρκεί να δείξουμε ότι είναι και ίσοι.

²²Ο Félix Édouard Justin Émile Borel (7 Ιανουαρίου 1871 - 3 Φεβρουαρίου 1956) υπήρξε Γάλλος μαθηματικός και πολιτικός γεννημένος στην κοινότητα Saint-Affrique του διαμερίσματος Aveyron στη νότια Γαλλία. Μαζί με τον René-Louis Baire και τον Henri Léon Lebesgue, ήταν ανάμεσα στους πρωτοπόρους της θεωρίας μέτρου και της θεωρίας πιθανοτήτων. Προς τιμήν του, η έννοια του συνόλου Borel φέρει το όνομά του. Μάλιστα, κατά τα έτη 1913–1914 με την εργασία του "expository", γεφύρωσε το χάσμα μεταξύ υπερβολικής γεωμετρίας και ειδικής θεωρίας της σχετικότητας.

Ας υποθέσουμε ότι η $\{A_n, n \geq 1\}$ είναι αύξουσα, τότε $A_j \subset A_{j+1}$ για $j \geq 1$.
Θέτοντας $B_1 = A_1$

$$B_2 = A_2 \setminus A_1$$

$$B_3 = A_3 \setminus A_2$$

$\vdots$

$\vdots$

$\vdots$

$$B_j = A_j \setminus A_{j-1} \quad \forall j = 2, 3, \dots$$

έχουμε ότι

$$B_i \cap B_j = \emptyset \text{ για } i \neq j,$$

επιπλέον, θα ισχύει

$$B_1 \cup B_2 \cup \dots \cup B_n = A_1 \cup A_2 \cup \dots \cup A_n = A_n,$$

καθώς και

$$\bigcup_{j=1}^{\infty} B_j = \bigcup_{j=1}^{\infty} A_j = \lim_{n \rightarrow +\infty} A_n.$$

Συνεπώς,

$$\begin{aligned} P\left(\lim_{n \rightarrow +\infty} A_n\right) &= P\left(\bigcup_{j=1}^{\infty} A_j\right) = P\left(\bigcup_{j=1}^{\infty} B_j\right) = \sum_{j=1}^{\infty} P(B_j) = \lim_{n \rightarrow +\infty} \sum_{j=1}^{\infty} P(B_j) \\ &= \lim_{n \rightarrow +\infty} P(B_1 \cup B_2 \cup \dots \cup B_n) = \lim_{n \rightarrow +\infty} P(A_n). \end{aligned}$$

Εάν η $\{A_n, n \geq 1\}$ είναι φθίνουσα, τότε η $\{A_n^c, n \geq 1\}$ είναι αύξουσα και λόγω των προηγούμενων,

$$P\left(\lim_{n \rightarrow +\infty} A_n^c\right) = \lim_{n \rightarrow +\infty} P(A_n^c).$$

Όμως,

$$\lim_{n \rightarrow +\infty} A_n^c = \bigcup_{n=1}^{\infty} A_n^c = (\bigcap_{n=1}^{\infty} A_n)^c = (\bigcap_{n=1}^{\infty} A_n^c)^c = (\lim_{n \rightarrow +\infty} A_n)^c$$

και

$$\lim_{n \rightarrow +\infty} P(A_n^c) = \lim_{n \rightarrow +\infty} [1 - P(A_n)] = 1 - \lim_{n \rightarrow +\infty} P(A_n).$$

Χρησιμοποιώντας τις τρεις τελευταίες σχέσεις, προκύπτει ότι

$$\begin{aligned} 1 - \lim_{n \rightarrow +\infty} P(A_n) &= \lim_{n \rightarrow +\infty} P(A_n^c) = P\left(\lim_{n \rightarrow +\infty} A_n^c\right) \\ &= P\left(\left(\lim_{n \rightarrow +\infty} A_n\right)^c\right) = 1 - P\left(\lim_{n \rightarrow +\infty} A_n\right) \end{aligned}$$

και το συμπέρασμα έπεται και για φθίνουσες ακολουθίες. ■

Πρόταση 1.2: Η X είναι τ.μ. στον $(\Omega, \mathcal{A})$ αν και μόνο αν $X^{-1}(B) \in \mathcal{A}$, για κάθε $B \in \mathcal{B}(\mathcal{R})$, δηλαδή, $X^{-1}(\mathcal{B}(\mathcal{R})) \subseteq \mathcal{A}$.

Απόδειξη: Καταρχάς $\mathcal{B}(\mathcal{R})$ είναι η σ -άλγεβρα του *Borel* στον $\mathcal{R}$. Ας υποθέσουμε πρώτα ότι

$$X^{-1}(B) \in \mathcal{A}, \text{ για κάθε } B \in \mathcal{B}(\mathcal{R}), \text{ δηλαδή, } X^{-1}(\mathcal{B}(\mathcal{R})) \subseteq \mathcal{A}.$$

Εφόσον, κάθε σύνολο B της μορφής $B = (-\infty, b]$ είναι, προφανώς, *Borel*, η προηγούμενη σχέση συνεπάγεται ότι $X^{-1}((-\infty, b]) \in \mathcal{A}$ για κάθε $b \in \mathcal{R}$, ήτοι, η X είναι τυχαία μεταβλητή.

Αντίστροφα, τώρα, υποθέτουμε ότι η X είναι τ.μ. και ας θεωρήσουμε την οικογένεια "καλών συνόλων"

$$\mathcal{B} = \{B \in \mathcal{B}(\mathcal{R}) : X^{-1}(B) \in \mathcal{A}\} \subseteq \mathcal{B}(\mathcal{R}).$$

Αφού η X είναι τ.μ., έπεται ότι $(-\infty, b] \in \mathcal{B}$ για κάθε $b \in \mathcal{R}$, δηλαδή, $\mathfrak{D}_1 \subset \mathcal{B}$, όπου $\mathfrak{D}_1 = \{(-\infty, x], x \in \mathcal{R}\}$.

Όμως, η $\mathcal{B}$ είναι σ -άλγεβρα στον $\mathcal{R}$, διότι:

(i) $\mathcal{R} \in \mathcal{B}$ (αφού $X^{-1}(\mathcal{R}) = \Omega \in \mathcal{A}$).

(ii) Αν $B \in \mathcal{B}$ τότε $X^{-1}(B) \in \mathcal{A}$, που σημαίνει ότι $X^{-1}(\mathcal{R} \setminus B) = (X^{-1}(B))^c \in \mathcal{A}$, και συνεπώς $\mathcal{R} \setminus B \in \mathcal{B}$.

(iii) Αν $\{B_n, n \geq 1\}$ είναι τυχούσα ακολουθία της $\mathcal{B}$, τότε $\{X^{-1}(B_n), n \geq 1\}$ είναι ακολουθία της $\mathcal{A}$, και

$$X^{-1}\left(\bigcup_{n=1}^{\infty} B_n\right) = \bigcup_{n=1}^{\infty} (X^{-1}(B_n)) \in \mathcal{A},$$

δηλαδή, $\bigcup_{n=1}^{\infty} B_n \in \mathcal{B}$.

Τώρα, $\mathfrak{D}_1 \subseteq \mathcal{B}$ και η $\mathcal{B}$ είναι σ -άλγεβρα, που σημαίνει ότι

$$\mathcal{B}(\mathcal{R}) = \sigma(\mathfrak{D}_1) \subseteq \mathcal{B},$$

και συνεπώς, εφόσον $\mathcal{B} \subseteq \mathcal{B}(\mathcal{R})$, έπεται ότι

$$\mathcal{B} = \mathcal{B}(\mathcal{R}),$$

δηλαδή, το ζητούμενο. ■

Ορισμός 1.8: (μετρήσιμη απεικόνιση (συνάρτηση))

Έστω $(\Omega_1, \mathcal{A})$ και $(\Omega_2, \mathcal{B})$ δυο μετρήσιμοι χώροι, και $f: \Omega_1 \rightarrow \Omega_2$ τυχούσα απεικόνιση (συνάρτηση). Η f θα λέγεται $\mathcal{A}/\mathcal{B}$ –μετρήσιμη (ή απλά μετρήσιμη όταν το $\mathcal{B}$ εννοείται, ή απλούστερα όταν τα $\mathcal{A}$ και $\mathcal{B}$ εννοούνται) όταν

$$f^{-1}(B) \in \mathcal{A}, \text{ για κάθε } B \in \mathcal{B},$$

δηλαδή, όταν $f^{-1}(B) \subset \mathcal{A}$.

Σύμφωνα με τους άνωθεν ορισμούς, και λόγω της προηγούμενης πρότασης, η συνάρτηση $X: \Omega \rightarrow \mathcal{R}$ είναι τ.μ. όταν και μόνο όταν είναι $\mathcal{A}/\mathcal{B}(\mathcal{R})$ –μετρήσιμη. Συγκεκριμένα όταν είναι $\mathcal{A}$ –μετρήσιμη, εφόσον η *Borel* σ -άλγεβρα εννοείται.

Στα μαθηματικά, η πιο ενδιαφέρουσα περίπτωση μετρήσιμης συνάρτησης είναι η *Borel*–μετρήσιμη συνάρτηση (ή απλά *Borel*–συνάρτηση), $f: \mathcal{R}^k \rightarrow \mathcal{R}^s$, που εξ’ ορισμού σημαίνει ότι η f είναι $\mathcal{B}(\mathcal{R}^k)/\mathcal{B}(\mathcal{R}^s)$ –μετρήσιμη, ήτοι $f^{-1}(\mathcal{B}(\mathcal{R}^s)) \subset \mathcal{B}(\mathcal{R}^k)$. Εδώ εννοούνται οι σ -άλγεβρες $\mathcal{B}(\mathcal{R}^k)$ και $\mathcal{B}(\mathcal{R}^s)$, αφού καθορίζονται από τις τιμές των k και s , που είναι γνωστές από το γεγονός ότι η f είναι συνάρτηση από τον $\mathcal{R}^k$ στον $\mathcal{R}^s$.

Οι συναρτήσεις *Borel* $f: \mathcal{R}^k \rightarrow \mathcal{R}^s$ είναι πολύ χρήσιμες, τόσο στη θεωρία μέτρου, όσο και στη θεωρία πιθανοτήτων. Το ενδιαφέρον παρουσιάζεται από το γεγονός ότι περιέχουν όλες τις συνεχείς συναρτήσεις $f: \mathcal{R}^k \rightarrow \mathcal{R}^s$, αλλά και από την σημαντικότερη ιδιότητα ότι αν η ακολουθία $\{f_n, n \geq 1\}$ είναι ακολουθία συναρτήσεων *Borel*, τότε η συνάρτηση (το κατά σημείο όριο των f_n)

$$f(x) = \lim_{n \rightarrow \infty} f_n(x)$$

είναι *Borel*. Μία ιδιότητα που δεν την έχουν οι συνεχείς συναρτήσεις. Άρα, η κλάση των *Borel* συναρτήσεων είναι πλουσιότερη και με καλές ιδιότητες.

Ορισμός 1.9: (εκτεταμένη τυχαία μεταβλητή)

Η τ.μ. X καλείται εκτεταμένη στην περίπτωση που λαμβάνει επιπλέον, και τις τιμές $\pm\infty$. Ουσιαστικά η μόνη επιπρόσθετη συνθήκη για μία εκτεταμένη τ.μ. X είναι η απαίτηση ότι τα σύνολα $\{\omega: X(\omega) = +\infty\}$ και $\{\omega: X(\omega) = -\infty\}$ είναι ενδεχόμενα. Παρατηρώντας ότι

$$\{\omega: X(\omega) = -\infty\} = \bigcap_{n=1}^{\infty} \{\omega: X(\omega) \leq -n\},$$

και

$$\{\omega: X(\omega) = +\infty\} = \bigcap_{n=1}^{\infty} \{\omega: X(\omega) \geq n\},$$

έπεται ότι ο ορισμός της συνήθους τ.μ.,

$$\{\omega: X(\omega) \leq b\} \in \mathcal{A}, \text{ για κάθε } b \in \mathcal{R},$$

καλύπτει και τις εκτεταμένες τ.μ.

Οι εκτεταμένες τ.μ. χρειάζονται όταν έχουμε όριο ακολουθίας (συνήθων) τ.μ., όπως,

$$X(\omega) = \lim_{n \rightarrow \infty} \inf X_n(\omega),$$

οπότε η $X(\omega)$ μπορεί να λάβει και τις τιμές $+\infty$ ή $-\infty$, ακόμη και όταν αυτό δεν συμβαίνει τις $X_n(\omega)$.

Ορισμός 1.10: Έστω X διακριτή τ.μ. με τιμές σε ένα αριθμήσιμο σύνολο $S \subseteq \mathcal{R}$ [$P(X \in S) = 1$]. Τότε η συνάρτηση μάζας πιθανότητας $f(x) = P(X = x)$, $x \in S$ καλείται και πυκνότητα ως προς το αριθμητικό μέτρο ν [$\nu(A)$ είναι ο πληθικός αριθμός του A , όπου A τυχόν υποσύνολο του $\mathcal{R}$] και ισχύει $B \subseteq S$, $P(X \in B) = \sum_{x \in B} P(X = x) = \sum_{x \in B} f(x) = \int_B f d\nu$.

Ορισμός 1.11: Έστω X (απόλυτα) συνεχής τ.μ. με τιμές στο $\mathcal{R}$. Τότε η συνάρτηση πυκνότητας πιθανότητας $f(x) (\neq P(X = x) = 0)$ καλείται και πυκνότητα ως προς το μέτρο Lebesgue λ ($\lambda(a, b) = b - a$) και ισχύει $B \subseteq \mathcal{R}$, $P(X \in B) = \int_B f(x) dx = \int_B f d\lambda$ και η συνάρτηση κατανομής $F(x) = P(X \leq x)$, $x \in \mathcal{R}$ είναι απόλυτα συνεχής συνάρτηση (άρα ομοιόμορφα συνεχής οπότε και συνεχής) πράγμα που σημαίνει ότι $F' = f$ σ.π. (σχεδόν παντού).

1.4 Θεωρία Πιθανοτήτων & Μαθηματική Στατιστική

1.4.1 Εισαγωγή

Η Θεωρία Πιθανοτήτων είναι ο κλάδος των μαθηματικών ο οποίος ασχολείται με την ανάλυση τυχαίων φαινομένων. Κεντρικό ρόλο στη Θεωρία Πιθανοτήτων παίζει η έννοια της πιθανότητας, ενώ σημαντικές είναι οι τυχαίες μεταβλητές, οι συναρτήσεις κατανομής, οι στοχαστικές διαδικασίες και τα γεγονότα: μαθηματικές αφαιρέσεις μη ντετερμινιστικών συμβάντων τα οποία είτε συμβαίνουν μία φορά είτε εξελίσσονται με το πέρασμα του χρόνου. Αν και τα γεγονότα που μελετώνται από τη Θεωρία Πιθανοτήτων, όπως π.χ. η ρίψη ενός ζαριού ή το στρίψιμο ενός κέρματος, είναι τυχαία, όταν επαναλαμβάνονται πολλές φορές η αλληλουχία των τυχαίων γεγονότων παρουσιάζει ορισμένα στατιστικά μοτίβα τα οποία μπορούν να μελετηθούν και να προβλεφθούν. Δύο αντιπροσωπευτικά μαθηματικά αποτελέσματα που περιγράφουν τέτοια μοτίβα είναι οι Νόμοι των Μεγάλων Αριθμών και το Κεντρικό Οριακό Θεώρημα.

Μαθηματική Στατιστική είναι η περιοχή των Πιθανοτήτων και της Στατιστικής η οποία, μελετά, εξετάζει και ερευνά τα θέματα του κλάδου αυτού κάτω από το φώς της μαθηματικής δομής και αυστηρότητας. Έτσι, μεταξύ άλλων περιλαμβάνει τη Θεωρία Πιθανοτήτων, τη θεωρία τυχαίων μεταβλητών και κατανομών, την εκτίμηση παραμέτρων, τον έλεγχο στατιστικών υποθέσεων, τη θεωρία μεγάλων δειγμάτων, την ακολουθιακή ανάλυση κτλ. Με άλλο χαρακτηρισμό η Μαθηματική Στατιστική είναι ταυτόσημη με τη Στατιστική Θεωρία.

Θεώρημα 1.4: (Ισχυρός Νόμος των Μεγάλων Αριθμών (του Kolmogorov) για ανεξάρτητη ακολουθία)

Έστω $\{X_n, n \geq 1\}$ μία ανεξάρτητη ακολουθία τυχαίων μεταβλητών για την οποία ισχύει ότι $E(X_n) = 0$ και $Var(X_n) < \infty$ για κάθε n . Αν

$$\sum_{n=1}^{\infty} \frac{Var(X_n)}{n^2} < \infty$$

τότε

$$\bar{X}_n = \frac{X_1 + X_2 + \dots + X_n}{n} \xrightarrow{a.s.} 0.$$

Απόδειξη: Θέτουμε $Y_n = X_n/n$, οπότε, η Y_n είναι ανεξάρτητη ακολουθία τ.μ. με $Var(Y_n) = Var(X_n)/n^2$, και άρα, $\sum_{n=1}^{\infty} Var(Y_n) < \infty$. Επίσης, έχουμε $E(Y_n) = 0$ για κάθε n . Επομένως, υπάρχει (μη εκτεταμένη) τ.μ. X , τέτοια ώστε

$$\sum_{k=1}^n \frac{X_k(\omega)}{k} \xrightarrow{n \rightarrow \infty} X(\omega),$$

εκτός ίσως από κάποια $\omega \in \Omega$, τα οποία ανήκουν σε ένα ενδεχόμενο A , με $P(A) = 0$. Φυσικά, η $X(\omega)$ είναι πεπερασμένη για $\omega \in A^c$ και έτσι αν $\omega \in A^c$ τότε

$$\frac{X_1(\omega) + X_2(\omega) + \dots + X_n(\omega)}{n} \xrightarrow{n \rightarrow \infty} 0,$$

που σημαίνει ότι

$$A^c \subset \left\{ \omega : \frac{X_1(\omega) + X_2(\omega) + \dots + X_n(\omega)}{n} \xrightarrow{n \rightarrow \infty} 0 \right\}.$$

Αφού, $P(A) = 0$, προφανώς, θα έχουμε

$$P(A^c) = 1,$$

και συνεπώς,

$$P\left(\frac{X_1 + X_2 + \dots + X_n}{n} \xrightarrow{a.s.} 0\right) = 1,$$

δηλαδή,

$$\bar{X}_n = \frac{X_1 + X_2 + \dots + X_n}{n} \xrightarrow{a.s.} 0.$$

■

Πόρισμα 1.1: (Ισχυρός Νόμος των Μεγάλων Αριθμών (του Kolmogorov) για ανεξάρτητη ακολουθία)

Έστω $\{X_n, n \geq 1\}$ μία ανεξάρτητη ακολουθία τ.μ. με $Var(X_n) < \infty$ και $E(X_n) = \mu_n, \forall n$. Εάν $\sum_{n=1}^{\infty} Var(X_n) / n^2 < \infty$, τότε

$$\frac{X_1 + X_2 + \dots + X_n}{n} - \frac{\mu_1 + \mu_2 + \dots + \mu_n}{n} \xrightarrow{a.s.} 0,$$

δηλαδή,

$$\bar{X}_n - \bar{\mu}_n \xrightarrow{a.s.} 0,$$

όπου

$$\bar{\mu}_n = \frac{\mu_1 + \mu_2 + \dots + \mu_n}{n}, \quad n = 1, 2, \dots$$

Αν, επιπλέον, ισχύει ότι, καθώς $n \rightarrow \infty$,

$$\frac{\mu_1 + \mu_2 + \dots + \mu_n}{n} \rightarrow \mu \in \mathcal{R},$$

τότε

$$\bar{X}_n = \frac{X_1 + X_2 + \dots + X_n}{n} \xrightarrow{a.s.} \mu.$$

Απόδειξη: Οι τυχαίες μεταβλητές $\tilde{X}_n = X_n - E(X_n) = X_n - \mu_n$ ικανοποιούν τις υποθέσεις του προηγούμενου Θεωρήματος, και άρα, ισχύει ότι

$$\frac{X_1 - \mu_1 + X_2 - \mu_2 + \dots + X_n - \mu_n}{n} \xrightarrow{a.s.} 0,$$

δηλαδή,

$$\bar{X}_n - \bar{\mu}_n = \frac{X_1 + X_2 + \dots + X_n}{n} - \frac{\mu_1 + \mu_2 + \dots + \mu_n}{n} \xrightarrow{a.s.} 0.$$

Αν, επιπροσθέτως, ισχύει $(\mu_1 + \mu_2 + \dots + \mu_n)/n \xrightarrow{n \rightarrow \infty} \mu \in \mathcal{R}$, τότε έχουμε ότι $\bar{\mu}_n \rightarrow \mu$ (άρα, και $\bar{\mu}_n \xrightarrow{a.s.} \mu$), $\bar{X}_n - \bar{\mu}_n \xrightarrow{a.s.} 0$, και τελικά,

$$\bar{X}_n = (\bar{X}_n - \bar{\mu}_n) + \bar{\mu}_n \xrightarrow{a.s.} 0 + \mu = \mu, \text{ διότι, αν } \bar{X}_n \xrightarrow{a.s.} X \text{ και } \bar{Y}_n \xrightarrow{a.s.} c \in \mathcal{R}, \text{ τότε}$$

$\bar{X}_n + \bar{Y}_n \xrightarrow{a.s.} X + c$ και εφαρμόσαμε αυτό για $X_n = \bar{X}_n - \bar{\mu}_n \xrightarrow{a.s.} 0$, και $Y_n(\omega) = \bar{\mu}_n \rightarrow \mu, \forall \omega \in \Omega$. ■

Ορισμός 1.12: (συνεπής εκτιμήτρια)

Η ακολουθία εκτιμητριών $\{T_n\}_{n \geq 1}$ (ή πιο απλά, η εκτιμήτρια T_n) μίας παραμετρικής συνάρτησης $g(\theta)$ λέγεται (ασθενικά) συνεπής εάν, για κάθε $\varepsilon > 0$ ισχύει

$$\lim_{n \rightarrow \infty} P(g(\theta) - \varepsilon < T_n < g(\theta) + \varepsilon) = 1 \text{ για κάθε } \theta \in \Theta.$$

Με άλλα λόγια, η T_n είναι συνεπής, εάν η πιθανότητα, η T_n να λάβει τιμή που θα διαφέρει αριθμητικά από την $g(\theta)$ περισσότερο από μία αυθαίρετη σταθερή ποσότητα ε , τείνει στο 0, όταν $n \rightarrow \infty$. Βλέπουμε ότι η συνέπεια είναι μία ασυμπτωτική ιδιότητα (ή οριακή ιδιότητα). Στην ουσία όμως, δεν είναι τίποτε άλλο παρά η σύγκλιση τυχαίων μεταβλητών κατά πιθανότητα που είναι γνωστή από τη Θεωρία Πιθανοτήτων, δηλαδή, $T_n \xrightarrow{p} g(\theta)$. Το επόμενο Θεώρημα δίνει επαρκείς συνθήκες, ώστε η εκτιμήτρια T_n να είναι συνεπής.

Θεώρημα 1.5: Έστω T_n μία εκτιμήτρια στατιστική συνάρτηση μίας παραμετρικής συνάρτησης $g(\theta)$. Εάν

- (i) η T_n είναι αμερόληπτη εκτιμήτρια της $g(\theta)$, δηλαδή, $E(T_n) = g(\theta)$ και
- (ii) $Var(T_n) \rightarrow 0$ όταν $n \rightarrow \infty$

τότε η T_n είναι συνεπής εκτιμήτρια της $g(\theta)$.

Παρατήρηση: Το συμπέρασμα ισχύει και αν ακόμη η T_n είναι ασυμπτωτικά αμερόληπτη εκτιμήτρια, δηλαδή, $E(T_n) \rightarrow g(\theta)$ καθώς $n \rightarrow \infty$.

Απόδειξη: Από την ανισότητα του *Chebyshev*²³ έχουμε

$$P(|T_n - g(\theta)| < \varepsilon) \geq 1 - \frac{E(T_n - g(\theta))^2}{\varepsilon^2} \quad \text{για κάποιο } \varepsilon > 0.$$

Όμως,

$$\begin{aligned} E(T_n - g(\theta))^2 &= E\{T_n - E(T_n) + E(T_n) - g(\theta)\}^2 \\ &= E\{(T_n - E(T_n))^2 + 2(T_n - E(T_n))(E(T_n) - g(\theta)) + (E(T_n) - g(\theta))^2\} \\ &= E\{T_n - E(T_n)\}^2 + 2(E(T_n) - E(T_n))(E(T_n) - g(\theta)) + \{E(T_n) - g(\theta)\}^2 \\ &= Var(T_n) + \{E(T_n) - g(\theta)\}^2 \xrightarrow{n \rightarrow \infty} 0 + 0 = 0, \end{aligned}$$

και έτσι,

$$\lim_{n \rightarrow \infty} P(g(\theta) - \varepsilon < T_n < g(\theta) + \varepsilon) \geq 1.$$

Επιπλέον, επειδή ισχύει πάντοτε $P(g(\theta) - \varepsilon < T_n < g(\theta) + \varepsilon) \leq 1$, έπεται ότι

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

Οπότε, βάσει του προηγούμενου ορισμού, η T_n είναι συνεπής εκτιμήτρια της $g(\theta)$. Για το β' μέρος της απόδειξης χρησιμοποιείται η ανισότητα του *Markov*²⁴,

²³Ρώσος μαθηματικός με πλήρες όνομα *Pafnuty Lvovich Chebyshev*, 16 Μαΐου 1821 - 8 Δεκεμβρίου 1894. Σπούδασε στο Πανεπιστήμιο της Μόσχας και είχε τεράστια συμβολή στους τομείς, Πιθανότητες, Στατιστική, Μηχανική και Θεωρία Αριθμών. Έμεινε γνωστός για την ανισότητα *Chebyshev*, τα πολυώνυμα *Chebyshev* και την μεροληψία *Chebyshev*. Δύο από τους πιο επιφανείς του μαθητές ήταν ο *Andrey Andreyevich Markov* (ο πρεσβύτερος) και ο *Aleksander Mikhailovich Lyapunov*.

$$P(X \geq \varepsilon) \leq \frac{E(X)}{\varepsilon}, \quad \text{για } X \geq 0, \text{ και } \varepsilon > 0. \quad \blacksquare$$

Παρατήρηση: Το αντίστροφο του προηγούμενου Θεωρήματος γενικά δεν ισχύει διότι η $T_n \xrightarrow{p} g(\theta)$ δεν μας δίνει πληροφορία για την ασυμπτωτική συμπεριφορά των ροπών της T_n καθώς $n \rightarrow \infty$.

Ορισμός 1.13: (ασυμπτωτική κανονικότητα)

Η ακολουθία εκτιμητριών $\{T_n\}_{n \geq 1}$ μίας παραμετρικής συνάρτησης $g(\theta)$ καλείται ασυμπτωτικά κανονική εκτιμήτρια της $g(\theta)$ όταν $\exists r(\theta) \in [0, +\infty)$:

$$\sqrt{n}(T_n - g(\theta)) \xrightarrow{d} \mathcal{N}(0, r^2(\theta)), \quad \forall \theta \in \Theta, \text{ καθώς } n \rightarrow \infty.$$

Πρόταση 1.3:

(i) Αν $X_n \xrightarrow{a.s.} X$ τότε $X_n \xrightarrow{p} X$ και $X_n \xrightarrow{d} X$. (τα αντίστροφα γενικά δεν ισχύουν)

(ii) $X_n \xrightarrow{p} c \Leftrightarrow X_n \xrightarrow{d} c, \quad c \in \mathcal{R}.$

(iii) Αν το Ω είναι αριθμήσιμο σύνολο και $\mathcal{A} = \mathcal{P}(\Omega)$ τότε $X_n \xrightarrow{p} X \Leftrightarrow X_n \xrightarrow{d} X$.

Θεώρημα 1.6: (Slutsky²⁵)

Έστω ότι $X_n \xrightarrow{d} X$ και $Y_n \xrightarrow{d} c$, όπου $c \in \mathcal{R}$, σταθερά. [Οι X_n, Y_n, X μπορεί να έχουν (ή να μην έχουν) οποιαδήποτε σχέση μεταξύ τους.] Τότε ισχύουν τα εξής:

(i) $X_n + Y_n \xrightarrow{d} X + c.$

(ii) $X_n Y_n \xrightarrow{d} cX.$

²⁴Ρώσος μαθηματικός με πλήρες όνομα *Andrey Andreyevich Markov*, 14 Ιουνίου 1856 - 20 Ιουλίου 1922. Είναι πολύ δημοφιλής για την εργασία του στις Στοχαστικές Διαδικασίες. Ένα πρωταρχικό αντικείμενο της έρευνάς του έγινε, μετέπειτα, γνωστό ως Μαρκοβιανές αλυσίδες και ως Μαρκοβιανές διαδικασίες.

²⁵Ρώσος μαθηματικός, στατιστικός και πολιτικός οικονομολόγος με πλήρες όνομα *Eugen Slutsky*, 7 Απριλίου 1880 - 10 Μαρτίου 1948.

(iii) $\frac{X_n}{Y_n} \xrightarrow{d} \frac{X}{c}$, εφόσον $c \neq 0$.

Απόδειξη: Έστω F_n, F οι συναρτήσεις κατανομών των X_n, X .

(i) Έστω $\varepsilon > 0$, έχουμε

$$\begin{aligned} F_{X_n+Y_n}(t) &= P(X_n + Y_n \leq t) \\ &= P(X_n + Y_n \leq t, |Y_n - c| \leq \varepsilon) + P(X_n + Y_n \leq t, |Y_n - c| > \varepsilon) \\ &\leq P(X_n + Y_n \leq t, c - \varepsilon \leq Y_n \leq c + \varepsilon) + P(|Y_n - c| > \varepsilon) \\ &\leq F_n(t - c + \varepsilon) + P(|Y_n - c| > \varepsilon). \end{aligned}$$

Παρόμοια,

$$\begin{aligned} F_n(t - c - \varepsilon) &= P(X_n \leq t - c - \varepsilon) \\ &= P(X_n \leq t - c - \varepsilon, |Y_n - c| \leq \varepsilon) + P(X_n \leq t - c - \varepsilon, |Y_n - c| > \varepsilon) \\ &\leq P(X_n + Y_n \leq t, |Y_n - c| > \varepsilon) + P(|Y_n - c| > \varepsilon) \\ &= F_{X_n+Y_n}(t) + P(|Y_n - c| > \varepsilon) \end{aligned}$$

άρα, για τυχόν $\varepsilon > 0$, ισχύει

$$F_n(t - c - \varepsilon) - P(|Y_n - c| > \varepsilon) \leq F_{X_n+Y_n}(t) \leq F_n(t - c + \varepsilon) + P(|Y_n - c| > \varepsilon).$$

Επειδή μία συνάρτηση κατανομής F είναι συνεχής στο $\mathcal{R} \setminus N$, όπου N είναι ένα (το πολύ) αριθμήσιμο σύνολο, μπορούμε να βρούμε αυθαίρετα μικρό $\varepsilon > 0$, έτσι ώστε τα $t - c \pm \varepsilon$ να είναι (και τα δύο) σημεία συνεχείας της F . Λόγω της τελευταίας σχέσης $\forall \varepsilon > 0$ θα είναι

$$F(t - c - \varepsilon) \leq \liminf_{n \rightarrow +\infty} F_{X_n+Y_n}(t) \leq \limsup_{n \rightarrow +\infty} F_{X_n+Y_n}(t) \leq F(t - c + \varepsilon)$$

Συνεπώς, καθώς $\varepsilon \rightarrow 0^+$ για το οποίο τα $t - c \pm \varepsilon$ είναι σημεία συνεχείας της F , προκύπτει ότι

$$F_{X_n+Y_n}(t) \rightarrow F(t - c) = P(X \leq t - c) = P(X + c \leq t), \text{ καθώς } n \rightarrow +\infty,$$

δεδομένου ότι το $t - c$ είναι σημείο συνεχείας της F . Επομένως, το ζητούμενο έπεται αρκεί να παρατηρήσουμε ότι το $t - c$ είναι σημείο συνεχείας της F αν και μόνο αν το t είναι σημείο συνεχείας της F_{c+X} .

(ii) Πρωτίστως θεωρούμε ότι $c = 0$. Τότε, για $\varepsilon > 0$, $\alpha > 0$,

$$P(|X_n Y_n| \geq \varepsilon) = P(|X_n Y_n| \geq \varepsilon, |Y_n| < \alpha) + P(|X_n Y_n| \geq \varepsilon, |Y_n| \geq \alpha)$$

$$\leq P\left(|X_n| > \frac{\varepsilon}{\alpha}\right) + P(|Y_n| \geq \alpha)$$

$$\leq F_n\left(-\frac{\varepsilon}{\alpha}\right) + 1 - F_n\left(\frac{\varepsilon}{\alpha}\right) + P(|Y_n| \geq \alpha).$$

Αφήνοντας το $n \rightarrow +\infty$, και εκλέγοντας $\alpha > 0$ τέτοιο ώστε τα $\pm \frac{\varepsilon}{\alpha}$ να είναι σημεία συνεχείας της F ,

$$\lim_{n \rightarrow +\infty} \sup P(|X_n Y_n| \geq \varepsilon) \leq F\left(-\frac{\varepsilon}{\alpha}\right) + 1 - F\left(\frac{\varepsilon}{\alpha}\right) \rightarrow 0, \text{ καθώς } \alpha \rightarrow 0^+,$$

δηλαδή, $P(|X_n Y_n| \geq \varepsilon) \rightarrow 0$, καθώς $n \rightarrow +\infty$. Οπότε, $X_n Y_n \xrightarrow{p} 0$, ήτοι $X_n Y_n \xrightarrow{d} 0 = 0 \cdot X$.

Στην γενική περίπτωση όπου $Y_n \xrightarrow{d} c$, συνάγεται ότι $Y_n - c \xrightarrow{d} 0$. Άρα, από το αποτέλεσμα που μόλις αποδείξαμε, προκύπτει ότι

$$X_n Y_n - c X_n = X_n (Y_n - c) \xrightarrow{d} 0.$$

Όμως, προφανώς, $c X_n \xrightarrow{d} c X$ και από το (i),

$$X_n Y_n = X_n (Y_n - c) + c X_n \xrightarrow{d} 0 + c X = c X.$$

(iii) Από το (ii), και το γεγονός ότι $\frac{1}{Y_n} \xrightarrow{d} \frac{1}{c}$ (διότι η συνάρτηση $x \mapsto 1/x$ είναι συνεχής στο $x = c \neq 0$), απορρέει ότι

$$\frac{X_n}{Y_n} = \frac{1}{Y_n} X_n \xrightarrow{d} \frac{1}{c} X = \frac{X}{c}.$$

■

Παρατήρηση: Αν $X_n \xrightarrow{d} X$ και $Y_n \xrightarrow{d} Y$ τότε, δεν ισχύει κατ' ανάγκην ότι $X_n \pm Y_n \xrightarrow{d} X \pm Y$. Στην ειδική περίπτωση που οι X_n, Y_n είναι ανεξάρτητες και οι X, Y είναι ανεξάρτητες τότε $X_n \pm Y_n \xrightarrow{d} X \pm Y$.

Ορισμός 1.14: (ομοιόμορφης ολοκληρωσιμότητας)

Η ακολουθία τ.μ. $\{X_n, n \geq 1\}$ (όπου κάθε τυχαία μεταβλητή X_n επιτρέπεται να ορίζεται σε διαφορετικό χώρο πιθανότητας $(\Omega_n, \mathcal{A}_n, \mathcal{P}_n)$) καλείται ομοιόμορφα ολοκληρώσιμη όταν

$$\sup_{n \geq 1} E[|X_n| I(|X_n| \geq \alpha)] \rightarrow 0, \text{ καθώς } \alpha \rightarrow \infty.$$

Ισοδύναμα χρησιμοποιείται και η διατύπωση «οι X_n είναι ομοιόμορφα ολοκληρώσιμες»

Θεώρημα 1.7: (ομοιόμορφης ολοκληρωσιμότητας)

Έστω μία ομοιόμορφα ολοκληρώσιμη ακολουθία τ.μ. $\{X_n, n \geq 1\}$, και ας υποθέσουμε ότι $X_n \rightarrow X$ (με πιθανότητα 1). Τότε ισχύουν τα εξής:

- (i) Η X είναι ολοκληρώσιμη (δηλ. $E|X| < \infty$).
- (ii) $E(X_n) \rightarrow E(X)$.
- (iii) $E|X_n - X| \rightarrow 0$.

Τώρα, εάν $X_n \geq 0$, $X_n \rightarrow X$ (με πιθανότητα 1), $E(X) < \infty$, $E(X_n) < \infty$ για κάθε $n \geq 1$ και $E(X_n) \rightarrow E(X)$ τότε η ακολουθία $\{X_n, n \geq 1\}$ είναι ομοιόμορφα ολοκληρώσιμη.

Θεώρημα 1.8: (Skorokhod²⁶)

Θεωρούμε τις τ.μ. $\{X_n, n \geq 1\}$ και X , ορισμένες σε αυθαίρετους χώρους πιθανότητας, και τέτοιες ώστε $X_n \xrightarrow{d} X$. Τότε, υπάρχει χώρος πιθανότητας $(\Omega, \mathcal{A}, \mathcal{P})$, και τ.μ. $\{Y_n, n \geq 1\}$ και Y , με τις εξής ιδιότητες:

- (i) Οι τ.μ. $\{Y_n, n \geq 1\}$ και Y είναι ορισμένες στον ίδιο χώρο πιθανότητας $(\Omega, \mathcal{A}, \mathcal{P})$.
- (ii) $Y_n \xleftrightarrow{d} X_n, n \geq 1$, και $Y \xleftrightarrow{d} X$.
- (iii) Για κάθε $\omega \in \Omega$, $Y_n(\omega) \rightarrow Y(\omega)$ (δηλ., $Y_n \rightarrow Y$ κατά σημείο).

Πρόταση 1.4: Έστω $X_n \xrightarrow{p} X$, $Y_n \xrightarrow{p} Y$ τότε

- (i) $X_n \pm Y_n \xrightarrow{p} X \pm Y$.
- (ii) $X_n Y_n \xrightarrow{p} XY$.

²⁶Ο *Anatoliy Volodymyrovych Skorokhod* (10 Σεπτεμβρίου 1930 – 3 Ιανουαρίου 2011) υπήρξε Ουκρανός μαθηματικός ο οποίος, εργάστηκε στο Πανεπιστήμιο του Κιέβου κατά τα έτη (1956 - 1964). Μετέπειτα, διετέλεσε καθηγητής και στο Πανεπιστήμιο του Μίσιγκαν της Αμερικής. Δημοσίευσε επιστημονικά έργα περί στοχαστικών διαφορικών εξισώσεων, οριακών θεωρημάτων των τυχαίων διαδικασιών, κατανομών σε απειροδιάστατους χώρους και Μαρκοβιανών διαδικασιών.

(iii) $\frac{X_n}{Y_n} \xrightarrow{p} \frac{X}{Y}$, εφόσον, $P(Y_n = 0) = P(Y = 0) = 0$.

Η τελευταία πρόταση υφίσταται και για την ισχυρή σύγκλιση $\left(\xrightarrow{a.s.}\right)$.

Πρόταση 1.5: (Slutsky – Fréchet²⁷)

Αν $X_n \xrightarrow{d} X$ και g συνεχής συνάρτηση τότε $g(X_n) \xrightarrow{d} g(X)$. Το ίδιο ισχύει και για την ισχυρή $\left(\xrightarrow{a.s.}\right)$ και για την κατά πιθανότητα σύγκλιση $\left(\xrightarrow{p}\right)$.

Μοναδικότητα των ορίων:

- (i) $\begin{cases} X_n \xrightarrow{p} X \\ Y_n \xrightarrow{p} Y \end{cases} \Rightarrow P(X = Y) = 1.$
- (ii) $\begin{cases} X_n \xrightarrow{a.s.} X \\ Y_n \xrightarrow{a.s.} Y \end{cases} \Rightarrow P(X = Y) = 1 \text{ ή } (X = Y \text{ a.s.}).$
- (iii) $\begin{cases} X_n \xrightarrow{d} X \\ Y_n \xrightarrow{d} Y \end{cases} \Rightarrow F_X = F_Y.$

Η σύγκλιση κατά νομό $\left(\xrightarrow{d}\right)$ δεν επηρεάζεται άμεσα από τις τ.μ. X_n, X αλλά από τις κατανομές τους.

Σχέσεις $O(\cdot)$ και $o(\cdot)$

Ορισμός 1.15: Αν $\{\alpha_n\}_{n \geq 1}, \{\beta_n\}_{n \geq 1}$ είναι ακολουθίες πραγματικών αριθμών τότε έχουμε τα κάτωθι:

- $\alpha_n = O(1)$ όταν η $\{\alpha_n\}_{n \geq 1}$ είναι φραγμένη (δηλ. $\exists M > 0: |\alpha_n| \leq M, \forall n$).
- $\alpha_n = o(1)$ όταν $\alpha_n \rightarrow 0$.

²⁷Γάλλος μαθηματικός με πλήρες όνομα *Réne Maurice Fréchet*, 21 Σεπτεμβρίου 1878 - 4 Ιουνίου 1973. Πραγματοποίησε σημαντικές συνεισφορές στην τοπολογία των σημειοσυνόλων και εισήγαγε την έννοια των μετρικών χώρων. Είχε επίσης, μεγάλη συμβολή στον τομέα πιθανοτήτων & στατιστικής, καθώς και στον τομέα του λογισμού. Η Διδακτορική του διατριβή διεύρυνε το πεδίο των συναρτήσεων πάνω στους μετρικούς χώρους και παρουσίασε την έννοια της συμπάγειας. Μάλιστα ανακάλυψε, ανεξάρτητα από τον *Riesz*, το θεώρημα αναπαράστασης στο χώρο των *Lebesgue* τετραγωνικών ολοκληρωτικών συναρτήσεων.

- $\alpha_n = O(\beta_n)$ όταν η $\left\{\frac{\alpha_n}{\beta_n}\right\}_{n \geq 1}$ είναι φραγμένη (δηλ. η έχει $\{\alpha_n\}$ τάξη μικρότερη ή ίση από αυτή της $\{\beta_n\}$).
- $\alpha_n = o(\beta_n)$ όταν $\frac{\alpha_n}{\beta_n} \rightarrow 0$ (δηλ. η έχει $\{\alpha_n\}$ τάξη μικρότερη της $\{\beta_n\}$).
- $\alpha_n \sim \beta_n$ όταν $\frac{\alpha_n}{\beta_n} \rightarrow 1$ (δηλ. η $\{\alpha_n\}$ ασυμπτωτικά ισοδύναμη της $\{\beta_n\}$).

Ορισμός 1.16: Έστω $\{X_n\}_{n \geq 1}, \{Y_n\}_{n \geq 1}$ είναι ακολουθίες τ.μ. ορισμένες στον ίδιο χώρο πιθανότητας $(\Omega, \mathcal{A}, \mathcal{P})$ τότε ισχύουν τα επόμενα:

- $X_n = O(1)$ *a.s.* όταν η $\{X_n\}_{n \geq 1}$ είναι φραγμένη *a.s.*
- $X_n = o(1)$ *a.s.* όταν $X_n \xrightarrow{a.s.} 0$
- $X_n = O(Y_n)$ *a.s.* όταν η $\left\{\frac{X_n}{Y_n}\right\}_{n \geq 1}$ είναι φραγμένη *a.s.*
- $X_n = o(Y_n)$ όταν $\frac{X_n}{Y_n} \xrightarrow{a.s.} 0$.
- $X_n \sim Y_n$ *a.s.* όταν $\frac{X_n}{Y_n} \xrightarrow{a.s.} 1$.

Ορισμός 1.17: Έστω ακολουθία τ.μ. $\{X_n\}_{n \geq 1}$ ορισμένη στον ίδιο χώρο πιθανότητας $(\Omega, \mathcal{A}, \mathcal{P})$. Λέμε ότι η $\{X_n\}_{n \geq 1}$ είναι φραγμένη κατά πιθανότητα $(X_n = O_p(1))$ εάν

$$\forall \varepsilon > 0, \exists M := M(\varepsilon) > 0: P(|X_n| \leq M) \geq 1 - \varepsilon, \forall n$$

ή ισοδύναμα

$$\forall \varepsilon > 0, \exists M := M(\varepsilon) > 0: P(|X_n| > M) < \varepsilon, \forall n.$$

Λήμμα 1.2:

- Αν $X_n = O(1)$ *a.s.* τότε $X_n = O_p(1)$. (το αντίστροφο γενικά δεν ισχύει)
- Αν $X_n \xrightarrow{d} X$ τότε $X_n = O_p(1)$.
- Αν $X_n = O_p(1)$ και $Y_n \xrightarrow{p} 0$ τότε $X_n Y_n \xrightarrow{p} 0$.

- Αν $X_n = O_p(1)$ και $Y_n = o_p(X_n)$ τότε $Y_n \xrightarrow{p} 0$.

Ορισμός 1.18: Έστω $\{X_n\}_{n \geq 1}, \{Y_n\}_{n \geq 1}$ είναι ακολουθίες τ.μ. ορισμένες στον ίδιο χώρο πιθανότητας $(\Omega, \mathcal{A}, \mathcal{P})$ τότε έχουμε τα ακόλουθα:

- $X_n = o_p(1) \Leftrightarrow X_n \xrightarrow{p} 0$.
- $X_n = O_p(Y_n) \Leftrightarrow \frac{X_n}{Y_n} = O_p(1)$ (δηλ. η $\left\{\frac{X_n}{Y_n}\right\}_{n \geq 1}$ είναι φραγμένη κατά πιθανότητα).
- $X_n = o_p(Y_n)$ όταν $\frac{X_n}{Y_n} \xrightarrow{p} 0$.
- $X_n \sim_p Y_n$ όταν $\frac{X_n}{Y_n} \xrightarrow{p} 1$.

Κεφάλαιο 2

Μέθοδος Δέλτα πρώτης τάξης

2.1 Η οριακή κατανομή των δειγματικών κεντρικών ροπών

2.1.1 Εισαγωγή

Έστω X είναι μία τυχαία μεταβλητή με συνάρτηση κατανομής F και $E(X^{2k}) < \infty$ για κάποιο σταθερό $k \in \{1, 2, \dots\}$, και θεωρούμε ένα τυχαίο δείγμα $X_1, X_2, \dots, X_n$ από την F . Προς αποφυγή τετριμμένων περιπτώσεων, υποθέτουμε επιπλέον, ότι η X είναι μη-εκφυλισμένη. Επομένως, το σύνολο των σημείων αύξησης της F ,

$$S_F := \{x \in \mathcal{R}: F(x + \varepsilon) - F(x - \varepsilon) > 0, \quad \forall \varepsilon > 0\},$$

περιέχει τουλάχιστον δύο σημεία. Υποθέτουμε ότι οι πρώτες $2k$ -κεντρικές ροπές της τυχαίας μεταβλητής X γύρω από τον μέσο $\mu := E(X)$, είναι καλώς-ορισμένες και πεπερασμένες. Αυτές εκφράζονται από

$$\mu_j := E(X - \mu)^j, \quad j = 0, 1, \dots, 2k.$$

Έτσι,

$$\mu_0 = E(X - \mu)^0 = 1,$$

$$\mu_1 = E(X - \mu)^1 = E(X) - \mu = E(X) - E(X) = 0,$$

$$\begin{aligned} \mu_2 &= E(X - \mu)^2 = E(X^2 - 2\mu X + \mu^2) = E(X^2) - 2\mu E(X) + \mu^2 \\ &= E(X^2) - E(X)^2 = \text{Var}(X) = \sigma^2 \text{ (που είναι αυστηρά θετική αφού υποθέσαμε} \\ &\quad \text{ότι η } X \text{ είναι μη-εκφυλισμένη),} \end{aligned}$$

$$\begin{aligned} \mu_3 &= E(X - \mu)^3 = E(X^3 - 3X^2\mu + 3X\mu^2 - \mu^3) \\ &= E(X^3) - 3E(X^2)E(X) + 3E(X)E(X)^2 - E(X)^3 \\ &= E(X^3) - 3E(X)[E(X^2) - E(X)^2] - E(X)^3 \\ &= E(X^3) - 3E(X)\text{Var}(X) - E(X)^3 \text{ κ.ο.κ.} \end{aligned}$$

Η φυσική εκτιμήτρια της μ_k , $\forall k \geq 2$ όταν το μ είναι άγνωστο (όπως στην περίπτωση αυτή), είναι η δειγματική κεντρική ροπή k -τάξης,

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

όπου $\bar{X} = \bar{X}_n := \frac{1}{n} \sum_{i=1}^n X_i$ είναι ο συνήθης δειγματικός μέσος. Είναι εύκολο να δούμε ότι η M_k μπορεί να εκφραστεί σε όρους των δειγματικών κεντρικών ροπών περί την (άγνωστη) μέση τιμή μ . Πράγματι, ορίζοντας

$$m_j = m_{j,n} := \frac{1}{n} \sum_{i=1}^n (X_i - \mu)^j, \quad j = 1, 2, \dots, k,$$

παρατηρούμε ότι $m_1 = \bar{X} - \mu$ και χρησιμοποιώντας το διώνυμο του *Newton*

$$(\alpha + \beta)^k = \sum_{j=0}^k \binom{k}{j} \alpha^j \beta^{k-j},$$

θα έχουμε διαδοχικά ότι

$$\begin{aligned} M_k &= \frac{1}{n} \sum_{i=1}^n (X_i - \mu - m_1)^k = \frac{1}{n} \sum_{i=1}^n \left[\sum_{j=0}^k \binom{k}{j} (X_i - \mu)^j (-1)^{k-j} m_1^{k-j} \right] \\ &= \frac{1}{n} \sum_{i=1}^n \left[(-1)^k m_1^k + k(X_i - \mu)(-1)^{k-1} m_1^{k-1} \right. \\ &\quad \left. + \sum_{j=2}^{k-1} \binom{k}{j} (X_i - \mu)^j (-1)^{k-j} m_1^{k-j} + (X_i - \mu)^k \right] \\ &= (-1)^k m_1^k + k(\bar{X} - \mu)(-1)^{k-1} m_1^{k-1} \\ &\quad + \sum_{j=2}^{k-1} \binom{k}{j} (-1)^{k-j} m_1^{k-j} \left(\frac{1}{n} \sum_{i=1}^n (X_i - \mu)^j \right) + \frac{1}{n} \sum_{i=1}^n (X_i - \mu)^k \\ &= (-1)^k m_1^k + k m_1 (-1)^{k-1} m_1^{k-1} + \sum_{j=2}^{k-1} \binom{k}{j} (-1)^{k-j} m_1^{k-j} m_j + m_k \\ &= (-1)^{k-1} m_1^k (k-1) + \sum_{j=2}^{k-1} \binom{k}{j} (-1)^{k-j} m_1^{k-j} m_j + m_k, \end{aligned}$$

όπου ένα κενό άθροισμα θα πρέπει να θεωρείται μηδέν. Εφόσον, $E|X|^k < \infty$, από τον Ισχυρό Νόμο των Μεγάλων Αριθμών συνεπάγεται ότι για κάθε $j \leq k$ ισχύει $m_{j,n} \xrightarrow{a.s.} \mu_j$ καθώς $n \rightarrow \infty$, και ιδίως, $m_{1,n} \xrightarrow{a.s.} 0$.

Έτσι, εύκολα παίρνουμε το γνωστό αποτέλεσμα ότι $M_{k,n} \xrightarrow{a.s.} \mu_k$.

Πράγματι,

$$M_{k,n} \xrightarrow{a.s.} (-1)^{k-1} (k-1)0^k + \sum_{j=2}^{k-1} \binom{k}{j} (-1)^{k-j} \mu_j 0^{k-j} + \mu_k = \mu_k,$$

καθώς $n \rightarrow \infty$. Δηλαδή, η M_k είναι μία (ισχυρά) συνεπής εκτιμήτρια της μ_k .

Παρακάτω θα αποδείξουμε ότι η M_k είναι ασυμπτωτικά κανονική εκτιμήτρια της μ_k υπό την υπόθεση ότι $E(X^{2k}) < \infty$, δηλαδή, καθώς $n \rightarrow \infty$,

$$\sqrt{n}(M_{k,n} - \mu_k) \xrightarrow{d} \mathcal{N}(0, v_k^2),$$

$$\text{όπου } v_k^2 = \mu_{2k} - \mu_k^2 + k\mu_{k-1}(k\sigma^2\mu_{k-1} - 2\mu_{k+1}). \quad (2.1)$$

Για παράδειγμα, οι πρώτες τρεις τιμές της οριακής διασποράς v_k^2 , εφόσον, $\mu_1 = 0$ και $\mu_2 = \sigma^2$, είναι

$$v_2^2 = \mu_4 - \sigma^4 \quad (\text{ισχύει όταν } E(X^4) < \infty),$$

$$v_3^2 = \mu_6 - \mu_3^2 + 9\sigma^6 - 6\sigma^2\mu_4 \quad (\text{ισχύει όταν } E(X^6) < \infty),$$

$$v_4^2 = \mu_8 - \mu_4^2 - 8\mu_3\mu_5 + 16\sigma^2\mu_3^2 \quad (\text{ισχύει όταν } E(X^8) < \infty).$$

Ο στόχος της παρούσας διατριβής είναι να εντρυφήσουμε στην οριακή συμπεριφορά του διανύσματος $\mathbf{M} = \mathbf{M}_n := (M_1, M_2, \dots, M_k)'$ καθώς $n \rightarrow \infty$. Ειδικότερα, θα διερευνηθούν λεπτομερώς οι ιδιάζουσες περιπτώσεις δηλαδή όταν $v_k^2 = 0$. Επιπρόσθετα, θα δείξουμε ότι, μεταξύ των κατανομών που έχουν πεπερασμένες ροπές οποιασδήποτε τάξης, η ασυμπτωτική ανεξαρτησία, του δειγματικού μέσου $\bar{X}$ και της ακολουθίας $\{M_k, k \geq 2\}$, χαρακτηρίζει την κανονική κατανομή. Αυτό το γεγονός αποδεικνύει, διαισθητικά, ένα οριακό ισοδύναμο του γνωστού αποτελέσματος ότι η ανεξαρτησία των ποσοτήτων $\bar{X}$ και $M_2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 = \frac{n-1}{n} S^2$ (για κάποια σταθερά $n \geq 2$), όπου S^2 η δειγματική διακύμανση, χαρακτηρίζει την κανονικότητα – βλ. *Geary (1936)*, *Zinger (1958)*, *Laha et al (1960)* και *Kagan et al (1973)*. Εδώ η υπόθεση της ανεξαρτησίας περιστέλλεται στην ασυμπτωτική ανεξαρτησία (καθώς $n \rightarrow \infty$) αλλά, η απαίτηση της ύπαρξης όλων των ροπών και η αξίωση ότι η $\bar{X}$ πρέπει να είναι ασυμπτωτικά ανεξάρτητη από όλες τις ροπές M_k , $k \geq 2$ (και όχι μόνο για $k = 2$), επιβάλλει επιπρόσθετους περιορισμούς. Όμως, θα αποδείξουμε ότι για οποιαδήποτε σταθερά $k \geq 2$, υπάρχουν (απείρως πολλές) μη-κανονικές κατανομές για τις οποίες $\bar{X}$ και $(M_2, \dots, M_k)'$ είναι ασυμπτωτικά ανεξάρτητα.

2.2 Η οριακή κατανομή και ο χαρακτηρισμός της κανονικότητας

Υποθέτουμε ότι $k \geq 2$ και ορίζουμε $\mathbf{m} = \mathbf{m}_n := (m_{1,n}, m_{2,n}, \dots, m_{k,n})'$. Το πολυπαραγοντικό Κεντρικό Οριακό Θεώρημα δίνει, πάραυτα, ότι καθώς $n \rightarrow \infty$,

$$\sqrt{n}(\mathbf{m}_n - \boldsymbol{\mu}) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}), \quad (2.2)$$

όπου $\mathbf{0} = (0, 0, \dots, 0)' \in \mathcal{R}^k$, $\boldsymbol{\mu} = (\mu_1, \mu_2, \dots, \mu_k)' = (0, \sigma^2, \dots, \mu_k)' \in \mathcal{R}^k$ και $\boldsymbol{\Sigma} = (\sigma_{ij}) \in \mathcal{R}^{k \times k}$ είναι ο πίνακας διασπορών–συνδιασπορών με $\sigma_{ij} = \mu_{i+j} - \mu_i \mu_j$. Παρατηρούμε ότι $M_{k,n} = g(\mathbf{m}_n)$, όπου $g: \mathcal{R}^k \rightarrow \mathcal{R}$ είναι Borel–μετρήσιμη συνάρτηση η οποία δίνεται από τον τύπο

$$g(x_1, x_2, \dots, x_k) = (-1)^{k-1}(k-1)x_1^k + \sum_{j=2}^{k-1} \binom{k}{j} (-1)^{k-j} x_1^{k-j} x_j + x_k. \quad (2.3)$$

Εφόσον, η g έχει συνεχείς μερικές παραγώγους, μία εφαρμογή της μεθόδου Δέλτα δίνει το όριο

$$\sqrt{n}(M_{k,n} - g(\boldsymbol{\mu})) = \sqrt{n}(g(\mathbf{m}_n) - g(\boldsymbol{\mu})) \xrightarrow{d} \mathcal{N}(\mathbf{0}, [\nabla g(\boldsymbol{\mu})]' \boldsymbol{\Sigma} [\nabla g(\boldsymbol{\mu})]), \quad (2.4)$$

όπου για $\mathbf{x} = (x_1, x_2, \dots, x_k)'$,

$$\nabla g(\mathbf{x}) = \begin{pmatrix} \frac{\partial g(x_1, x_2, \dots, x_k)}{\partial x_1} \\ \frac{\partial g(x_1, x_2, \dots, x_k)}{\partial x_2} \\ \vdots \\ \frac{\partial g(x_1, x_2, \dots, x_k)}{\partial x_k} \end{pmatrix}.$$

Συγκεκριμένα,

$$\begin{aligned} \bullet \frac{\partial g(x_1, x_2, \dots, x_k)}{\partial x_1} &= (-1)^{k-1}(k-1)kx_1^k + \sum_{j=2}^{k-1} \binom{k}{j} (-1)^{k-j} x_1^{k-j-1} x_j (k-j) \\ &= \sum_{j=1}^{k-1} \binom{k}{j} (-1)^{k-j} x_1^{k-j-1} x_j (k-j) \\ &= \sum_{j=1}^{k-2} \binom{k}{j} (-1)^{k-j} x_1^{k-j-1} x_j (k-j) + (-1)^{k+1-k} \binom{k}{k-1} x_{k-1} (k+1-k) x_1^{k+1-k-1} \end{aligned}$$

$$\begin{aligned}
&= \sum_{j=1}^{k-2} \binom{k}{j} (-1)^{k-j} x_1^{k-j-1} x_j (k-j) + (-1) \frac{k!}{(k-1)! 1!} x_{k-1} \\
&= \sum_{j=1}^{k-2} \binom{k}{j} (-1)^{k-j} x_1^{k-j-1} x_j (k-j) + (-1) \frac{(k-1)! k}{(k-1)! 1!} x_{k-1} \\
&= \sum_{j=1}^{k-2} \binom{k}{j} (-1)^{k-j} x_1^{k-j-1} x_j (k-j) - k x_{k-1},
\end{aligned}$$

$$\bullet \frac{\partial g(x_1, x_2, \dots, x_k)}{\partial x_i} = (-1)^{k-i} \binom{k}{i} x_1^{k-i}, \quad i = 2, 3, \dots, k-1,$$

$$\bullet \frac{\partial g(x_1, x_2, \dots, x_k)}{\partial x_k} = 1.$$

Έτσι, συνοπτικά έχουμε

$$\begin{aligned}
\frac{\partial g(x_1, x_2, \dots, x_k)}{\partial x_i} &= \\
&= \begin{cases} \sum_{j=1}^{k-2} \binom{k}{j} (-1)^{k-j} x_1^{k-j-1} x_j (k-j) - k x_{k-1}, & \text{για } i = 1, \\ (-1)^{k-i} \binom{k}{i} x_1^{k-i}, & \text{για } i = 2, 3, \dots, k-1, \\ 1, & \text{για } i = k. \end{cases} \quad (2.5)
\end{aligned}$$

Ως εκ τούτου, αφού $\mu_1 = 0$, από την σχέση (2.3) έπεται ότι $g(0, x_2, \dots, x_k) = x_k$ και συγκεκριμένα $g(\boldsymbol{\mu}) = g(\mu_1, \mu_2, \dots, \mu_k) = \mu_k$. Επιπλέον, από την (2.5) παίρνουμε ότι

$$\nabla g(\boldsymbol{\mu}) = \begin{pmatrix} -k\mu_{k-1} \\ 0 \\ \vdots \\ \vdots \\ \vdots \\ 0 \\ 1 \end{pmatrix}. \quad (2.6)$$

Σημειώνοντας ότι $\sigma_{11} = \mu_2 - \mu_1^2 = \mu_2 = \sigma^2$, $\sigma_{1k} = \mu_{1+k} - \mu_1 \mu_k = \mu_{1+k}$, $\sigma_{kk} = \mu_k - \mu_k^2$, παρατηρούμε ότι

$$[\nabla g(\boldsymbol{\mu})]' \boldsymbol{\Sigma} [\nabla g(\boldsymbol{\mu})] = (-k\mu_{k-1} \ 0 \ \cdots \ \cdots \ 0 \ 1) \begin{pmatrix} \sigma_{11} & \cdots & \cdots & \cdots & \sigma_{1k} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \sigma_{k1} & \cdots & \cdots & \cdots & \sigma_{kk} \end{pmatrix} \begin{pmatrix} -k\mu_{k-1} \\ 0 \\ \vdots \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

$$= (-k\mu_{k-1}\sigma_{11} + \sigma_{k1} \ \cdots \ \cdots \ \cdots -k\mu_{k-1}\sigma_{1k} + \sigma_{kk}) \begin{pmatrix} -k\mu_{k-1} \\ 0 \\ \vdots \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

$$= k^2\mu_{k-1}^2\sigma_{11} - k\sigma_{k1}\mu_{k-1} - k\mu_{k-1}\sigma_{1k} + \sigma_{kk}$$

$$= k^2\mu_{k-1}^2\mu_2 - 2k\mu_{k-1}\mu_{k+1} + \mu_{2k} - \mu_k^2 = \mu_{2k} - \mu_k^2 + k\mu_{k-1}(k\mu_{k-1}\sigma^2 - 2\mu_{k+1})$$

$$= v_k^2.$$

Έτσι, η σχέση (2.4) γίνεται

$$\sqrt{n}(M_{k,n} - \mu_k) \xrightarrow{d} \mathcal{N}(0, v_k^2), \quad (2.7)$$

συνεπώς, η δειγματική κεντρική ροπή $M_{k,n}$ είναι ασυμπτωτικά κανονική εκτιμήτρια της κεντρικής ροπής μ_k .

Η παραπάνω τεχνική μπορεί εύκολα να εφαρμοστεί για να λάβουμε την από κοινού οριακή κατανομή, για παράδειγμα, της ποσότητας $(\bar{X} - \mu, M_k)'$ και, επίσης, της ποσότητας $(M_r, M_k)'$ για $2 \leq r < k$. Γενικότερα, η μέθοδος αυτή μας επιτρέπει να λάβουμε την k -διάστατη οριακή κατανομή της $\mathbf{M} = \mathbf{M}_n$, καθώς $n \rightarrow \infty$. Η γενική k -διάστατη περίπτωση είναι ήδη δημοσιευμένη στην εργασία του Αφένδρα (2013), μολονότι κάποιες ενδιαφέρουσες περιπτώσεις δεν έχουν αποσαφηνιστεί.

Πρώτα, ας θεωρήσουμε το ζεύγος $(M_r, M_k)'$ για $2 \leq r < k$. Σαφώς, $M_{k,n} = g(\mathbf{m}_n)$ με g όπως δίνεται από τη σχέση (2.3) και $M_{r,n} = h(\mathbf{m}_n)$, όπου $h: \mathcal{R}^k \rightarrow \mathcal{R}$ είναι Borel-μετρήσιμη συνάρτηση η οποία δίνεται από τον τύπο

$$h(x_1, x_2, \dots, x_r, \dots, x_k) = (-1)^{r-1}(r-1)x_1^r + \sum_{j=2}^{r-1} \binom{r}{j} (-1)^{r-j} x_1^{r-j} x_j + x_r. \quad (2.8)$$

Παρόμοια με πριν, θα προκύψει

$$\frac{\partial h(x_1, x_2, \dots, x_r, \dots, x_k)}{\partial x_i} = \begin{cases} \sum_{j=1}^{r-2} \binom{r}{j} (-1)^{r-j} x_1^{r-j-1} x_j (r-j) - r x_{r-1}, & \text{για } i = 1, \\ (-1)^{r-i} \binom{r}{i} x_1^{r-i}, & \text{για } i = 2, 3, \dots, r-1, \\ 1, & \text{για } i = r, \\ 0, & \text{για } i = r+1, \dots, k. \end{cases} \quad (2.9)$$

Επομένως, δεδομένου ότι $\mu_1 = 0$, θα ισχύει, $h(\boldsymbol{\mu}) = h(\mu_1, \mu_2, \dots, \mu_r, \dots, \mu_k) = \mu_r$ και έτσι θα έχουμε

$$\nabla h(\boldsymbol{\mu}) = \begin{pmatrix} -r\mu_{r-1} \\ 0 \\ \vdots \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ \vdots \\ 0 \end{pmatrix} \in \mathcal{R}^k. \quad (2.10)$$

Εφαρμόζοντας τη μέθοδο Δέλτα στη σχέση (2.2) θα πάρουμε

$$\sqrt{n} \begin{pmatrix} h(\mathbf{m}_n) - h(\boldsymbol{\mu}) \\ g(\mathbf{m}_n) - g(\boldsymbol{\mu}) \end{pmatrix} \xrightarrow{d} \mathcal{N} \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \mathcal{J}_{\boldsymbol{\mu}} \boldsymbol{\Sigma} \mathcal{J}_{\boldsymbol{\mu}}' \right), \quad (2.11)$$

όπου

$$\mathcal{J}_{\boldsymbol{\mu}} = \begin{pmatrix} [\nabla h(\boldsymbol{\mu})]' \\ [\nabla g(\boldsymbol{\mu})]' \end{pmatrix} \in \mathcal{R}^{2 \times k}.$$

Ωστόσο, χρησιμοποιώντας τις σχέσεις (2.6), (2.10) θα λάβουμε

$$\mathcal{J}_{\boldsymbol{\mu}} \boldsymbol{\Sigma} \mathcal{J}_{\boldsymbol{\mu}}' = \begin{pmatrix} [\nabla h(\boldsymbol{\mu})]' \boldsymbol{\Sigma} [\nabla h(\boldsymbol{\mu})] & [\nabla h(\boldsymbol{\mu})]' \boldsymbol{\Sigma} [\nabla g(\boldsymbol{\mu})] \\ [\nabla g(\boldsymbol{\mu})]' \boldsymbol{\Sigma} [\nabla h(\boldsymbol{\mu})] & [\nabla g(\boldsymbol{\mu})]' \boldsymbol{\Sigma} [\nabla g(\boldsymbol{\mu})] \end{pmatrix} = \begin{pmatrix} v_r^2 & v_{rk} \\ v_{kr} & v_k^2 \end{pmatrix},$$

όπου v_r^2 και v_k^2 είναι όπως στη σχέση (2.1) και

$$[\nabla g(\boldsymbol{\mu})]' \boldsymbol{\Sigma} [\nabla h(\boldsymbol{\mu})] = v_{kr} = v_{rk} = [\nabla h(\boldsymbol{\mu})]' \boldsymbol{\Sigma} [\nabla g(\boldsymbol{\mu})]$$

$$= (-r\mu_{r-1} \ 0 \ \cdots \ \cdots \ \cdots \ 0 \ 1 \ 0 \ \cdots \ \cdots \ \cdots \ 0) \begin{pmatrix} \sigma_{11} & \cdots & \sigma_{1r} & \cdots & \sigma_{1k} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \sigma_{r1} & \cdots & \sigma_{rr} & \cdots & \sigma_{rk} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \sigma_{k1} & \cdots & \sigma_{kr} & \cdots & \sigma_{kk} \end{pmatrix} \begin{pmatrix} -k\mu_{k-1} \\ 0 \\ \vdots \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

$$= (-r\mu_{r-1}\sigma_{11} + \sigma_{r1} \ \cdots \ \cdots \ \cdots \ \sigma_{rr} \ \cdots \ \cdots \ \cdots -r\mu_{r-1}\sigma_{1k} + \sigma_{rk}) \begin{pmatrix} -k\mu_{k-1} \\ 0 \\ \vdots \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

$$= rk\mu_{r-1}\mu_{k-1}\sigma_{11} - k\mu_{k-1}\sigma_{r1} - r\mu_{r-1}\sigma_{1k} + \sigma_{rk}$$

$$= rk\mu_{r-1}\mu_{k-1}\sigma^2 - k\mu_{k-1}\mu_{r+1} - r\mu_{r-1}\mu_{k+1} + \mu_{r+k} - \mu_r\mu_k \quad (2.12)$$

(παρατηρούμε ότι η v_{kk} δίνεται από την (2.1) διότι $v_{kk} = v_k^2$). Τελικά, από την (2.11) καταλήγουμε ότι για $2 \leq r < k$,

$$\sqrt{n} \begin{pmatrix} M_{r,n} - \mu_r \\ M_{k,n} - \mu_k \end{pmatrix} \xrightarrow{d} \mathcal{N} \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} v_r^2 & v_{rk} \\ v_{kr} & v_k^2 \end{pmatrix} \right). \quad (2.13)$$

Θεωρούμε τώρα, το ζεύγος $(M_{1,n}, M_{k,n})' = (\bar{X}_n - \mu, M_{k,n})'$, όπου εξ' ορισμού $M_{1,n} = \bar{X}_n - \mu$. Έπειτα, αντικαθιστώντας την συνάρτηση h από την $h(x_1, x_2, \dots, x_k) = x_1$ (με $\nabla h(x) = (1, 0, \dots, 0)' \in \mathcal{R}^k$) μπορούμε να εφαρμόσουμε τα ίδια επιχειρήματα με συνάρτηση g όπως δίνεται παραπάνω. Έτσι, εύκολα καταλήγουμε ότι

$$[\nabla h(\boldsymbol{\mu})]' \boldsymbol{\Sigma} [\nabla h(\boldsymbol{\mu})] = (1 \ 0 \ \cdots \ \cdots \ \cdots \ 0) \begin{pmatrix} \sigma_{11} & \cdots & \cdots & \cdots & \sigma_{1k} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \sigma_{k1} & \cdots & \cdots & \cdots & \sigma_{kk} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ \vdots \\ 0 \end{pmatrix}$$

$$= (\sigma_{11} \dots \sigma_{1k}) \begin{pmatrix} 1 \\ 0 \\ \vdots \\ \vdots \\ 0 \end{pmatrix} = \sigma_{11} = \mu_2 = \sigma^2,$$

$$[\nabla h(\boldsymbol{\mu})]' \boldsymbol{\Sigma} [\nabla g(\boldsymbol{\mu})] = (1 \ 0 \ \dots \ 0) \begin{pmatrix} \sigma_{11} & \dots & \dots & \dots & \sigma_{1k} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \sigma_{k1} & \dots & \dots & \dots & \sigma_{kk} \end{pmatrix} \begin{pmatrix} -k\mu_{k-1} \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

$$= (\sigma_{11} \dots \sigma_{1k}) \begin{pmatrix} -k\mu_{k-1} \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} = -k\sigma^2\mu_{k-1} + \mu_{k+1}.$$

Επομένως,

$$\sqrt{n} \begin{pmatrix} \bar{X}_n - \mu \\ M_{k,n} - \mu_k \end{pmatrix} \xrightarrow{d} \mathcal{N} \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2 & \mu_{k+1} - k\sigma^2\mu_{k-1} \\ \mu_{k+1} - k\sigma^2\mu_{k-1} & v_k^2 \end{pmatrix} \right). \quad (2.14)$$

Όλα τα προηγούμενα ασυμπτωτικά αποτελέσματα αποτελούν ειδικές περιπτώσεις της ακόλουθης πρότασης.

Πρόταση 2.1: Αν $E(X^{2k}) < \infty$ τότε

$$\sqrt{n} \begin{pmatrix} \bar{X}_n - \mu \\ M_{2,n} - \sigma^2 \\ M_{3,n} - \mu_3 \\ \vdots \\ M_{k,n} - \mu_k \end{pmatrix} \xrightarrow{d} \mathcal{N} \left(\begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2 & \mu_3 & \mu_4 - 3\sigma^2 & \dots & \mu_{k+1} - k\sigma^2\mu_{k-1} \\ \mu_3 & v_2^2 & v_{2,3} & \dots & v_{2,k} \\ \mu_4 - 3\sigma^2 & v_{2,3} & v_3^2 & \dots & v_{3,k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mu_{k+1} - k\sigma^2\mu_{k-1} & v_{2,k} & v_{3,k} & \dots & v_k^2 \end{pmatrix} \right). \quad (2.15)$$

Απόδειξη: Η απόδειξη είναι προφανής αφού μπορούμε να εφαρμόσουμε τη μέθοδο Δέλτα για το διάνυσμα των *Borel*–μετρησίμων συναρτήσεων $\mathbf{g} = (g_1, g_2, \dots, g_k)': \mathcal{R}^k \rightarrow \mathcal{R}^k$, όπου οι g_r είναι κατάλληλα ορισμένες συναρτήσεις ούτως ώστε $g_r(\mathbf{m}_n) = M_{r,n}$, $r = 1, 2, \dots, k$, δηλαδή, g_k είναι ορισμένες όπως στην (2.3), g_r όπως στην (2.8) (για $r \in \{2, 3, \dots, k-1\}$) και $g_1(x) = x_1$. Το επιθυμητό αποτέλεσμα είναι άμεση εφαρμογή της (2.2) και του γεγονότος ότι (βλ., *Van der Vaart* (1998), Θεώρημα 3.1, βλ. και Θεώρημα 2.6, παρακάτω)

$$\sqrt{n}(\mathbf{g}(\mathbf{m}_n) - \mathbf{g}(\boldsymbol{\mu})) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \mathbf{V}),$$

όπου $\mathbf{0} = (0, 0, \dots, 0)'$ και $\mathbf{V} = \mathbf{J}_\mu \boldsymbol{\Sigma} \mathbf{J}_\mu' \in \mathcal{R}^{k \times k}$ με $\mathbf{J}_\mu = (\partial g_i(x)/\partial x_j)|_{x=\mu}$.

■

Παρατήρηση: Η σχέση (2.11) είναι υποπερίπτωση της (2.15) για την *Borel*–μετρήσιμη συνάρτηση $\mathbf{f} = (h, g)': \mathcal{R}^k \rightarrow \mathcal{R}^2$.

2.3 Σύγκλιση ροπών

Είναι γνωστό ότι η ασθενής σύγκλιση στην (2.15) δεν συνεπάγεται τη σύγκλιση των αντίστοιχων ροπών. Συγκεκριμένα, ενδέχεται να μην υφίσταται κατ' ανάγκην, είτε το όριο $E[\sqrt{n}(M_{k,n} - \mu_k)] \rightarrow 0$, καθώς $n \rightarrow \infty$, είτε το όριο $Var[\sqrt{n}(M_{k,n} - \mu_k)] \rightarrow v_k^2$, καθώς $n \rightarrow \infty$. Ωστόσο, το γεγονός ότι η (2.15) υποδεικνύει ότι τα όρια των αντίστοιχων αναμενόμενων τιμών, διασπορών και συνδιακυμάνσεων είναι καλώς–ορισμένα, καθιστά την περίπτωση αυτή πολύ ενδιαφέρουσα. Οι τρεις Προτάσεις που ακολουθούν ισχυρίζονται ότι αυτές οι συγκλίσεις πραγματικά πληρούνται, όταν το ελάχιστο (φυσικό) σύνολο των υποθέσεων επιβάλλεται πάνω στις ροπές της X . Πρώτα όμως, παραθέτουμε το Θεώρημα 2.1 και το Λήμμα 2.3 τα οποία διευκολύνουν τις αποδείξεις των τριών αυτών Προτάσεων.

Θεώρημα 2.1: Έστω $X, X_1, X_2, \dots, X_n$ μία ακολουθία ανεξάρτητων και ισόνομων τυχαίων μεταβλητών με αναμενόμενη τιμή $E(X) = \mu \in (-\infty, +\infty)$, πεπερασμένη διασπορά $Var(X) = \sigma^2 > 0$ και $E|X|^\delta < \infty$ για κάποιο $\delta \geq 2$. Αν $S_n = X_1 + X_2 + \dots + X_n$ τότε η ακολουθία $\left\{ \left| \frac{S_n - n\mu}{\sigma\sqrt{n}} \right|^\delta \right\}_{n=1}^\infty$ είναι ομοιόμορφα ολοκληρώσιμη.

Το Θεώρημα αυτό είναι πολύ σημαντικό επειδή συνεπάγεται τη σύγκλιση των ροπών, δηλαδή, ότι για οποιονδήποτε εκθέτη $\alpha \in (0, \delta]$,

$$\lim_{n \rightarrow \infty} E \left| \frac{S_n - n\mu}{\sqrt{n}} \right|^\alpha = \sigma^\alpha E|Z|^\alpha, \quad (2.16)$$

όπου Z η τυπική κανονική τ.μ.²⁸. Γνωρίζουμε ότι αν η ακολουθία $\{Y_n\}_{n=1}^\infty$ είναι ομοιόμορφα ολοκληρώσιμη τότε $\sup_{n \geq 1} E|Y_n| < \infty$. Επίσης, ξέρουμε ότι αν μπορούμε να βρούμε μία σταθερά $\gamma > 0$ έτσι ώστε $\sup_{n \geq 1} E|Y_n|^{1+\gamma} < \infty$ τότε η ακολουθία $\{Y_n\}_{n=1}^\infty$ είναι ομοιόμορφα ολοκληρώσιμη. Έτσι, το Θεώρημα 2.1 συνεπάγεται ότι για οποιοδήποτε $\alpha \in (0, \delta)$, η ακολουθία

$$\left\{ \left| \frac{S_n - n\mu}{\sigma\sqrt{n}} \right|^\alpha \right\}_{n=1}^\infty$$

είναι ομοιόμορφα ολοκληρώσιμη. Πράγματι, για $\gamma = \frac{\delta-\alpha}{\alpha} > 0$ έχουμε

$$\sup_{n \geq 1} E \left| \frac{S_n - n\mu}{\sigma\sqrt{n}} \right|^\alpha^{1+\gamma} = \sup_{n \geq 1} E \left| \frac{S_n - n\mu}{\sigma\sqrt{n}} \right|^\delta < \infty,$$

από το Θεώρημα 2.1.

Είναι ολοφάνερο ότι για $\alpha = \delta$, η (2.16) προκύπτει από το Κεντρικό Οριακό Θεώρημα και το Θεώρημα 2.1. Έτσι, η (2.16) πληρούται για οποιοδήποτε $\alpha \in (0, \delta]$. Αφετέρου, το Θεώρημα της ομοιόμορφης ολοκληρωσιμότητας και το γεγονός ότι

$$0 \leq \left| \frac{S_n - n\mu}{\sigma\sqrt{n}} \right|^\delta \xrightarrow{d} |Z|^\delta,$$

σε συνδυασμό με το Θεώρημα *Skorokhod*, δείχνουν ότι η (2.16) για $\alpha = \delta$ ικανοποιείται εάν και μόνο εάν η ακολουθία $\left\{ \left| \frac{S_n - n\mu}{\sigma\sqrt{n}} \right|^\delta \right\}_{n=1}^\infty$ είναι ομοιόμορφα ολοκληρώσιμη.

Βασιζόμενοι στις προηγούμενες παρατηρήσεις, καταλήγουμε ότι η σχέση (2.16) ισχύει για όλα τα $\alpha \in (0, \delta]$ αν και μόνο αν το Θεώρημα 2.1 είναι αληθές. Στη συνέχεια, θα προβούμε στην αυστηρή απόδειξη αυτού του Θεωρήματος στηριζόμενοι στην ανισότητα *Marcinkiewicz*²⁹ – *Zygmund*³⁰.

²⁸Ως γνωστόν, $E|Z|^p = \int_{-\infty}^{+\infty} |z|^p \varphi(z) dz = 2 \int_0^{+\infty} |z|^p \varphi(z) dz = 2 \int_{-\infty}^0 |z|^p \varphi(z) dz$, αφού, η $|z|^p \varphi(z)$ είναι άρτια συνάρτηση και $\varphi(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$.

²⁹Ο *Józef Marcinkiewicz* (30 Μαρτίου 1910 - 1940) υπήρξε μαθηματικός Πολωνικής καταγωγής και μαθητής του *Zygmund*. Ο ακριβής τόπος και η ημερομηνία θανάτου παραμένουν άγνωστα, εικάζεται όμως, ότι σκοτώθηκε στο *Χάρκοβο* της *Ουκρανίας*. Οι γονείς του, στους οποίους έδωσε χειρόγραφες μαθηματικές εργασίες που έγραψε πριν από την έναρξη του Β' Παγκοσμίου Πολέμου, είχαν μεταφερθεί, το 1940, στην *Σοβιετική Ένωση* όπου αργότερα αμφότεροι απεβίωσαν σε ένα στρατόπεδο λόγω πείνας, και έτσι, δε βρέθηκε ποτέ κανένα ίχνος των χειρογράφων αυτών.

Θεώρημα 2.2: (ανισότητα Marcinkiewicz – Zygmund)

Για κάθε $p \geq 1$ μπορούμε να βρούμε πεπερασμένες θετικές σταθερές A_p και B_p εξαρτώμενες αποκλειστικά και μόνο από το p έτσι ώστε για κάθε $n \in \{1, 2, \dots\}$, και για οποιεσδήποτε ανεξάρτητες τυχαίες μεταβλητές $X_1, X_2, \dots, X_n$ με $E|X_i|^p < \infty$ και $E(X_i) = 0$, $i = 1, 2, \dots, n$,

$$A_p E \left[\left(\sum_{i=1}^n X_i^2 \right)^{p/2} \right] \leq E \left| \sum_{i=1}^n X_i \right|^p \leq B_p E \left[\left(\sum_{i=1}^n X_i^2 \right)^{p/2} \right]. \quad (2.17)$$

Στην ουσία, χρειαζόμαστε μόνο το άνω φράγμα της (2.17). Συγκεκριμένα, για την απόδειξη του Θεωρήματος 2.1, το μόνο που χρειαζόμαστε από την ανισότητα Marcinkiewicz – Zygmund είναι το άνω φράγμα της σχέσης (2.20) η οποία παρατίθεται στο Πρόγραμμα 2.1, παρακάτω.

Λήμμα 2.1: Για οποιουδήποτε πραγματικούς αριθμούς $x_1, x_2, \dots, x_n$ και για τυχόν $p \geq 1$,

$$\left(\sum_{i=1}^n x_i^2 \right)^p \leq n^{p-1} \sum_{i=1}^n |x_i|^{2p}. \quad (2.18)$$

Απόδειξη: Έστω Y τυχαία μεταβλητή που ακολουθεί την διακριτή ομοιόμορφη κατανομή, δηλαδή, $Y \sim U\{1, 2, \dots, n\}$, όπου $P(Y = i) = 1/n$, $i = 1, 2, \dots, n$ και θέτουμε $X = g(Y)$ όπου $g(i) = x_i^2$, $i = 1, 2, \dots, n$. Τότε σύμφωνα με την ανισότητα του *Lyapunov*³¹ θα έχουμε

$$E|X| \leq (E|X|^p)^{1/p}, \quad p \geq 1,$$

ήτοι,

$$E|g(Y)| \leq (E|g(Y)|^p)^{1/p}, \quad p \geq 1. \quad (2.19)$$

Όμως,

³⁰Ο *Antoni Zygmund* (25 Δεκεμβρίου 1900 – 30 Μαΐου 1992) ήταν Πολωνός μαθηματικός με την αρμονική ανάλυση να αποτελεί την κύρια περιοχή του ενδιαφέροντός του. Μάλιστα, θεωρείται ένας από τους μεγαλύτερους αναλύστες του 20^{ου} αιώνα.

³¹Ο *Aleksandr Mikhailovich Lyapunov* (6 Ιουνίου 1857 - 3 Νοεμβρίου 1918) υπήρξε Ρώσος μαθηματικός, φυσικός και μηχανικός. Απέκτησε φήμη για την ανάπτυξη της θεωρίας της σταθερότητας ενός δυναμικού συστήματος, καθώς και για τις πολλές συνεισφορές του στην μαθηματική φυσική και στη θεωρία πιθανοτήτων.

$$E|g(Y)| = \sum_{i=1}^n |g(i)| P(Y = i) = \frac{1}{n} \sum_{i=1}^n |g(i)| = \frac{1}{n} \sum_{i=1}^n x_i^2$$

και, παρόμοια,

$$E|g(Y)|^p = \frac{1}{n} \sum_{i=1}^n |x_i|^{2p}.$$

Έτσι, αντικαθιστώντας στην (2.19), παίρνουμε

$$\frac{1}{n} \sum_{i=1}^n x_i^2 \leq \left(\frac{1}{n} \sum_{i=1}^n |x_i|^{2p} \right)^{1/p}, \quad p \geq 1,$$

δηλαδή,

$$\left(\sum_{i=1}^n x_i^2 \right)^p \leq n^{p-1} \sum_{i=1}^n |x_i|^{2p}.$$

■

Πόρισμα 2.1: Αν $p \geq 2$ για οποιεσδήποτε ανεξάρτητες και ισόνομες τυχαίες μεταβλητές $X_1, X_2, \dots, X_n$ με $E|X|^p < \infty$ και $E(X) = 0$, ισχύει

$$E \left| \sum_{i=1}^n X_i \right|^p \leq B_p n^{p/2} E|X|^p, \quad (2.20)$$

όπου B_p όπως στο Θεώρημα 2.2.

Απόδειξη: Εφόσον, $X, X_1, X_2, \dots, X_n$ είναι ισόνομα κατανομημένες τυχαίες μεταβλητές θα ισχύει

$$E \left[\sum_{i=1}^n |X_i|^p \right] = \sum_{i=1}^n E|X_i|^p = nE|X|^p.$$

Έτσι, από τις σχέσεις (2.17), (2.18) συμπεραίνουμε ότι

$$E \left| \sum_{i=1}^n X_i \right|^p \leq B_p E \left[\left(\sum_{i=1}^n X_i^2 \right)^{p/2} \right] \leq B_p n^{p/2-1} E \left[\sum_{i=1}^n |X_i|^p \right] = B_p n^{p/2} E|X|^p.$$

■

Κατόπιν, πρέπει να χρησιμοποιήσουμε τις ακόλουθες ανισότητες.

Λήμμα 2.2:

(i) Αν x, y είναι πραγματικοί αριθμοί και $p > 0$ τότε

$$|x + y|^p \leq 2^p |x|^p + 2^p |y|^p. \quad (2.21)$$

(ii) Αν U και V είναι δύο αυθαίρετες τυχαίες μεταβλητές, τότε για οποιοδήποτε $\alpha > 0$ και για τυχόν $p > 0$,

$$E[|U + V|^p I(|U + V| > 2\alpha)] \leq 2^p E[|U|^p I(|U| > \alpha)] + 2^p E[|V|^p I(|V| > \alpha)]. \quad (2.22)$$

(iii) Για οποιαδήποτε τυχαία μεταβλητή W και τυχούσες σταθερές $p > 0$ και $\beta > 0$ ισχύει

$$E[|W|^p I(|W| > \beta)] \leq \beta^{-p} E|W|^{2p}. \quad (2.23)$$

Απόδειξη:

(i) Από την τριγωνική ανισότητα, έχουμε

$$|x + y| \leq |x| + |y| = \min\{|x|, |y|\} + \max\{|x|, |y|\} \leq 2 \max\{|x|, |y|\},$$

και υψώνοντας στο p θα λάβουμε

$$|x + y|^p \leq (2 \max\{|x|, |y|\})^p = 2^p \max\{|x|^p, |y|^p\} \leq 2^p (|x|^p + |y|^p).$$

(ii) Αν $E|U|^p = \infty$ τότε $E[|U|^p I(|U| > \alpha)] = \infty$ για όλα τα $\alpha > 0$, και δεν έχουμε να αποδείξουμε τίποτα, διότι το δεξιό μέρος της (2.22) γίνεται ∞ . Παρόμοια, αν $E|V|^p = \infty$ τότε, και πάλι, το δεξιό μέρος της (2.22) γίνεται ∞ . Ως εκ τούτου, δεν υπάρχει απώλεια της γενικότητας αν υποθέσουμε ότι $E|U|^p < \infty$ και $E|V|^p < \infty$. Τώρα, αρκεί να δείξουμε ότι

$$|U + V|^p I(|U + V| > 2\alpha) \leq 2^p |U|^p I(|U| > \alpha) + 2^p |V|^p I(|V| > \alpha), \quad (2.24)$$

αφού, παίρνοντας μέσες τιμές στην (2.24), καταλήγουμε στην (2.22). Αν $|U| \leq \alpha$ και $|V| \leq \alpha$ τότε $|U + V| \leq |U| + |V| \leq 2\alpha$, και η (2.24) πληρούται κατά τον τετριμμένο τρόπο $0 = 0$. Αν $|U| > \alpha$ και $|V| > \alpha$ τότε το δεξιό μέρος της (2.24) ισούται με $2^p |U|^p + 2^p |V|^p$, ενώ το αριστερό μέρος είναι το πολύ $|U + V|^p$, και τελικά, η (2.24) έπεται από την (2.21) στην περίπτωση αυτή. Έπειτα, υποθέτουμε ότι $|U| \leq \alpha < |V|$. Εδώ, το δεξιό μέρος της (2.24) ισούται με $2^p |V|^p$, ενώ, εφόσον $|U| < |V|$, το αριστερό μέρος είναι το πολύ $|U + V|^p \leq (|U| + |V|)^p \leq (|V| + |V|)^p = 2^p |V|^p$. Επομένως, η (2.24) και πάλι

ικανοποιείται. Τέλος, η περίπτωση $|V| \leq a < |U|$ είναι παρεμφερής με την προηγούμενη, και έτσι, η απόδειξη είναι πλήρης.

(iii) Η ανισότητα αληθεύει, αρκεί να πάρουμε τις αναμενόμενες τιμές στην κάτωθι σχέση:

$$|W|^p I(|W| > \beta) \leq \beta^{-p} |W|^{2p} I(|W| > \beta) \leq \beta^{-p} |W|^{2p}. \quad \blacksquare$$

Είμαστε σε θέση πλέον, να επαληθεύσουμε το Θεώρημα 2.1.

Απόδειξη Θεωρήματος 2.1: Θέτοντας, $Y = (X - \mu)/\sigma$ και $Y_i = (X_i - \mu)/\sigma$, $i \geq 1$, παρατηρούμε ότι

$$\left| \frac{S_n - n\mu}{\sqrt{n}} \right|^\delta = \left| \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{X_i - \mu}{\sigma} \right|^\delta = \left| \frac{1}{\sqrt{n}} \sum_{i=1}^n Y_i \right|^\delta,$$

και οι τυχαίες μεταβλητές $Y, Y_1, Y_2, \dots, Y_n$ είναι ανεξάρτητες και ισόνομες με μέσο 0 και διασπορά 1. Επιπλέον, $E|X|^\delta < \infty$ εάν και μόνο εάν $E|Y|^\delta < \infty$. Πράγματι, επικαλούμενοι την (2.21) βλέπουμε ότι

$$|Y|^\delta = \sigma^{-\delta} |X - \mu|^\delta \leq \sigma^{-\delta} 2^\delta |X|^\delta + \sigma^{-\delta} 2^\delta |\mu|^\delta,$$

άρα, $E|Y|^\delta < \infty$ όταν $E|X|^\delta < \infty$,

και

$$|X|^\delta = |\sigma Y + \mu|^\delta \leq 2^\delta \sigma^\delta |Y|^\delta + 2^\delta |\mu|^\delta,$$

οπότε, $E|X|^\delta < \infty$ όταν $E|Y|^\delta < \infty$.

Έτσι, χωρίς βλάβη της γενικότητας θεωρούμε ότι $\mu = 0$, $\sigma = 1$. Συνεπώς, από' δω και στο εξής, υποθέτουμε ότι $\mu = E(X) = 0$, $\sigma^2 = E(X^2) = 1$ και $E|X|^\delta < \infty$ για τυχόν $\delta \geq 2$, και ας είναι $S_n = X_1 + X_2 + \dots + X_n$. Ο στόχος μας είναι να αποδείξουμε ότι η ακολουθία $\left\{ |n^{-1/2} S_n|^\delta \right\}_{n=1}^\infty$ είναι ομοιόμορφα ολοκληρώσιμη, με άλλα λόγια, θα δείξουμε ότι

$$g(\alpha) := \sup_{n \geq 1} \left\{ \frac{1}{n^{\delta/2}} E[|S_n|^\delta I(|S_n|^\delta > n^{\delta/2} \alpha)] \right\} \rightarrow 0, \text{ καθώς } \alpha \rightarrow \infty. \quad (2.25)$$

Παρατηρώντας ότι $I(|S_n|^\delta > n^{\delta/2} \alpha) = I(n^{-1/2} |S_n| > \alpha^{1/\delta})$ και αφού $g(\alpha) \rightarrow 0$, καθώς $\alpha \rightarrow \infty$, αν και μόνο αν $g(\alpha^{1/\delta}) \rightarrow 0$, καθώς $\alpha \rightarrow \infty$, μπορούμε να γράψουμε την σχέση (2.25) ως $\lim_{\alpha \rightarrow \infty} g(\alpha^{1/\delta}) = 0$. Οπότε, επιδιώκουμε να επιβεβαιώσουμε το ακόλουθο όριο:

$$h(a) := \sup_{n \geq 1} E \left[\left| \frac{S_n}{\sqrt{n}} \right|^\delta I \left(\left| \frac{S_n}{\sqrt{n}} \right| > a \right) \right] \rightarrow 0, \text{ καθώς } a \rightarrow \infty. \quad (2.26)$$

Δηλαδή, θα αποδείξουμε ότι $\lim_{a \rightarrow \infty} h(a) = 0$.

Προς αυτή την κατεύθυνση, λοιπόν, για μία αυθαίρετη σταθερά $\lambda > 0$ (που θα καθοριστεί αργότερα) ορίζουμε τις τυχαίες μεταβλητές

$$X'_i := X_i I(|X_i| \leq \lambda) - E[X_i I(|X_i| \leq \lambda)] \text{ και } X''_i := X_i - X'_i, \quad i = 1, 2, \dots,$$

και θέτουμε

$$S'_n := \sum_{i=1}^n X'_i \text{ και } S''_n := \sum_{i=1}^n X''_i,$$

ούτως ώστε $S_n = S'_n + S''_n$.

Παρατηρούμε ότι

$$E(X'_i) = E(X''_i) = E(S'_n) = E(S''_n) = 0 \text{ για όλα τα } n \text{ και } i,$$

και ότι οι τυχαίες μεταβλητές $X'_1, X'_2, \dots, X'_n$ είναι ανεξάρτητες και ισόνομα κατανεμημένες, και μάλιστα, το ίδιο ισχύει και για τις τ.μ. $X''_1, X''_2, \dots, X''_n$. Συγκεκριμένα, κάθε X'_i έχει την ίδια κατανομή όπως η

$$X' := XI(|X| \leq \lambda) - E[XI(|X| \leq \lambda)]$$

και κάθε X''_i έχει την ίδια κατανομή με την

$$X'' := XI(|X| > \lambda) - E[XI(|X| > \lambda)].$$

Βλέπουμε επίσης ότι $|X'_i| \leq 2\lambda$ και $|X'| \leq 2\lambda$.

Εφαρμόζοντας την (2.20) για τις τυχαίες μεταβλητές $X', X'_1, X'_2, \dots, X'_n$ και για $p = 2\delta \geq 2$, λαμβάνουμε ότι

$$E|S'_n|^{2\delta} \leq B_{2\delta} n^\delta E|X'|^{2\delta} \leq B_{2\delta} n^\delta (2\lambda)^{2\delta}. \quad (2.27)$$

Από τη άλλη, εφαρμόζοντας την (2.23) για την τυχαία μεταβλητή $W = n^{-1/2} S'_n$ και για $p = \delta$, προκύπτει

$$E \left[\left| \frac{S'_n}{\sqrt{n}} \right|^\delta I \left(\left| \frac{S'_n}{\sqrt{n}} \right| > \beta \right) \right] \leq \beta^{-\delta} E \left[\left| \frac{S'_n}{\sqrt{n}} \right|^{2\delta} \right] = (n\beta)^{-\delta} E|S'_n|^{2\delta}, \quad (2.28)$$

όπου β είναι μία σταθερά που θα προσδιοριστεί προσεχώς. Συνδυάζοντας την (2.27) και την (2.28),

$$E \left[\left| \frac{S'_n}{\sqrt{n}} \right|^\delta I \left(\left| \frac{S'_n}{\sqrt{n}} \right| > \beta \right) \right] \leq (n\beta)^{-\delta} E|S'_n|^{2\delta} \leq (n\beta)^{-\delta} B_{2\delta} n^\delta (2\lambda)^{2\delta},$$

και, συνεπώς,

$$E \left[\left| \frac{S'_n}{\sqrt{n}} \right|^\delta I \left(\left| \frac{S'_n}{\sqrt{n}} \right| > \beta \right) \right] \leq \frac{B_{2\delta} (2\lambda)^{2\delta}}{\beta^\delta}, \quad \beta > 0, \quad \lambda > 0. \quad (2.29)$$

[Το σημαντικό σημείο είναι ότι το άνω φράγμα της (2.29) δεν εξαρτάται από το n .]

Τώρα, εφαρμόζοντας ξανά την (2.20) για τις τυχαίες μεταβλητές $X'', X''_1, X''_2, \dots, X''_n$ και για $p = \delta \geq 2$, απορρέει ότι

$$E|S''_n|^\delta \leq B_\delta n^{\delta/2} E|X''|^\delta. \quad (2.30)$$

Υπενθυμίζουμε ότι $X'' := XI(|X| > \lambda) - E[XI(|X| > \lambda)]$. Επιπλέον, από την ανισότητα³² του *Lyapounov*, θα έχουμε

$$|E[XI(|X| > \lambda)]|^\delta \leq (E[|X|I(|X| > \lambda)])^\delta \leq E[|X|^\delta I(|X| > \lambda)], \quad (2.31)$$

και η (2.31), μαζί με την (2.21), οδηγεί στην

$$\begin{aligned} |X''|^\delta &\leq 2^\delta |X|^\delta I(|X| > \lambda) + 2^\delta |E[XI(|X| > \lambda)]|^\delta \\ &\leq 2^\delta |X|^\delta I(|X| > \lambda) + 2^\delta E[|X|^\delta I(|X| > \lambda)]. \end{aligned} \quad (2.32)$$

Λαμβάνοντας μέσες τιμές στην (2.32) διαπιστώνουμε ότι

$$E|X''|^\delta \leq 2^\delta E[|X|^\delta I(|X| > \lambda)] + 2^\delta E[|X|^\delta I(|X| > \lambda)] = 2^{\delta+1} E[|X|^\delta I(|X| > \lambda)], \quad (2.33)$$

και η (2.30), σε συνδυασμό με την (2.33), οδηγεί στην ανισότητα

$$E|S''_n|^\delta \leq B_\delta n^{\delta/2} 2^{\delta+1} E[|X|^\delta I(|X| > \lambda)]. \quad (2.34)$$

Ισοδύναμα,

$$E \left| \frac{S''_n}{\sqrt{n}} \right|^\delta \leq B_\delta 2^{\delta+1} E[|X|^\delta I(|X| > \lambda)], \quad (2.35)$$

και αφού για οποιοδήποτε $\beta > 0$,

$$\left| \frac{S''_n}{\sqrt{n}} \right|^\delta I \left(\left| \frac{S''_n}{\sqrt{n}} \right| > \beta \right) \leq \left| \frac{S''_n}{\sqrt{n}} \right|^\delta,$$

συμπεραίνουμε ότι

³²Για τυχόν $p \geq 1$ και για οποιαδήποτε τυχαία μεταβλητή Y ισχύει $(E|Y|)^p \leq E|Y|^p$.

$$E \left[\left| \frac{S_n''}{\sqrt{n}} \right|^\delta I \left(\left| \frac{S_n''}{\sqrt{n}} \right| > \beta \right) \right] \leq B_\delta 2^{\delta+1} E[|X|^\delta I(|X| > \lambda)], \quad \beta > 0, \quad \lambda > 0. \quad (2.36)$$

[Και πάλι, το σπουδαίο γεγονός είναι ότι η τελευταία ανισότητα φράσσεται από μία ποσότητα ανεξάρτητη του n .]

Τώρα, παρατηρούμε ότι

$$\left| \frac{S_n}{\sqrt{n}} \right|^\delta I \left(\left| \frac{S_n}{\sqrt{n}} \right| > 2\beta \right) = \left| \frac{S_n' + S_n''}{\sqrt{n}} \right|^\delta I \left(\left| \frac{S_n' + S_n''}{\sqrt{n}} \right| > 2\beta \right).$$

Έτσι, παίρνοντας αναμενόμενες τιμές στην προηγούμενη ανισότητα και χρησιμοποιώντας την (2.22) με $\alpha = \beta$, $p = \delta$, $U = n^{-1/2}S_n'$ και $V = n^{-1/2}S_n''$, προκύπτει ότι

$$\begin{aligned} & E \left[\left| \frac{S_n}{\sqrt{n}} \right|^\delta I \left(\left| \frac{S_n}{\sqrt{n}} \right| > 2\beta \right) \right] \\ & \leq 2^\delta E \left[\left| \frac{S_n'}{\sqrt{n}} \right|^\delta I \left(\left| \frac{S_n'}{\sqrt{n}} \right| > \beta \right) \right] + 2^\delta E \left[\left| \frac{S_n''}{\sqrt{n}} \right|^\delta I \left(\left| \frac{S_n''}{\sqrt{n}} \right| > \beta \right) \right]. \end{aligned} \quad (2.37)$$

Λαμβάνοντας υπόψη τις σχέσεις (2.29), (2.36) και (2.37) καταλήγουμε ότι για οποιοδήποτε $\beta > 0$ και για οποιοδήποτε $\lambda > 0$,

$$E \left[\left| \frac{S_n}{\sqrt{n}} \right|^\delta I \left(\left| \frac{S_n}{\sqrt{n}} \right| > 2\beta \right) \right] \leq \frac{B_{2\delta} 2^{3\delta} \lambda^{2\delta}}{\beta^\delta} + B_\delta 2^{2\delta+1} E[|X|^\delta I(|X| > \lambda)]. \quad (2.38)$$

Εφόσον, το δεξιό μέρος της (2.38) δεν εξαρτάται από το n , συμπεραίνουμε ότι για οποιεσδήποτε σταθερές $\beta > 0, \lambda > 0$, ισχύει

$$\sup_{n \geq 1} E \left[\left| \frac{S_n}{\sqrt{n}} \right|^\delta I \left(\left| \frac{S_n}{\sqrt{n}} \right| > 2\beta \right) \right] \leq \frac{B_{2\delta} 2^{3\delta} \lambda^{2\delta}}{\beta^\delta} + B_\delta 2^{2\delta+1} E[|X|^\delta I(|X| > \lambda)], \quad (2.39)$$

δηλαδή,

$$h(2\beta) \leq \frac{B_{2\delta} 2^{3\delta} \lambda^{2\delta}}{\beta^\delta} + B_\delta 2^{\delta+1} E[|X|^\delta I(|X| > \lambda)], \quad \beta > 0, \quad \lambda > 0 \quad (2.40)$$

(βλ. σχέση (2.26) για το πως ορίζεται η συνάρτηση h). Τώρα, επειδή αμφότερες οι τιμές των $\beta > 0, \lambda > 0$ είναι αυθαίρετες, μπορούμε να αντικαταστήσουμε στην (2.40) όπου $\beta = \alpha/2$ και όπου $\lambda = \alpha^{1/4}$. Τοιουτοτρόπως, θα προκύψει

$$0 \leq h(\alpha) \leq \frac{B_{2\delta} 2^{4\delta}}{\alpha^{\delta/2}} + B_\delta 2^{2\delta+1} E[|X|^\delta I(|X| > \alpha^{1/4})], \quad \alpha > 0, \quad (2.41)$$

και αφού, $E|X|^\delta < \infty$, αποφαινόμεσθε, εξαιτίας της κυριαρχημένης σύγκλισης, ότι

$$\lim_{\alpha \rightarrow \infty} E[|X|^\delta I(|X| > \alpha^{1/4})] = 0.$$

Η απόδειξη ολοκληρώνεται παίρνοντας όρια στην τελευταία ανισότητα, καθώς $\alpha \rightarrow \infty$. ■

Λήμμα 2.3: Αν $X, X_1, X_2, \dots, X_n$ είναι ανεξάρτητες και ισόνομες τυχαίες μεταβλητές με $E(X) = \mu$ και $E|X|^v < \infty$ για κάποιο $v \in \{2, 3, \dots\}$ τότε για οποιοδήποτε $j \in \{2, 3, \dots, v\}$ ισχύει,

$$E|m_{j,n}|^{v/j} \leq E|X - \mu|^v, \quad (2.42)$$

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

Απόδειξη:

• Αν $j = v$ τότε $|m_{j,n}| \leq \frac{1}{n} \sum_{i=1}^n |X_i - \mu|^j = \frac{1}{n} \sum_{i=1}^n |X_i - \mu|^v$ και παίρνοντας την αναμενόμενη μέση τιμή και στα δύο μέλη της ανίσωσης θα προκύψει τελικά

$$E|m_{j,n}|^{v/j} \leq E|X - \mu|^v.$$

• Αν $j < v$ (και έτσι, $v \geq 3$) τότε από την ανισότητα του Hölder³³ έχουμε

$$\left| \sum_{i=1}^n x_i \right|^p \leq \left(\sum_{i=1}^n |x_i| \right)^p \leq n^{p-1} \sum_{i=1}^n |x_i|^p, \quad p > 1.$$

Θέτοντας $p = v/j$ και $x_i = (X_i - \mu)^j$ θα λάβουμε διαδοχικά

$$E|m_{j,n}|^{v/j} = \frac{1}{n^{v/j}} E \left| \sum_{i=1}^n (X_i - \mu)^j \right|^{v/j} \leq \frac{1}{n^{v/j}} E \left[\sum_{i=1}^n |X_i - \mu|^j \right]^{v/j}$$

³³Ο Otto Ludwig Hölder (22 Δεκεμβρίου 1859 – 29 Αυγούστου 1937) ήταν Γερμανός μαθηματικός γεννημένος στη Στουτγάρδη. Το 1877, εισήλθε στο Πανεπιστήμιο του Βερολίνου και πήρε το διδακτορικό του από το Πανεπιστήμιο του Tübingen το 1882. Ο τίτλος της διδακτορικής διατριβής του ήταν «Beiträge zur Potentialtheorie» ("Συμβολή στην θεωρία δυναμικού"). Εργάστηκε στο Πανεπιστήμιο της Λειψίας από το 1899 μέχρι τη συνταξιοδότησή του. Είναι διάσημος για πολλά πράγματα όπως η ανισότητα Hölder, το Θεώρημα Jordan-Hölder, η ταξινόμηση των απλών ομάδων της τάξης έως 200, το Θεώρημα Hölder κ.α. Μια άλλη σημαντική έννοια που σχετίζεται με το όνομά του, είναι η συνθήκη Hölder (ή συνέχειας Hölder), η οποία χρησιμεύει σε πολλούς τομείς της ανάλυσης, συμπεριλαμβανομένων των θεωριών περί μερικών διαφορικών εξισώσεων και περί συναρτησιακών χώρων.

$$\leq \frac{1}{n^{v/j}} E \left[n^{\frac{v}{j}-1} \sum_{i=1}^n |X_i - \mu|^v \right] = \frac{1}{n} E \left[\sum_{i=1}^n |X_i - \mu|^v \right] = E |X_i - \mu|^v.$$

■

Πρόταση 2.2: Αν $E|X|^k < \infty$ για σταθερό $k \in \{2, 3, \dots\}$ τότε

$$E(M_{k,n}) = \mu_k + o(1/\sqrt{n}), \text{ καθώς } n \rightarrow \infty. \quad (2.43)$$

Απόδειξη: Παρατηρούμε ότι η (2.43) ισοδυναμεί με τη σχέση

$$E[\sqrt{n}(M_{k,n} - \mu_k)] \rightarrow 0. \quad (2.44)$$

Γράφοντας

$$M_{k,n} - \mu_k = (m_{k,n} - \mu_k) + (-1)^{k-1}(k-1)m_{1,n}^k + \sum_{j=2}^{k-1} \binom{k}{j} (-1)^{k-j} m_{1,n}^{k-j} m_{j,n}, \quad (2.45)$$

αρκεί πλέον, να επαληθεύσουμε ότι, καθώς $n \rightarrow \infty$

$$(i) \quad \sqrt{n}E(m_{k,n} - \mu_k) = 0.$$

$$(ii) \quad \sqrt{n}E(m_{1,n}^k) \rightarrow 0.$$

$$(iii) \quad \sqrt{n}E(m_{1,n}^{k-j} m_{j,n}) \rightarrow 0, \quad j = 2, 3, \dots, k-2 \text{ (υπό την προϋπόθεση ότι } k \geq 4).$$

$$(iv) \quad \sqrt{n}E(m_{1,n} m_{k-1,n}) \rightarrow 0 \text{ (υπό τον όρο ότι } k \geq 3).$$

Τώρα το (i) είναι προφανές (αφού $E(m_{k,n}) = \mu_k$), το (iv) έπεται από $E(m_{1,n} m_{k-1,n}) = \mu_k/n$ και το (ii) μπορεί να αποδειχθεί χρησιμοποιώντας το Λήμμα 2.1 για $\alpha = \delta = k$, το οποίο δείχνει ότι

$$|n^{k/2} E(m_{1,n}^k)| \leq n^{k/2} E|m_{1,n}|^k = E|\sqrt{n}(\bar{X}_n - \mu)|^k \rightarrow \sigma^k E|Z|^k < \infty,$$

και έτσι,

$$|\sqrt{n}E(m_{1,n}^k)| \leq n^{-(k-1)/2} E|\sqrt{n}(\bar{X}_n - \mu)|^k \rightarrow 0.$$

Για να αποδείξουμε το (iii) υποθέτουμε ότι $k \geq 4$ και $2 \leq j \leq k-2$. Από την ανισότητα του Hölder για $p = \frac{k}{k-j} > 1$, από το Θεώρημα 2.1 για $v = k$ και από το Λήμμα 2.3 για $\alpha = \delta = k$ θα λάβουμε διαδοχικά

$$\begin{aligned}
|\sqrt{n}E(m_{1,n}^{k-j}m_{j,n})| &\leq \sqrt{n}E|m_{1,n}|^{k-j}|m_{j,n}| \leq \sqrt{n}\left(E|m_{1,n}|^k\right)^{\frac{k-j}{k}}\left(E|m_{j,n}|^{\frac{k}{j}}\right)^{\frac{j}{k}} \\
&\leq \sqrt{n}\left(n^{-\frac{k}{2}}E|\sqrt{n}(\bar{X}_n - \mu)|^k\right)^{\frac{k-j}{k}}(E|X - \mu|^k)^{\frac{j}{k}} \\
&= n^{-\frac{(k-1-j)}{2}}\left(E|\sqrt{n}(\bar{X}_n - \mu)|^k\right)^{\frac{k-j}{k}}(E|X - \mu|^k)^{\frac{j}{k}} \\
&= n^{-(k-1-j)/2} \cdot O(1) \rightarrow 0, \\
\text{αφού } E|\sqrt{n}(\bar{X}_n - \mu)|^k &\rightarrow \sigma^k E|Z|^k < \infty. \quad \blacksquare
\end{aligned}$$

Πρόταση 2.3: Αν $E|X|^{k+1} < \infty$ για σταθερό $k \in \{2, 3, \dots\}$ τότε

$$\text{Cov}(\bar{X}_n, M_{k,n}) = \frac{1}{n}(\mu_{k+1} - k\sigma^2\mu_{k-1}) + o(1/n), \text{ καθώς } n \rightarrow \infty. \quad (2.46)$$

Απόδειξη: Παρατηρούμε ότι η (2.46) είναι ισοδύναμη με τη σχέση

$$\text{Cov}[\sqrt{n}(\bar{X}_n - \mu), \sqrt{n}(M_{k,n} - \mu_k)] \rightarrow \mu_{k+1} - k\sigma^2\mu_{k-1}, \quad (2.47)$$

και εφόσον, $E(\bar{X}_n - \mu) = 0$, αρκεί να επαληθεύσουμε ότι

$$nE[(\bar{X}_n - \mu)(M_{k,n} - \mu_k)] = nE[m_{1,n}(M_{k,n} - \mu_k)] \rightarrow \mu_{k+1} - k\sigma^2\mu_{k-1}. \quad (2.48)$$

• Αν $k = 2$ τότε

$$\begin{aligned}
nE[m_{1,n}(M_{2,n} - \mu_2)] &= nE[(\bar{X}_n - \mu)(m_{2,n} - \mu_2)] - nE(\bar{X}_n - \mu)^3 \\
&= \mu_3 - nE(\bar{X}_n - \mu)^3,
\end{aligned}$$

και είναι εύκολο να δούμε, από το Θεώρημα 2.1 για $\alpha = \delta = 3$ ότι

$$|nE(\bar{X}_n - \mu)^3| \leq n^{-1/2}E|\sqrt{n}(\bar{X}_n - \mu)|^3 \rightarrow 0.$$

Έτσι, έχουμε

$$nE[m_{1,n}(M_{2,n} - \mu_2)] \rightarrow \mu_3.$$

Εφόσον, $\mu_1 = 0$, η (2.48) ικανοποιείται για $k = 2$.

• Αν $k = 3$ τότε

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

$$-3nE \left[m_{2,n}(\bar{X}_n - \mu)^2 \right].$$

Όμως, $nE[(\bar{X}_n - \mu)(m_{3,n} - \mu_3)] = \mu_4$ και από το Θεώρημα 2.1 για $\alpha = \delta = 4$ θα έχουμε $2nE(\bar{X}_n - \mu)^4 \rightarrow 0$. Εν κατακλείδι, προκύπτει ότι

$$-3nE \left[m_{2,n}(\bar{X}_n - \mu)^2 \right] = -\frac{3(\mu_4 + (n-1)\mu_2^2)}{n} \rightarrow -3\mu_2^2 = -3\sigma^4,$$

που επαληθεύει την (2.48) για $k = 3$.

Στην γενική περίπτωση όπου $k \geq 4$ γράφουμε την ποσότητα $M_{k,n} - \mu_k$ όπως στη σχέση (2.45) και παρατηρούμε ότι η (2.48) θα προκύψει αν δείξουμε ότι, καθώς $n \rightarrow \infty$

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

$$(ii) \quad nE(m_{1,n}^{k+1}) \rightarrow 0.$$

$$(iii) \quad nE(m_{1,n}^{k+1-j} m_{j,n}) \rightarrow 0, \quad j = 2, 3, \dots, k-2.$$

$$(iv) \quad nE(m_{1,n}^2 m_{k-1,n}) \rightarrow \sigma^2 \mu_{k-1}.$$

Υπολογίζουμε

$$\begin{aligned} E[m_{1,n}(m_{k,n} - \mu_k)] &= E[(\bar{X}_n - \mu)(m_{k,n} - \mu_k)] = \\ &= \frac{1}{n^2} \sum_{i_1=n}^n \sum_{i_2=n}^n E[(X_{i_1} - \mu)(X_{i_2} - \mu)^k] = \frac{\mu_{k+1}}{n}, \end{aligned}$$

και έτσι συμπεραίνουμε το (i), ενώ το (ii) απορρέει από το Θεώρημα 2.1 για $\alpha = \delta = k+1$. Επιπλέον,

$$\begin{aligned} nE(m_{1,n}^2 m_{k-1,n}) &= \frac{1}{n^2} \sum_{i_1=n}^n \sum_{i_2=n}^n \sum_{i_3=n}^n E[(X_{i_1} - \mu)(X_{i_2} - \mu)(X_{i_3} - \mu)^{k-1}] \\ &= \frac{1}{n^2} \{n\mu_{k+1} + n(n-1)\sigma^2 \mu_{k-1}\} \rightarrow \sigma^2 \mu_{k-1}, \end{aligned}$$

επομένως, η (iv) ικανοποιείται, και απομένει να επαληθευθεί η (iii). Προς αυτή την κατεύθυνση, θέτοντας $p = \frac{k+1}{k+1-j} > 1$ στην ανισότητα του Hölder και από το Λήμμα 2.3 για $v = k+1$ λαμβάνουμε διαδοχικά

$$|nE(m_{1,n}^{k+1-j} m_{j,n})| \leq nE|m_{1,n}|^{k+1-j} |m_{j,n}|$$

$$\begin{aligned}
&\leq n \left(E|m_{1,n}|^{k+1} \right)^{\frac{k+1-j}{k+1}} \left(E|m_{j,n}|^{\frac{k+1}{j}} \right)^{\frac{j}{k+1}} \\
&\leq n \left(n^{-(k+1)/2} E \left(|\sqrt{n}(\bar{X}_n - \mu)|^{k+1} \right) \right)^{\frac{k+1-j}{k+1}} (E|X - \mu|^{k+1})^{\frac{j}{k+1}} \\
&= n^{-(k-1-j)/2} \left(E|\sqrt{n}(\bar{X}_n - \mu)|^{k+1} \right)^{\frac{k+1-j}{k+1}} (E|X - \mu|^{k+1})^{\frac{j}{k+1}} \rightarrow 0,
\end{aligned}$$

επειδή $n^{-(k-1-j)/2} \rightarrow 0$ και χρησιμοποιώντας το Θεώρημα 2.1 για $\alpha = \delta = k + 1$ θα έχουμε ότι $E|\sqrt{n}(\bar{X}_n - \mu)|^{k+1} \rightarrow \sigma^{k+1} E|Z|^{k+1} < \infty$. ■

Πρόταση 2.4: Αν $E|X|^{k+r} < \infty$ για σταθερές $k, r \in \{2, 3, \dots\}$ τότε

$$Cov(M_{r,n}, M_{k,n}) = \frac{1}{n} v_{rk} + o(1/n), \text{ καθώς } n \rightarrow \infty, \quad (2.49)$$

και ιδίως, αν $E(X^{2k}) < \infty$ τότε

$$Var[\sqrt{n}(M_{k,n} - \mu_k)] \rightarrow v_k^2, \text{ καθώς } n \rightarrow \infty, \quad (2.50)$$

όπου v_{rk} και v_k^2 δίνονται από τις σχέσεις (2.12) και (2.1), αντίστοιχα.

Απόδειξη: Χωρίς βλάβη της γενικότητας υποθέτουμε ότι $2 \leq r \leq k$ και βλέπουμε ότι η (2.49) ισοδυναμεί με

$$Cov[\sqrt{n}(M_{r,n} - \mu_r), \sqrt{n}(M_{k,n} - \mu_k)] \rightarrow v_{rk}. \quad (2.51)$$

Εφόσον, $E|X|^{k+r} < \infty$, η (2.44) δείχνει ότι

$$E[\sqrt{n}(M_{k,n} - \mu_k)] \rightarrow 0 \text{ και } E[\sqrt{n}(\bar{X}_n - \mu)] \rightarrow 0,$$

και τούτο αρκεί για να επιβεβαιώσουμε ότι

$$\begin{aligned}
&nE[(M_{r,n} - \mu_r)(M_{k,n} - \mu_k)] \rightarrow v_{rk} = \\
&= \mu_{r+k} - \mu_r \mu_k - r\mu_{r-1}\mu_{k+1} - k\mu_{r+1}\mu_{k-1} + rk\sigma^2\mu_{r-1}\mu_{k-1}. \quad (2.52)
\end{aligned}$$

Η απόδειξη μπορεί να συναχθεί αποδεικνύοντας την (2.52) για κάθε μία από τις περιπτώσεις $r = k = 2$, $r = 2 \ \& \ k = 3$, $r = k = 3$, $r = 2 \ \& \ k \geq 4$, $r = 3 \ \& \ k \geq 4$, $4 \leq r \leq k$. Ακολούθως, θα παρουσιάσουμε τις λεπτομέρειες μόνο στην περίπτωση όπου $4 \leq r \leq k$, οι υπόλοιπες περιπτώσεις αντιμετωπίζονται χρησιμοποιώντας παρόμοια και απλούστερα επιχειρήματα.

Υποθέτουμε ότι $4 \leq r \leq k$. Από την (2.45) έχουμε

$$M_{r,n} - \mu_r = (m_{r,n} - \mu_r) - r m_{1,n} m_{r-1,n} + \sum_{j_1=2}^{r-2} \binom{r}{j_1} (-1)^{r-j_1} m_{1,n}^{r-j_1} m_{j_1,n} + (-1)^{r-1} (r-1) m_{1,n}^r, \quad (2.53)$$

$$M_{k,n} - \mu_k = (m_{k,n} - \mu_k) - k m_{1,n} m_{k-1,n} + \sum_{j_2=2}^{k-2} \binom{k}{j_2} (-1)^{k-j_2} m_{1,n}^{k-j_2} m_{j_2,n} + (-1)^{k-1} (k-1) m_{1,n}^k. \quad (2.54)$$

Στη συνέχεια, θα δείξουμε ότι η ασυμπτωτική διακύμανση στην (2.51) μπορεί να προσδιοριστεί χρησιμοποιώντας μόνο τους δύο πρώτους όρους από τις (2.53) και (2.54). Πράγματι, είναι εύκολο να δει κανείς ότι η (2.52) είναι αληθής εάν αποδειχθούν τα κάτωθι, καθώς $n \rightarrow \infty$:

- (i) $nE[(m_{r,n} - \mu_r)(m_{k,n} - \mu_k)] = \mu_{r+k} - \mu_r \mu_k.$
- (ii) $nE[m_{1,n} m_{k-1,n} (m_{r,n} - \mu_r)] \rightarrow \mu_{r+1} \mu_{k-1}.$
- (iii) $nE[m_{1,n} m_{r-1,n} (m_{k,n} - \mu_k)] \rightarrow \mu_{r-1} \mu_{k+1}.$
- (iv) $nE[m_{1,n}^2 m_{r-1,n} m_{k-1,n}] \rightarrow \sigma^2 \mu_{r-1} \mu_{k-1}.$
- (v) $nE[m_{1,n}^{k-j_2} m_{j_2,n} (m_{r,n} - \mu_r)] \rightarrow 0, \quad j_2 = 2, 3, \dots, k-2.$
- (vi) $nE[m_{1,n}^k (m_{r,n} - \mu_r)] \rightarrow 0.$
- (vii) $nE[m_{1,n}^{k+1-j_2} m_{j_2,n} m_{r-1,n}] \rightarrow 0, \quad j_2 = 2, 3, \dots, k-2.$
- (viii) $nE[m_{1,n}^{k+1} m_{r-1,n}] \rightarrow 0.$
- (ix) $nE[m_{1,n}^{r-j_1} m_{j_1,n} (m_{k,n} - \mu_k)] \rightarrow 0, \quad j_1 = 2, 3, \dots, r-2.$
- (x) $nE[m_{1,n}^{r+1-j_1} m_{j_1,n} m_{k-1,n}] \rightarrow 0, \quad j_1 = 2, 3, \dots, r-2.$
- (xi) $nE[m_{1,n}^{r+k-j_1-j_2} m_{j_1,n} m_{j_2,n}] \rightarrow 0, \quad j_1 = 2, 3, \dots, r-2, \quad j_2 = 2, 3, \dots, k-2.$
- (xii) $nE[m_{1,n}^{r+k-j_1} m_{j_1,n}] \rightarrow 0, \quad j_1 = 2, 3, \dots, r-2.$
- (xiii) $nE[m_{1,n}^r (m_{k,n} - \mu_k)] \rightarrow 0.$
- (xiv) $nE[m_{1,n}^{r+1} m_{k-1,n}] \rightarrow 0.$

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

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

Προχωρούμε, ευθύς, στην επιβεβαίωση των σχέσεων (i) – (xvi). Επειδή $E(m_{r,n}) = \mu_r$ και $E(m_{k,n}) = \mu_k$, θα ισχύει

$$\begin{aligned} nE[(m_{r,n} - \mu_r)(m_{k,n} - \mu_k)] &= n\{E[m_{r,n}m_{k,n}] - \mu_r\mu_k\} \\ &= n\left\{\frac{1}{n^2} \sum_{i_1=n}^n \sum_{i_2=n}^n E[(X_{i_1} - \mu)^r (X_{i_2} - \mu)^k] - \mu_r\mu_k\right\} \\ &= n\left\{\frac{1}{n^2} [n\mu_{r+k} + n(n-1)\mu_r\mu_k] - \mu_r\mu_k\right\} = \mu_{r+k} - \mu_r\mu_k, \end{aligned}$$

πράγμα που δείχνει την (i). Επίσης, οι σχέσεις (ii), (iii) και (iv) απορρέουν από απλούς υπολογισμούς. Για παράδειγμα για την (ii) έχουμε

$$\begin{aligned} nE[m_{1,n}m_{k-1,n}(m_{r,n} - \mu_r)] &= -\mu_r\mu_k + \frac{\mu_{r+k} + (n-1)(\mu_{r+1}\mu_{k-1} + \mu_r\mu_k)}{n} \rightarrow \\ &\rightarrow \mu_{r+1}\mu_{k-1}, \end{aligned}$$

ενώ η (iii) είναι παρεμφερής με την (ii), και η (iv) μπορεί να συναχθεί από την

$$\begin{aligned} nE[m_{1,n}^2 m_{r-1,n} m_{k-1,n}] &= \frac{1}{n^3} \{n(n-1)(n-2)\sigma^2 \mu_{r-1}\mu_{k-1} + o(n^3)\} \rightarrow \\ &\rightarrow \sigma^2 \mu_{r-1}\mu_{k-1}. \end{aligned}$$

Τα μηδενικά όρια των σχέσεων (vi) – (viii) και (x) – (xvi) προκύπτουν από το Θεώρημα 2.1 και το Λήμμα 2.3, θέτοντας $\alpha = \delta = v = r+k$, διότι $E|X|^{r+k} < \infty$. Οντως, έχουμε

$$|nE(m_{1,n}^{r+k})| \leq nE|m_{1,n}|^{r+k} = n^{-(r+k-2)/2} E|\sqrt{n}(\bar{X}_n - \mu)|^{r+k} \rightarrow 0,$$

γεγονός που επαληθεύει την (xvi). Τώρα για την (xv), από την ανισότητα του Hölder για $p = \frac{r+k}{r+k-j_2} > 1$, θα πάρουμε

$$\begin{aligned} |nE(m_{1,n}^{r+k-j_2} m_{j_2,n})| &\leq nE|m_{1,n}|^{r+k-j_2} |m_{j_2,n}| \\ &\leq n(E|m_{1,n}|^{r+k})^{\frac{r+k-j_2}{r+k}} \left(E|m_{j_2,n}|^{\frac{r+k}{j_2}}\right)^{\frac{j_2}{r+k}} \end{aligned}$$

$$\begin{aligned} &\leq n \left(n^{-r+k/2} E |\sqrt{n}(\bar{X}_n - \mu)|^{r+k} \right)^{\frac{r+k-j_2}{r+k}} (E|X - \mu|^{r+k})^{\frac{j_2}{r+k}} \\ &= n^{-(r+k-j_2-2)/2} \left(E |\sqrt{n}(\bar{X}_n - \mu)|^{r+k} \right)^{\frac{r+k-j_2}{r+k}} (E|X - \mu|^{r+k})^{\frac{j_2}{r+k}} \rightarrow 0, \end{aligned}$$

αφού $n^{-(r+k-j_2-2)/2} \rightarrow 0$ και $E|X - \mu|^{r+k} \rightarrow \sigma^{r+k} E|Z|^{r+k} < \infty$.

Με ανάλογο τρόπο, προκύπτει και η σχέση (xii). Για το όριο (xvi) έχουμε

$$\begin{aligned} |nE[m_{1,n}^{r+1} m_{k-1,n}]| &\leq nE|m_{1,n}|^{r+1} |m_{k-1,n}| \\ &\leq n \left(E|m_{1,n}|^{r+k} \right)^{\frac{r+1}{r+k}} \left(E|m_{k-1,n}|^{\frac{r+k}{k-1}} \right)^{\frac{k-1}{r+k}} \\ &\leq n^{-(r-1)/2} \left(E |\sqrt{n}(\bar{X}_n - \mu)|^{r+k} \right)^{\frac{r+1}{r+k}} (E|X - \mu|^{r+k})^{\frac{k-1}{r+k}} \rightarrow 0, \end{aligned}$$

και παρομοίως για την (viii). Προκειμένου να αποδείξουμε την (xiii), αρκεί να δειχτεί ότι $nE[m_{1,n}^r m_{k,n}] \rightarrow 0$ και $nE[m_{1,n}^r] \rightarrow 0$. Το δεύτερο όριο είναι προφανές, καθώς από την (xvi), μπορούμε εύκολα να πάρουμε ότι

$$|nE[m_{1,n}^r]| \leq n^{-(r-2)/2} E |\sqrt{n}(\bar{X}_n - \mu)|^r = n^{-(r-2)/2} \cdot O(1) \rightarrow 0.$$

Τώρα, για το πρώτο όριο λαμβάνουμε

$$\begin{aligned} |nE[m_{1,n}^r m_{k,n}]| &\leq nE|m_{1,n}|^r |m_{k,n}| \leq n \left(E|m_{1,n}|^{r+k} \right)^{\frac{r}{r+k}} \left(E|m_{k,n}|^{\frac{r+k}{k}} \right)^{\frac{k}{r+k}} \\ &\leq n^{-(r-2)/2} \left(E |\sqrt{n}(\bar{X}_n - \mu)|^{r+k} \right)^{\frac{r}{r+k}} (E|X - \mu|^{r+k})^{\frac{k}{r+k}} \rightarrow 0. \end{aligned}$$

Το όριο (vi) είναι παρεμφερές με το (xiii) και έτσι η απόδειξή του παραλείπεται. Παρατηρώντας τη σχέση (xi) έχουμε

$$\begin{aligned} |nE[m_{1,n}^{r+k-j_1-j_2} m_{j_1,n} m_{j_2,n}]| &\leq nE|m_{1,n}|^{r+k-j_1-j_2} |m_{j_1,n} m_{j_2,n}| \\ &\leq n \left(E|m_{1,n}|^{r+k} \right)^{\frac{r+k-j_1-j_2}{r+k}} \left(E|m_{j_1,n} m_{j_2,n}|^{\frac{r+k}{j_1+j_2}} \right)^{\frac{j_1+j_2}{r+k}} \\ &\leq n \left(E|m_{1,n}|^{r+k} \right)^{\frac{r+k-j_1-j_2}{r+k}} \left\{ \left(E|m_{j_1,n}|^{\frac{r+k}{j_1}} \right)^{\frac{j_1}{j_1+j_2}} \left(E|m_{j_2,n}|^{\frac{r+k}{j_2}} \right)^{\frac{j_2}{j_1+j_2}} \right\}^{\frac{j_1+j_2}{r+k}} \end{aligned}$$

$$\begin{aligned}
&\leq n \left(E|m_{1,n}|^{r+k} \right)^{\frac{r+k-j_1-j_2}{r+k}} (E|X - \mu|^{r+k})^{\frac{j_1+j_2}{r+k}} \\
&= n^{-(r+k-j_1-j_2-2)/2} \left(E|\sqrt{n}(\bar{X}_n - \mu)|^{r+k} \right)^{\frac{r+k-j_1-j_2}{r+k}} (E|X - \mu|^{r+k})^{\frac{j_1+j_2}{r+k}} \rightarrow 0.
\end{aligned}$$

Παρόμοια, για το όριο (x) ισχύει

$$\begin{aligned}
&|nE[m_{1,n}^{r+1-j_1} m_{j_1,n} m_{k-1,n}]| \leq nE|m_{1,n}|^{r+1-j_1} |m_{j_1,n} m_{k-1,n}| \\
&\leq n \left(E|m_{1,n}|^{r+k} \right)^{\frac{r+1-j_1}{r+k}} \left(E|m_{j_1,n} m_{k-1,n}|^{\frac{r+k}{j_1+k-1}} \right)^{\frac{j_1+k-1}{r+k}} \\
&\leq n \left(E|m_{1,n}|^{r+k} \right)^{\frac{r+1-j_1}{r+k}} \left\{ \left(E|m_{j_1,n}|^{\frac{r+k}{j_1}} \right)^{\frac{j_1}{j_1+k-1}} \left(E|m_{k-1,n}|^{\frac{r+k}{k-1}} \right)^{\frac{k-1}{j_1+k-1}} \right\}^{\frac{j_1+k-1}{r+k}} \\
&\leq n \left(E|m_{1,n}|^{r+k} \right)^{\frac{r+1-j_1}{r+k}} (E|X - \mu|^{r+k})^{\frac{j_1+k-1}{r+k}} \\
&= n^{-(r-j_1-1)/2} \left(E|\sqrt{n}(\bar{X}_n - \mu)|^{r+k} \right)^{\frac{r+k-j_1}{r+k}} (E|X - \mu|^{r+k})^{\frac{j_1+k-1}{r+k}} \rightarrow 0,
\end{aligned}$$

ενώ το (vii) προκύπτει ανάλογα με το (x).

Υπολείπονται προς επιβεβαίωση οι σχέσεις (v) και (ix), αλλά εφόσον τούτες είναι παρεμφερείς, αρκεί να αποδείξουμε τη μία από τις δύο, έστω την (v). Ας υποθέσουμε, λοιπόν, ότι $j_2 = \{2, 3, \dots, k-3\}$, έτσι παίρνουμε $k \geq 5$ και $j_2 < k-2$. Επομένως, θα λάβουμε

$$\begin{aligned}
&|nE[m_{1,n}^{k-j_2} m_{j_2,n} (m_{r,n} - \mu_r)]| \\
&\leq nE|m_{1,n}|^{k-j_2} |m_{j_2,n} m_{r,n}| + n|\mu_r| E|m_{1,n}|^{k-j_2} |m_{j_2,n}|,
\end{aligned}$$

και αρκεί να δείξουμε ότι

$$nE|m_{1,n}|^{k-j_2} |m_{j_2,n} m_{r,n}| \rightarrow 0 \text{ και } nE|m_{1,n}|^{k-j_2} |m_{j_2,n}| \rightarrow 0.$$

Για το πρώτο όριο πράγματι θα έχουμε

$$nE|m_{1,n}|^{k-j_2} |m_{j_2,n} m_{r,n}| \leq n \left(E|m_{1,n}|^{r+k} \right)^{\frac{k-j_2}{r+k}} \left(E|m_{j_2,n} m_{r,n}|^{\frac{r+k}{r+j_2}} \right)^{\frac{r+j_2}{r+k}}$$

$$\leq n \left(E|m_{1,n}|^{r+k} \right)^{\frac{k-j_2}{r+k}} \left\{ \left(E|m_{j_2,n}|^{\frac{r+k}{j_2}} \right)^{\frac{j_2}{r+j_2}} \left(E|m_{r,n}|^{\frac{r+k}{r+j_2}} \right)^{\frac{r}{j_1+j_2}} \right\}^{\frac{r+j_2}{r+k}}$$

$$\leq n^{-(k-j_2-2)/2} \left(E|\sqrt{n}(\bar{X}_n - \mu)|^{r+k} \right)^{\frac{k-j_2}{r+k}} (E|X - \mu|^{r+k})^{\frac{r+j_2}{r+k}} \rightarrow 0,$$

αφού $k - j_2 - 2 > 0$. Ομοίως, για το δεύτερο όριο ισχύει

$$nE|m_{1,n}|^{k-j_2}|m_{j_2,n}| \leq n \left(E|m_{1,n}|^{r+k} \right)^{\frac{k-j_2}{r+k}} \left(E|m_{j_2,n}|^{\frac{r+k}{j_2}} \right)^{\frac{r+j_2}{r+k}}$$

$$\leq n \left(E|m_{1,n}|^{r+k} \right)^{\frac{k-j_2}{r+k}} \left(E|m_{j_2,n}|^{\frac{r+k}{j_2}} \right)^{\frac{j_2}{r+j_2}}$$

$$\leq n^{-(k-j_2-2)/2} \left(E|\sqrt{n}(\bar{X}_n - \mu)|^{r+k} \right)^{\frac{k-j_2}{r+k}} (E|X - \mu|^{r+k})^{\frac{j_2}{r+k}} \rightarrow 0,$$

διότι $k - j_2 - 2 > 0$. Τέλος, απομένει να μελετήσουμε το όριο (v) όταν $j_2 = k - 2$. Σε αυτήν την περίπτωση τα όρια δεν μηδενίζονται απαραίτητα. Ως εκ τούτου, εφόσον $j_2 = k - 2$ θα λάβουμε

$$nE[m_{1,n}^{k-j_2} m_{j_2,n} (m_{r,n} - \mu_r)] = nE[m_{1,n}^2 m_{r,n} m_{k-2,n}] - n\mu_r E[m_{1,n}^2 m_{k-2,n}],$$

και με άμεσους υπολογισμούς καταλήγουμε ότι

$$nE[m_{1,n}^2 m_{r,n} m_{k-2,n}] = \frac{1}{n^3} \{n(n-1)(n-2)\sigma^2 \mu_r \mu_{k-2} + o(n^3)\} \rightarrow \sigma^2 \mu_r \mu_{k-2}$$

και

$$nE[m_{1,n}^2 m_{k-2,n}] = \frac{1}{n^2} \{n(n-1)\sigma^2 \mu_{k-2} + o(n^2)\} \rightarrow \sigma^2 \mu_{k-2}.$$

Έτσι, όταν $j_2 = k - 2$ παίρνουμε

$$\begin{aligned} nE[m_{1,n}^{k-j_2} m_{j_2,n} (m_{r,n} - \mu_r)] &= nE[m_{1,n}^2 m_{r,n} m_{k-2,n}] - n\mu_r E[m_{1,n}^2 m_{k-2,n}] \\ &\rightarrow \sigma^2 \mu_r \mu_{k-2} - \mu_r (\sigma^2 \mu_{k-2}) = 0, \end{aligned}$$

και η απόδειξη είναι πλήρης. ■

Συνεπώς, είναι λογικό να εισάγουμε τον ακόλουθο ορισμό.

Ορισμός 2.1: Υποθέτουμε ότι $k \in \{2, 3, \dots\}$. Τότε, η δειγματική μέση τιμή $\bar{X}_n$ είναι ασυμπτωτικά ανεξάρτητη της δειγματικής κεντρικής ροπής $M_{k,n}$, αν υπάρχουν ανεξάρτητες τυχαίες μεταβλητές W_1 και W_k τέτοιες ώστε, καθώς $n \rightarrow \infty$

$$\sqrt{n} \begin{pmatrix} \bar{X}_n - \mu \\ M_{k,n} - \mu_k \end{pmatrix} \xrightarrow{d} \begin{pmatrix} W_1 \\ W_k \end{pmatrix}.$$

Επιπλέον, η $\bar{X}_n$ είναι ασυμπτωτικά ανεξάρτητη του τυχαίου διανύσματος $(M_{2,n}, \dots, M_{k,n})'$ όταν υπάρχουν τυχαίες μεταβλητές $W_1, W_2, \dots, W_k$ ούτως ώστε W_1 και $(W_2, \dots, W_k)'$ να είναι ανεξάρτητα και καθώς $n \rightarrow \infty$, να ισχύει

$$\sqrt{n} \begin{pmatrix} \bar{X}_n - \mu \\ M_{2,n} - \mu_2 \\ M_{3,n} - \mu_3 \\ \vdots \\ M_{k,n} - \mu_k \end{pmatrix} \xrightarrow{d} \begin{pmatrix} W_1 \\ W_2 \\ W_3 \\ \vdots \\ W_k \end{pmatrix}.$$

Τέλος, οι ποσότητες $\bar{X}_n$ και $M_{k,n}$ καλούνται ασυμπτωτικά ασυσχέτιστες όταν

$$\text{Cov}(\sqrt{n} \bar{X}_n, \sqrt{n} M_{k,n}) \rightarrow 0, \text{ καθώς } n \rightarrow \infty.$$

Παρατήρηση: Θεωρούμε ότι $E(X^{2k}) < \infty$ για κάποιο σταθερό $k \in \{2, 3, \dots\}$. Σύμφωνα με την σχέση (2.14) και την Πρόταση 2.3 (βλ. επίσης, και την (2.47)), η $\bar{X}_n$ και η $M_{k,n}$ είναι ασυμπτωτικά ανεξάρτητες αν και μόνο αν είναι ασυμπτωτικά ασυσχέτιστες. Επιπλέον, η ασυμπτωτική κανονικότητα για την οποία κάνει λόγο η Πρόταση 2.1 δείχνει ότι, $\bar{X}_n$ και $(M_{2,n}, \dots, M_{k,n})'$ είναι ασυμπτωτικά ανεξάρτητα αν και μόνο αν $\bar{X}_n$ και $M_{r,n}$ είναι ασυμπτωτικά ασυσχέιστα για όλα τα $r \in \{2, 3, \dots, k\}$. Αν υποθέσουμε, απλώς, ότι $E|X|^{k+1} < \infty$ τότε, ακόμη και για εκείνες τις περιπτώσεις όπου η οριακή κατανομή της ποσότητας $\sqrt{n}(M_{k,n} - \mu_k)$ δεν υπάρχει, η Πρόταση 2.3 επιτρέπει να αποφασίσουμε αν τα $\bar{X}_n, M_{k,n}$ είναι ασυμπτωτικά ασυσχέιστα ή όχι.

Υποθέτουμε τώρα ότι $X \sim \mathcal{N}(\mu, \sigma^2)$. Κοιτάζοντας τον πίνακα διασποράς στην (2.15), καθίσταται σαφές ότι η πρώτη στήλη, εξαιρουμένου του πρώτου στοιχείου σ^2 , μηδενίζεται. Αυτό γίνεται επειδή $\mu_k = 0$ για όλα τα περιττά k και $\mu_{2r} = \frac{(2r)!}{2^r r!} \sigma^{2r}$, έτσι για οποιεσδήποτε σταθερές $k \in \{2, 3, \dots\}$ ισχύει $\mu_{k+1} - k\sigma^2 \mu_{k-1}$. Σύμφωνα με τον Ορισμό 2.1, έχουμε ότι ο δειγματικός μέσος $\bar{X}_n$ είναι ασυμπτωτικά ανεξάρτητος (ασυσχέτιστος) από όλες τις δειγματικές

κεντρικές ροπές $M_{k,n}$, καθώς $n \rightarrow \infty$. Τούτο όμως, δεν είναι παράξενο γεγονός για την κανονική κατανομή, αφού είναι κοινώς γνωστό ότι για οποιαδήποτε σταθερά $n \geq 2$, η $\bar{X}$ είναι ανεξάρτητη από το διάνυσμα $Y := (X_1 - \bar{X}, X_2 - \bar{X}, \dots, X_n - \bar{X})'$. Όντως, αρκεί να παρατηρήσουμε ότι το διάνυσμα $(X_1 - \bar{X}, X_2 - \bar{X}, \dots, X_n - \bar{X})'$ ακολουθεί την πολυπαραγοντική κανονική κατανομή και $Cov(\bar{X}, X_i - \bar{X}) = 0$ για όλα τα i και ως εκ τούτου, η $\bar{X}$ είναι στοχαστικά ανεξάρτητη (και ασυσχέτιστη) από οποιαδήποτε ακολουθία της μορφής $\{h_r(Y), r = 2, 3, \dots\}$, όπου $h_r: \mathcal{R}^n \rightarrow \mathcal{R}$ είναι αυθαίρετες Borel–μετρήσιμες συναρτήσεις. Εφόσον, $M_r = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^r = h_r(Y)$ προκύπτει ότι $\bar{X}$ και $(M_2, M_3, \dots, M_k)'$ είναι ανεξάρτητα άρα, $\bar{X}$ και M_k είναι ασυσχέτιστα για όλα τα k και n . Μάλιστα, το ίδιο είναι αληθές και για τα όρια (καθώς $n \rightarrow \infty$) των κατανομών τους. Ιδιαίτερο ενδιαφέρον παρουσιάζει και το αντίστροφο, ήτοι, η ασυμπτωτική ανεξαρτησία των $\bar{X}, M_k$ για όλα τα k χαρακτηρίζει την κανονικότητα.

Θεώρημα 2.3: Θεωρούμε μία μη–εκφυλισμένη τυχαία μεταβλητή X που έχει πεπερασμένες ροπές οποιασδήποτε τάξης. Αν $\bar{X}$ και M_k είναι ασυμπτωτικά ανεξάρτητες (ή απλώς, ασυμπτωτικά ασυσχέτιστες) για όλα τα $k \in \{2, 3, \dots\}$ τότε η X ακολουθεί την κανονική κατανομή.

Απόδειξη: Έστω $\mu = E(X)$, $\sigma^2 = Var(X) > 0$ και $\mu_k = E(X - \mu)^k$. Λόγω της Πρότασης 2.3 αλλά και των σχέσεων (2.14), (2.15) έπεται ότι η $\bar{X}$ και η M_k είναι ασυμπτωτικά ασυσχέτιστες αν και μόνο αν $\mu_{k+1} = k\sigma^2\mu_{k-1}$. Όμως, υποθέσαμε ότι αυτή η σχέση υφίσταται για όλα τα $k \geq 2$, έτσι, προκύπτει ότι $\mu_1 = \mu_3 = \mu_5 = \dots = 0$ και παρόμοια για όλα τα $r \in \{1, 2, \dots\}$ έχουμε $\mu_{2r} = \frac{(2r)!}{2^r r!} \sigma^{2r}$. Ωστόσο, οι τελευταίες σχέσεις αναπαριστούν τις ροπές της $\mathcal{N}(0, \sigma^2)$ και αφού οι κανονικές κατανομές χαρακτηρίζονται από αυτήν την ακολουθία ροπών, καταλήγουμε ότι $X - \mu \sim \mathcal{N}(0, \sigma^2)$. ■

Αντιπαραβάλλοντας, τον συνήθη σπουδαίο χαρακτηρισμό της κανονικότητας, δηλαδή, την ασυμπτωτική ανεξαρτησία, έναντι της ανεξαρτησίας των $\bar{X}, S^2 = \frac{n}{n-1} M_2$, αποφαινόμαστε ότι η ασυμπτωτική ανεξαρτησία είναι μία πολύ ασθενέστερη προϋπόθεση για να επιτρέψει ένα αποτέλεσμα χαρακτηρισμού. Για παράδειγμα, η σχέση (2.14) και η Πρόταση 2.3 για $k = 2$ (βλ. και την (2.47)) δείχνουν ότι $\bar{X}$ και S^2 είναι ασυμπτωτικά ανεξάρτητα αν και μόνο αν $\mu_3 = 0$, υπό τον όρο ότι $E|X|^3 < \infty$. Αναφανδόν, η

σχέση $E(X - \mu)^3 = 0$ πληρούται από οποιαδήποτε συμμετρική κατανομή (με πεπερασμένη ροπή τρίτης τάξης), καθώς επίσης, και από πολλές άλλες. Αφετέρου, η αξίωση ότι η ασυμπτωτική ανεξαρτησία πρέπει να ικανοποιείται για όλα τα $k \geq 2$, ενδεχομένως να θεωρηθεί υπέρ του δέοντος περιοριστική. Τα ακόλουθα αποτελέσματα αποδεικνύουν ότι η απαίτηση αυτή δεν μπορεί να παραληφθεί.

Θεώρημα 2.4: Για οποιοδήποτε $k \geq 2$ υπάρχουν (απεριόριστα πολλές) μη-εκφυλισμένες, μη-κανονικές τυχαίες μεταβλητές X με πεπερασμένες ροπές οποιασδήποτε τάξης τέτοιες ώστε $\bar{X}$ και $(M_2, M_3, \dots, M_k)'$ να είναι ασυμπτωτικά ανεξάρτητα.

Απόδειξη: Έστω $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ η συνάρτηση πυκνότητας πιθανότητας της τυπικής κανονικής κατανομής $\mathcal{N}(0, 1)$ και ας θεωρήσουμε το πολυώνυμο $P_m(x) := \frac{d^m}{dx^m}(x^m(1 - x^m))$ του Legendre³⁴ βαθμού m . Είναι γνωστό ότι για όλα τα $m \geq k + 2$, το P_m είναι ορθογώνιο ως προς τη βάση $\{1, x, x^2, \dots, x^{k+1}\}$ στο διάστημα $[0, 1]$, ήτοι,

$$\int_0^1 x^j P_m(x) dx = 0, \quad j = 0, 1, \dots, k + 1.$$

Εφόσον, η συνάρτηση P_m είναι συνεχής στο $[0, 1]$, έπεται ότι $0 < \max_{x \in [0, 1]} |P_m(x)| := a_m < \infty$. Επιπλέον, η συνάρτηση πυκνότητας πιθανότητας φ είναι φθίνουσα στο $[0, +\infty)$ οπότε ισχύει $\min_{x \in [0, 1]} \varphi(x) = \varphi(1) = \frac{1}{\sqrt{2\pi e}} > 0$. Βεβαίως, μπορούμε να επιλέξουμε έναν αρκετά μικρό θετικό αριθμό ε_m , όπως $\varepsilon_m = \frac{1}{2a_m\sqrt{2\pi e}} > 0$, που εγγυάται ότι $\varphi(x) + \varepsilon_m P_m(x) > 0, \forall x \in [0, 1]$. Τώρα, ορίζουμε μία ακολουθία πυκνοτήτων $\{f_m, m \geq k + 2\}$ ως εξής

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

Αν $X_m \sim f_m$ τότε για $j = 0, 1, \dots, k + 1$ έχουμε

$$E(X_m^j) = \int_{-\infty}^{+\infty} x_m^j f_m(x) dx$$

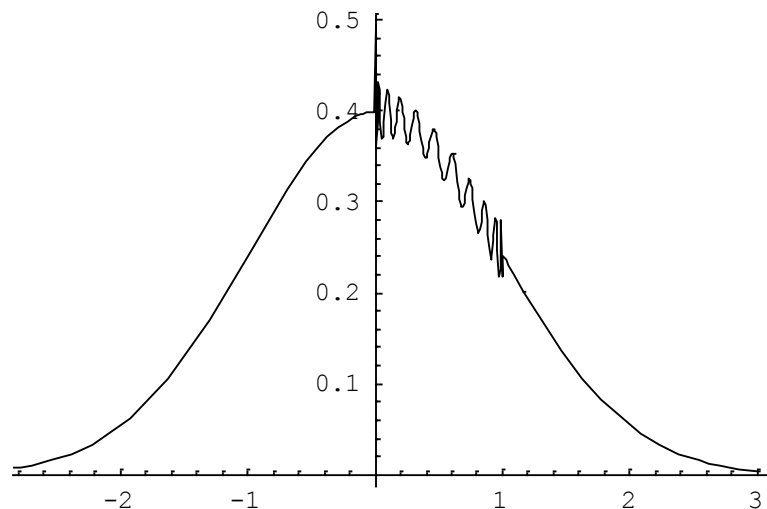
³⁴Γάλλος μαθηματικός με πλήρες όνομα *Adrien-Marie Legendre*, 18 Σεπτεμβρίου 1752 – 10 Ιανουαρίου 1833. Γνωστές και σημαντικές έννοιες που φέρουν το όνομά του είναι τα πολυώνυμα *Legendre* και ο μετασχηματισμός *Legendre*.

$$\begin{aligned}
&= \int_{-\infty}^0 x_m^j \varphi(x) dx + \int_0^1 x_m^j (\varphi(x) + \varepsilon_m P_m(x)) dx + \int_1^{+\infty} x_m^j \varphi(x) dx \\
&= \int_{-\infty}^0 x_m^j \varphi(x) dx + \int_0^1 x_m^j \varphi(x) dx + \int_1^{+\infty} x_m^j \varphi(x) dx + \varepsilon_m \int_0^1 x_m^j P_m(x) dx \\
&= \int_{-\infty}^{+\infty} x_m^j \varphi(x) dx = E(Z^j),
\end{aligned}$$

όπου $Z \sim \mathcal{N}(0,1)$. Επίσης, είναι σαφές ότι κάθε X_m είναι μη-κανονική, μη-εκφυλισμένη τυχαία μεταβλητή με πεπερασμένες ροπές οποιασδήποτε τάξης και από τον Ορισμό 2.1 απορρέει το συμπέρασμα ότι $\bar{X}$ και $(M_2, M_3, \dots, M_k)'$ είναι ασυμπτωτικά ανεξάρτητα.

Η καμπύλη της άνωθεν ακολουθίας πυκνοτήτων $\{f_m, m \geq k+2\}$ συνάδει με την κωδωνοειδή μορφή της τυπικής κανονικής κατανομής $\mathcal{N}(0,1)$ σε ολόκληρη την πραγματική ευθεία εκτός του διαστήματος $[0,1]$ για το οποίο, στην τυποποιημένη κανονική πυκνότητα $\varphi(x)$ προσθέτουμε κατάλληλο θετικό (αρκετά μικρό) πολλαπλάσιο ε_m του πολυωνύμου *Legendre*, $P_m(x)$, 21^{ου} βαθμού με συνέπεια, στο διάστημα αυτό, να μην υφίσταται η κανονικότητα.

Σχηματικά, η γραφική παράσταση θα έχει τη μορφή



"Μια μη-κανονική πυκνότητα για την οποία ο δειγματικός μέσος είναι ασυμπτωτικά ανεξάρτητος από τις πρώτες δεκαεννέα δειγματικές κεντρικές ροπές."

όπου αυτή η πυκνότητα έχει τις είκοσι πρώτες ροπές ίσες με αυτές της τυποποιημένης κανονικής. ■

2.4 Οι ιδιάζουσες κατανομές

Για οποιαδήποτε $r, k \in \{2, 3, \dots\}$ ισχύουν οι κάτωθι σχέσεις

$$\begin{aligned}
 v_k^2 &= \mu_{2k} - \mu_k^2 + k\mu_{k-1}(k\mu_{k-1}\sigma^2 - 2\mu_{k+1}) \\
 &= E(X - \mu)^{2k} - [E(X - \mu)^k]^2 + k^2\mu_{k-1}^2\sigma^2 - 2k\mu_{k-1}\mu_{k+1} \\
 &= \text{Var}(X - \mu)^k + k^2\mu_{k-1}^2\sigma^2 - 2k\mu_{k-1}[E(X - \mu)^{k+1} - E(X - \mu)^k E(X - \mu)] \\
 &= \text{Var}(X - \mu)^k + k^2\mu_{k-1}^2\text{Var}(X) - 2k\mu_{k-1}\text{Cov}[(X - \mu)^k, (X - \mu)] \\
 &= \text{Var}(X - \mu)^k + k^2\mu_{k-1}^2\text{Var}(X - \mu) - 2\text{Cov}[(X - \mu)^k, k\mu_{k-1}(X - \mu)] \\
 &= \text{Var}[(X - \mu)^k - k\mu_{k-1}(X - \mu)], \tag{2.55}
 \end{aligned}$$

$$\begin{aligned}
 v_{rk} &= \mu_{r+k} - k\mu_{k-1}\mu_{r+1} - r\mu_{r-1}\mu_{k+1} + kr\mu_{r-1}\mu_{k-1}\sigma^2 - \mu_r\mu_k \\
 &= E(X - \mu)^{r+k} - k\mu_{k-1}E(X - \mu)^{r+1} - r\mu_{r-1}E(X - \mu)^{k+1} + kr\mu_{r-1}\mu_{k-1}\mu_2 \\
 &\quad - E(X - \mu)^r E(X - \mu)^k \\
 &= E[(X - \mu)^{r+k} - k\mu_{k-1}(X - \mu)^{r+1} - r\mu_{r-1}(X - \mu)^{k+1} + kr\mu_{r-1}\mu_{k-1}(X - \mu)^2] \\
 &\quad - E[(X - \mu)^r - r\mu_{r-1}(X - \mu)]E[(X - \mu)^k - k\mu_{k-1}(X - \mu)] \\
 &= \text{Cov}[(X - \mu)^r - r\mu_{r-1}(X - \mu), (X - \mu)^k - k\mu_{k-1}(X - \mu)], \tag{2.56}
 \end{aligned}$$

και

$$\begin{aligned}
 \mu_{k+1} - k\mu_{k-1}\sigma^2 &= \mu_{k+1} - k\mu_{k-1}\mu_2 = E(X - \mu)^{k+1} - k\mu_{k-1}E(X - \mu)^2 \\
 &= E[(X - \mu)^{k+1} - k\mu_{k-1}(X - \mu)^2] - E(X - \mu)E[(X - \mu)^k - k\mu_{k-1}(X - \mu)] \\
 &= \text{Cov}[(X - \mu), (X - \mu)^k - k\mu_{k-1}(X - \mu)], \tag{2.57}
 \end{aligned}$$

υπό την προϋπόθεση ότι $E(X^{2k}) < \infty$, $E|X|^{r+k} < \infty$ και $E|X|^{r+k} < \infty$, αντίστοιχα. Ωστόσο, θέτοντας $Y = X - \mu$ με $E(Y^k) = \mu_k$, ο πίνακας διασποράς της οριακής κατανομής όπως αυτός δίνεται από τη σχέση (2.15) συμπίπτει με τον πίνακα διακύμανσης του παρακάτω τυχαίου διανύσματος

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

Παρατήρηση: Παίρνοντας διασπορές στα στοιχεία αυτού του διανύσματος θα λάβουμε τα στοιχεία της κύριας διαγωνίου του πίνακα της σχέσης (2.15). Πράγματι, δεδομένου ότι $Var(X \pm Y) = Var(X) + Var(Y) \pm 2Cov(X, Y)$ και $Cov(X, Y) = E(XY) - E(X)E(Y)$ έχουμε

$$Var(Y) = Var(X - \mu) = Var(X) = \sigma^2,$$

$$Var(Y^2) = E(Y^4) - (E(Y^2))^2 = E(X - \mu)^4 - (E(X - \mu)^2)^2 = \mu_4 - \sigma^4 = v_2^2,$$

$$Var(Y^3 - 3\sigma^2 Y) = Var(Y^3) + 9\sigma^4 Var(Y) - 2Cov(Y^3, 3\sigma^2 Y)$$

$$= E(X - \mu)^6 - (E(X - \mu)^3)^2 + 9\sigma^4 \sigma^2 - 6\sigma^2 Cov(Y^3, Y)$$

$$= \mu_6 - \mu_3^2 + 9\sigma^6 - 6\sigma^2(E(Y^4) - E(Y^3)E(Y)) = \mu_6 - \mu_3^2 + 9\sigma^6 - 6\sigma^2 E(Y^4)$$

$$= \mu_6 - \mu_3^2 + 9\sigma^6 - 6\sigma^2 \mu_4 = v_3^2,$$

⋮

⋮

⋮

$$Var(Y^k - k\mu_{k-1}Y) = Var((X - \mu)^k - k\mu_{k-1}(X - \mu))$$

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

$$= \mu_{2k} - \mu_k^2 + k^2 \mu_{k-1}^2 \sigma^2 - 2k\mu_{k-1}(E(X - \mu)^{k+1} - E(X - \mu)^k E(X - \mu))$$

$$= \mu_{2k} - \mu_k^2 + k^2 \mu_{k-1}^2 \sigma^2 - 2k\mu_{k-1}E(X - \mu)^{k+1}$$

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

Η σχέση (2.55) υποδεικνύει ότι $v_k^2 \geq 0$. Φυσικά, η μη-αρνητικότητα της v_k^2 είναι άμεση συνέπεια του γεγονότος ότι, από την Πρόταση 2.4, ισχύει $v_k^2 = \lim_{n \rightarrow \infty} Var[\sqrt{n}(M_{k,n} - \mu_k)]$ (βλ. την (2.45)) αλλά, το θέμα εδώ είναι ότι δεν έχουμε όριο για να αναφερθούμε σε αυτό. Επιπροσθέτως, η έκφραση (2.55) παρέχει την δυνατότητα να σκιαγραφήσουμε όλες τις κατανομές στην περίπτωση που $v_k^2 = 0$. Κατανομές που είναι συνυφασμένες με την τελευταία περίπτωση θα ονομάζονται ιδιάζουσες, σύμφωνα με τον ορισμό που ακολουθεί.

Ορισμός 2.2: Για σταθερό $k \geq 2$, η μη-εκφυλισμένη τυχαία μεταβλητή X , ή η συνάρτηση κατανομής της, ας είναι F , ονομάζεται ιδιάζουσα (τάξης k) αν $E(X^{2k}) < \infty$ και

$$\sqrt{n}(M_{k,n} - \mu_k) \xrightarrow{p} 0.$$

Το σύνολο όλων των ιδιαζόντων τυχαίων μεταβλητών τάξης k θα αναπαριστάται με $\mathcal{F}_k$ και το υποσύνολο όλων των τυποποιημένων (ήτοι, με μέσο 0 και διακύμανση 1) ιδιαζόντων τυχαίων μεταβλητών τάξης k θα συμβολίζεται με $\mathcal{F}_k^0$.

Σημειώνοντας ότι $Y \in \mathcal{F}_k^0$ αν και μόνο αν $X := \mu + \sigma Y \in \mathcal{F}_k$, όπου $\mu \in \mathcal{R}$ παράμετρος θέσης και $\sigma > 0$ παράμετρος κλίμακας, έπεται ότι το σύνολο $\mathcal{F}_k$ περιέχει ακριβώς την οικογένεια θέσης–κλίμακας των τυχαίων μεταβλητών που ανήκουν στο $\mathcal{F}_k^0$. Σύμφωνα με τις σχέσεις (2.7), (2.55) και τις Προτάσεις 2.2, 2.4 (βλ. επίσης, και τις (2.44), (2.50)), $X \in \mathcal{F}_k$ όταν και μόνο όταν $v_k^2 = 0$ ή ισοδύναμα,

$$\text{Var}[(X - \mu)^k - k\mu_{k-1}(X - \mu)] = 0 \Rightarrow$$

$$(X - \mu)^k - k\mu_{k-1}(X - \mu) = E[(X - \mu)^k - k\mu_{k-1}(X - \mu)] = E(X - \mu)^k = \mu_k$$

$$\Rightarrow (X - \mu)^k = \mu_k + k\mu_{k-1}(X - \mu) \text{ με πιθανότητα ένα.} \quad (2.58)$$

Επιπλέον, σημειώνουμε ότι το $\mathcal{F}_k$ είναι μη–κενό για όλα τα $k \geq 2$. Όντως, είναι εύκολο να δούμε ότι η τυχαία μεταβλητή Y με συνάρτηση μάζας πιθανότητας $P(Y = \pm 1) = 1/2$ ανήκει στο $\mathcal{F}_{2k}^0 \subseteq \mathcal{F}_{2k}$, $k = 1, 2, \dots$, διότι $\mu = E(Y) = 0$, $\sigma^2 = E(Y^2) = 1$, $\mu_{2k} = E(Y^{2k}) = 1$, $\mu_{2k-1} = E(Y^{2k-1}) = 0$ και $Y^{2k} = \mu_{2k} + 2k\mu_{2k-1}Y$ με πιθανότητα ένα. Παρόμοια, για κάθε $k \in \{1, 2, \dots\}$, η τρίτιμη συμμετρική τυχαία μεταβλητή Y_{2k+1} με συνάρτηση μάζας πιθανότητας $P(Y_{2k+1} = \pm\sqrt{2k+1}) = \frac{1}{2(2k+1)}$ και $P(Y_{2k+1} = 0) = \frac{2k}{2k+1}$ ανήκει στο $\mathcal{F}_{2k+1}^0$. Εξάλλου, κάτω από το Λήμμα 3.2, που εμφανίζεται στις επόμενες σελίδες, θα αποδείξουμε ότι μπορούμε να βρούμε μία μοναδική τιμή $p = p_{2k+1} \in (\frac{1}{2}, 1)$, για την οποία η δίτιμη τυχαία μεταβλητή W_{2k+1} με συνάρτηση μάζας πιθανότητας

$$P\left(W_{2k+1} = \sqrt{\frac{1-p}{p}}\right) = p = 1 - P\left(W_{2k+1} = -\sqrt{\frac{p}{1-p}}\right)$$

είναι τέτοια ώστε $W_{2k+1} \in \mathcal{F}_{2k+1}^0$.

Πράγματι, σε γενικές γραμμές, παρατηρούμε ότι η εξίσωση της μορφής $y^k = \alpha + \beta y$ (με $\alpha, \beta \in \mathcal{R}$) έχει το πολύ δύο πραγματικές λύσεις για άρτιο k , και το πολύ τρεις για περιττό k . Υποθέτοντας ότι $X \sim F$ και $X \in \mathcal{F}_k$, προκύπτει από την (2.58) ότι η X παίρνει το πολύ δύο τιμές όταν k άρτιος και εφόσον η X είναι μη–εκφυλισμένη, θα λαμβάνει τελικά ακριβώς δύο τιμές, και επίσης, παίρνει δύο ή τρεις τιμές όταν k περιττός. Τούτο είναι απόρροια του γεγονότος

ότι τα σημεία αύξησης της F δεν μπορεί να είναι περισσότερα από τρία, όταν k περιττός, και πάνω από δύο, όταν k άρτιος. Για να επιβεβαιώσουμε αυτόν τον ισχυρισμό, ας θεωρήσουμε ότι k είναι περιττός, $X \sim \mathcal{F}_k$, $E(X) = \mu$, $E(X - \mu)^k = \mu_k$ και $E(X - \mu)^{k-1} = \mu_{k-1}$. Έστω $x_1 < x_2 < x_3 < x_4$ είναι τέσσερα διακριτά σημεία αύξησης της F . Τότε, υπάρχει τουλάχιστον ένα x_i για το οποίο $(x_i - \mu)^k - \mu_k - k\mu_{k-1}(x_i - \mu) \neq 0$, και έτσι δυνάμεθα να βρούμε ένα μικρό θετικό αριθμό $\varepsilon > 0$, έτσι ώστε, $(x - \mu)^k \neq \mu_k + k\mu_{k-1}(x - \mu)$ για όλα τα $x \in (x_i - \varepsilon, x_i + \varepsilon]$. Οπότε,

$$P[x_i - \varepsilon < X \leq x_i + \varepsilon] \leq P[(X - \mu)^k \neq \mu_k + k\mu_{k-1}(X - \mu)].$$

Εντούτοις, αφού το x_i είναι σημείο αύξησης της F , έχουμε

$$\begin{aligned} 0 < F(x_i + \varepsilon) - F(x_i - \varepsilon) &= P[x_i - \varepsilon < X \leq x_i + \varepsilon] \\ &\leq P[(X - \mu)^k \neq \mu_k + k\mu_{k-1}(X - \mu)], \end{aligned}$$

αυτό το πράγμα, όμως, είναι άτοπο διότι, αντιτίθεται στην (2.58). Άρα η X λαμβάνει το πολύ τρεις τιμές. Τώρα, τα ίδια επιχειρήματα εφαρμόζονται και στην περίπτωση όπου το k είναι άρτιος. Έτσι, πλέον, είμαστε σε θέση να διατυπώσουμε την εξής πρόταση.

Πρόταση 2.5: Αν $k \geq 2$ είναι άρτιος τότε το $\mathcal{F}_k$ περιέχει μόνο δίτιμες τυχαίες μεταβλητές. Αν $k \geq 3$ είναι περιττός τότε το $\mathcal{F}_k$ περιέχει είτε δίτιμες είτε τρίτιμες τυχαίες μεταβλητές.

Ο στόχος μας είναι να περιγράψουμε όλες τις ιδιάζουσες κατανομές και να λάβουμε ένα, δεύτερης τάξης, μη-εκφυλισμένο, όριο κατανομής για την ποσότητα $M_{k,n} - \mu_k$, καθώς $n \rightarrow \infty$. Πρώτα, θεωρούμε τις δίτιμες κατανομές επειδή αυτές είναι πιθανά στοιχεία του $\mathcal{F}_k$.

Λήμμα 2.4: Έστω $X \sim b(p)$, δηλαδή, $P(X = 1) = p = 1 - P(X = 0)$ για κάποιο $p \in (0, 1)$. Τότε, $X \in \mathcal{F}_2$ όταν και μόνο όταν $p = 1/2$. Επίσης, αν $k \geq 4$ είναι άρτιος τότε $X \in \mathcal{F}_k$ αν και μόνο αν $p \in \{1 - p_k, \frac{1}{2}, p_k\}$, όπου p_k είναι η μοναδική ρίζα της εξίσωσης

$$\left(\frac{p}{1-p}\right)^{k-1} = \frac{(k+1)p-1}{k-(k+1)p}, \quad \frac{k-2}{k-1} < p < \frac{k}{k+1}, \quad (2.59)$$

για παράδειγμα, $p_4 = \frac{1}{2} + \frac{\sqrt{3}}{6}$ και $p_6 = \frac{1}{2} + \frac{\sqrt{15(4\sqrt{10}-5)}}{30}$.

Απόδειξη: Έχουμε $\mu = E(X) = p$, επομένως,

$$\begin{aligned}\mu_k &= E(X - \mu)^k = E(X - p)^k = \sum_{x=0}^1 (x - p)^k P(X = x) \\ &= (-p)^k(1 - p) + (1 - p)^k p = (-1)^k p^k(1 - p) + (1 - p)^k p \\ &= p(1 - p)\{(-1)^k p^{k-1} + (1 - p)^{k-1}\}, \quad k = 1, 2, \dots\end{aligned}\quad (2.60)$$

Η (2.58) δείχνει ότι $X \in \mathcal{F}_k$ αν και μόνο αν

$$(x - p)^k = \mu_k + k\mu_{k-1}(x - p) \quad \text{για } x = 0 \text{ και } x = 1. \quad (2.61)$$

Χρησιμοποιώντας την (2.60) και το γεγονός ότι $k \geq 2$ είναι άρτιος, για $x = 0$ στην (2.61) προκύπτει ότι

$$\begin{aligned}(-p)^k &= \mu_k + k\mu_{k-1}(-p) \\ &= p(1 - p)^k + (-p)^k(1 - p) + [kp(1 - p)^{k-1} + (-p)^{k-1}k(1 - p)](-p) \\ &= p(1 - p)^k + (-p)^k(1 - p) - kp^2(1 - p)^{k-1} + k(-p)^k(1 - p) \\ &\Rightarrow p(1 - p)^k + (-p)^k(1 - p) = kp^2(1 - p)^{k-1} - k(-p)^k(1 - p) + (-p)^k, \quad (2.62)\end{aligned}$$

και για $x = 1$ στην (2.61) θα έχουμε

$$\begin{aligned}(1 - p)^k &= \mu_k + k\mu_{k-1}(1 - p) \\ &= p(1 - p)^k + (-p)^k(1 - p) + kp(1 - p)^k + k(-p)^{k-1}(1 - p)^2 \\ &\Rightarrow p(1 - p)^k + (-p)^k(1 - p) = (1 - p)^k - kp(1 - p)^k - k(-p)^{k-1}(1 - p)^2.\end{aligned}\quad (2.63)$$

Από τις δύο τελευταίες σχέσεις αντιλαμβανόμαστε ότι τα πρώτα μέλη είναι ίσα άρα, και τα δεύτερα, οπότε

$$\begin{aligned}kp^2(1 - p)^{k-1} - k(-p)^k(1 - p) + (-p)^k \\ &= (1 - p)^k - kp(1 - p)^k - k(-p)^{k-1}(1 - p)^2 \\ &\Rightarrow (1 - p)^{k-1}[kp^2 - (1 - p) + kp(1 - p)] \\ &= (-p)^{k-1}[-kp(1 - p) - (-p) - k(1 - p)^2] \\ &\Rightarrow (1 - p)^{k-1}[kp^2 - 1 + p + kp - kp^2] \\ &= (-p)^{k-1}[-kp + kp^2 + p - k + 2kp - kp^2] \\ &\Rightarrow (1 - p)^{k-1}[-1 + p(k + 1)] = (-p)^{k-1}[kp - k + p] \\ &\Rightarrow (1 - p)^{k-1}[-1 + p(k + 1)] = (-1)^{k-1}p^{k-1}[kp - k + p],\end{aligned}$$

όμως, k είναι άρτιος, έτσι, $(-1)^{k-1} = -1$. Επομένως, τελικά, οι σχέσεις (2.62), (2.63) ανάγονται σε μία

$$(1-p)^{k-1}[-1+p(k+1)] = p^{k-1}[-kp+k-p]$$

$$\Rightarrow p^{k-1}[k-(k+1)p] = (1-p)^{k-1}[-1+p(k+1)], \quad 0 < p < 1. \quad (2.64)$$

Γίνεται αντιληπτό ότι η τιμή $p = \frac{k}{k+1}$ αποτελεί ρίζα του αριστερού μέλους της εξίσωσης (2.64), γεγονός που δεν υφίσταται και για το δεξιό. Έτσι, καταλήγουμε ότι οι σχέσεις (2.59) και (2.64) είναι ισοδύναμες. Προφανώς, το $p = \frac{1}{2}$ είναι ρίζα της (2.64). Προκειμένου να βρούμε όλες τις λύσεις της (2.64), δημιουργούμε τον μετασχηματισμό $t = \frac{p}{1-p} \Rightarrow p = \frac{t}{t+1}$, ο οποίος απεικονίζει μονότονα το $p \in (0, 1)$ σε $t \in (0, +\infty)$. Με αυτό τον τρόπο από την (2.59) παίρνουμε τη νέα εξίσωση

$$t^{k-1} = \frac{-1 + (k+1)\frac{t}{t+1}}{k - (k+1)\frac{t}{t+1}} = \frac{(-1-t+kt+t)/(t+1)}{(k+kt-kt-t)/(t+1)}$$

$$= \frac{-1-t+kt+t}{k+kt-kt-t} = \frac{-1+kt}{k-t} \Rightarrow$$

$$\Rightarrow kt^{k-1} - t^k = -1 + kt \Rightarrow$$

$$\Rightarrow p_k(t) := t^k - kt^{k-1} + kt - 1 = 0, \quad 0 < t < \infty \quad (2.65)$$

που έχει προφανή ρίζα την $t = 1$ (αντίστοιχα, $p = \frac{1}{2}$ για την (2.64)). Αν $k = 2$ τότε η (2.65) γράφεται ως $t^2 - 1 = 0 \Rightarrow t = \pm 1$. Ωστόσο, η $t = -1$ απορρίπτεται επειδή $0 < t < \infty$, επομένως, η $t = 1$ είναι μοναδική λύση για την (2.65) (αντιστοίχως, $p = \frac{1}{2}$ για την (2.64)). Όμως, για κάθε $t > 0$ ισχύει

$$p_k\left(\frac{1}{t}\right) := t^{-k} - kt^{1-k} + kt^{-1} - 1 = -\frac{t^k - kt^{k-1} + kt - 1}{t^k} = \frac{-p_k(t)}{t^k},$$

άρα, έπεται ότι το $1/t$ είναι μία ρίζα της (2.65) για οποιοδήποτε ρίζα t (αντίστοιχα, $1-p$ είναι μία λύση της (2.64)). Για κάθε άρτιο $k \geq 4$ παρατηρούμε ότι

$$p_k(0) := 1, \quad p_k(1) := 0 \quad \text{και} \quad p_k(+\infty) := \lim_{t \rightarrow +\infty} p_k(t) = +\infty.$$

Επίσης, η πρώτη παράγωγος της συνάρτησης $p_k(t)$ είναι

$$p'_k(t) = kt^{k-1} - k(k-1)t^{k-2} + k = k(t^{k-1} - (k-1)t^{k-2} + 1) = kq_k(t),$$

όπου για την $q_k(t) = t^{k-1} - (k-1)t^{k-2} + 1$ εφαρμόζουμε, διαδοχικά, στα διαστήματα $[0, 1]$, $[1, +\infty]$ το Θεώρημα του Bolzano³⁵. Συγκεκριμένα, η συνάρτηση q_k είναι ορισμένη στα κλειστά διαστήματα $[0, 1]$, $[1, +\infty]$. Επιπροσθέτως,

- η q_k είναι συνεχής στο $[0, 1]$ και στο $[1, +\infty]$ (ως πολυωνυμική)
- $q_k(0) \cdot q_k(1) < 0$ (αφού $q_k(0) = 1 > 0$ και $q_k(1) = 3 - k \leq -1 < 0$)
- $q_k(1) \cdot q_k(+\infty) < 0$ (διότι $q_k(1) < 0$ και $q_k(+\infty) = +\infty > 0$)

τότε σε αμφότερα ανοικτά διαστήματα $(0, 1)$ και $(1, +\infty)$ αντίστοιχα, για την εξίσωση $q_k(t) = 0$, υπάρχει από μία, τουλάχιστον, ρίζα $\rho_1 \in (0, 1)$ και $\rho_2 \in (1, +\infty)$ έτσι ώστε $q_k(\rho_1) = 0 = q_k(\rho_2)$. Επιπλέον, διαπιστώνουμε ότι η πρώτη παράγωγος της $q_k(t)$, ήτοι, η

$$q'_k(t) = (k-1)t^{k-2} - (k-1)(k-2)t^{k-3} = (k-1)t^{k-3}[t - (k-2)]$$

είναι αρνητική για κάθε $t \in (0, k-2)$ και θετική για κάθε $t \in (k-2, +\infty)$. Αυτό έχει ως επακόλουθο, η $q_k(t)$ να παρουσιάζει μείωση στο $(0, k-2)$ και αύξηση στο $(k-2, +\infty)$. Οπότε, η $q_k(t)$ είναι γνησίως αύξουσα και γνησίως φθίνουσα αντίστοιχα, στα ανοικτά διαστήματα $(0, k-2)$ και $(k-2, +\infty)$. Με άλλα λόγια, υπάρχουν δύο λύσεις $\rho_1 < \rho_2$, με $0 < \rho_1 < 1 < k-2 < \rho_2 < +\infty$, τέτοιες ώστε $q_k(t) > 0$ για κάθε $t \in (0, \rho_1) \cup (\rho_2, +\infty)$ και $q_k(t) < 0$ για κάθε $t \in (\rho_1, \rho_2)$.

Από τη σχέση $q_k(t) = \frac{p'_k(t)}{k}$ απορρέει ότι η $p_k(t)$ αυξάνεται στο $(0, \rho_1)$, μειώνεται στο (ρ_1, ρ_2) και αυξάνεται στο $(\rho_2, +\infty)$. Όμως, λόγω του ότι $1 \in (0, \rho_1)$ και $p_k(1) = 0$ συμπεραίνουμε ότι $p_k(\rho_1) > 0$ και $p_k(\rho_2) < 0$, έτσι, υπάρχει μοναδικό $t_1 \in (0, \rho_1)$ και μοναδικό $t_2 \in (\rho_2, +\infty)$ τέτοια ώστε $p_k(t_1) = 0 = p_k(t_2)$. Επειδή, $t_1 = 1/t_2$, το σύνολο όλων των λύσεων της (2.65) είναι το $\{\frac{1}{t_2}, 1, t_2\}$ και τοιουτοτρόπως, το σύνολο όλων των λύσεων της (2.64) είναι το $\{1 - p_k, 1/2, p_k\}$, όπου $p_k = \frac{t_2}{1+t_2}$. Τελικά, η σχέση $t_2 > \rho_2 > k-2$ καταδεικνύει ότι $p_k = \frac{t_2}{1+t_2} > \frac{k-2}{1+(k-2)} = \frac{k-2}{k-1}$. Επίσης, ισχύει $p_k < \frac{k}{k+1}$ αφού για κάθε $p \geq \frac{k}{k+1}$ το αριστερό μέλος της (38) είναι μη-θετικό ενώ το δεξιό μέλος είναι αυστηρά θετικό. ■

³⁵Βοημός μαθηματικός, δηλαδή, κάτοικος μίας ιστορικής περιοχής της κεντρικής Ευρώπης, της Βοημίας που καταλαμβάνει σχεδόν τα κεντρικά και δυτικά 2/3 της σημερινής έκτασης της Τσεχίας. Επιπλέον, ήταν δάσκαλος της λογικής, θεολόγος, καθολικός ιερέας και αντιμιλιταριστής της Γερμανικής μητρικής γλώσσας με πλήρες όνομα *Bernhard Placidus Johann Nepomuk Bolzano*, 5 Οκτωβρίου 1781 - 18 Δεκεμβρίου 1848.

Από την (2.59) βλέπουμε ότι $\frac{1}{2} < p_4 < p_6 < \dots$ και $p_{2k} = 1 - \frac{1}{2k} + o(1/k)$, καθώς $k \rightarrow \infty$. Το προηγούμενο Λήμμα περιγράφει πλήρως όλα τα σύνολα των ιδιαζόντων τυχαίων μεταβλητών $\mathcal{F}_k$ στην περίπτωση που ο k είναι άρτιος.

Πόρισμα 2.2: Εάν $k \geq 2$ είναι άρτιος τότε $X \in \mathcal{F}_k^0$ όταν και μόνο όταν

$$P(X = \pm 1) = \frac{1}{2}$$

ή

$$P\left(X = -\sqrt{\frac{p_k}{1-p_k}}\right) = 1 - p_k \text{ και } P\left(X = \sqrt{\frac{1-p_k}{p_k}}\right) = p_k \text{ για } k \in \{4, 6, \dots\}$$

ή

$$P\left(X = -\sqrt{\frac{1-p_k}{p_k}}\right) = p_k \text{ και } P\left(X = \sqrt{\frac{p_k}{1-p_k}}\right) = 1 - p_k \text{ για } k \in \{4, 6, \dots\},$$

όπου το $p_k \in \left(\frac{k-2}{k-1}, \frac{k}{k+1}\right)$ δίνεται από την σχέση (2.59).

Το Πόρισμα 2.2 υποδεικνύει ότι το $\mathcal{F}_2^0$ είναι ένα μονοσύνολο και ότι για κάθε $k \in \{4, 6, \dots\}$ το $\mathcal{F}_k^0$ συνίσταται ακριβώς από τρεις δίτιμες κατανομές. Από την άλλη, όταν k είναι περιττός τότε η φύση του $\mathcal{F}_k$ είναι αρκετά διαφορετική γιατί περιέχει τόσο δίτιμες όσο και τρίτιμες κατανομές. Πρώτα εξετάζουμε την περίπτωση των δίτιμων κατανομών.

Λήμμα 2.5: Έστω $X \sim b(p)$ για κάποιο $p \in (0, 1)$. Αν $X \in \mathcal{F}_k$ και $k \geq 3$ είναι περιττός τότε $p \in \{1 - p_k, p_k\}$, όπου p_k είναι η μοναδική ρίζα της εξίσωσης

$$\left(\frac{p}{1-p}\right)^{k-1} = \frac{(k+1)p-1}{(k+1)p-k}, \quad \frac{k}{k+1} < p < 1, \quad (2.66)$$

για παράδειγμα, $p_3 = \frac{1}{2} + \frac{\sqrt{3}}{6}$ και $p_5 = \frac{1}{2} + \frac{\sqrt{5\sqrt{5}}}{10}$. Αντίστροφα, αν $k \geq 3$ είναι περιττός και $X \sim b(p_k)$ ή $X \sim b(1 - p_k)$, με p_k όπως παραπάνω, τότε $X \in \mathcal{F}_k$.

Απόδειξη: Υποθέτουμε ότι $k \geq 3$ είναι περιττός, $X \sim b(p)$ και $X \in \mathcal{F}_k$. Αυτό σημαίνει ότι η (2.61) ικανοποιείται. Χρησιμοποιώντας την (2.60) και δεδομένου ότι $k \geq 3$ είναι περιττός, απορρέει με τον ίδιο τρόπο, η σχέση

$$(1-p)^{k-1}[-1+p(k+1)] = (-1)^{k-1}p^{k-1}[kp-k+p],$$

όμως, k είναι περιττός, άρα, $(-1)^{k-1} = 1$. Επομένως, τελικά, οι δύο αυτές σχέσεις που προκύπτουν από τις τιμές $x = 0$ και $x = 1$, ανάγονται σε μία

$$(1-p)^{k-1}[-1+p(k+1)] = p^{k-1}[kp-k+p]$$

$$\Rightarrow p^{k-1}[(k+1)p-k] = (1-p)^{k-1}[(k+1)p-1], \quad 0 < p < 1. \quad (2.67)$$

Καθίσταται σαφές, ότι η τιμή $p = \frac{k}{k+1}$ είναι ρίζα του αριστερού μέλους της εξίσωσης (2.67), κάτι που όμως, δεν υφίσταται και για το δεξιό. Οπότε, συμπεραίνουμε ότι οι σχέσεις (2.66) και (2.67) είναι ισοδύναμες. Όπως στο Λήμμα 2.4, έτσι και' δω, προκειμένου να βρούμε όλες τις λύσεις της (2.67), θεωρούμε τον μετασχηματισμό $t = \frac{p}{1-p} \Rightarrow p = \frac{t}{t+1}$, που απεικονίζει μονοσήμαντα το $p \in (0, 1)$ σε $t \in (0, +\infty)$. Με αυτό τον τρόπο, από την (2.66) λαμβάνουμε τη νέα εξίσωση

$$t^{k-1} = \frac{(k+1)\frac{t}{t+1} - 1}{(k+1)\frac{t}{t+1} - k} = \frac{(kt+t-t-1)/(t+1)}{(kt+t-kt-k)/(t+1)}$$

$$= \frac{kt+t-t-1}{kt+t-kt-k} = \frac{kt-1}{t-k} \Rightarrow$$

$$\Rightarrow t^k - kt^{k-1} = kt-1 \Rightarrow$$

$$\Rightarrow p_k(t) := t^k - kt^{k-1} - kt + 1 = 0, \quad 0 < t < \infty. \quad (2.68)$$

Εφόσον, για κάθε $t > 0$ ισχύει

$$p_k\left(\frac{1}{t}\right) := t^{-k} - kt^{1-k} - kt^{-1} + 1 = \frac{t^k - kt^{k-1} - kt + 1}{t^k} = \frac{p_k(t)}{t^k},$$

προκύπτει ότι το $1/t$ είναι μία λύση της (2.68) για οποιοδήποτε ρίζα t (αντίστοιχα, $1-p$ είναι μία ρίζα της (2.67) για κάθε ρίζα p). Για κάθε περιττό $k \geq 3$ παρατηρούμε ότι

$$p_k(0) := 1 > 0, \quad p_k(1) := 2 - 2k < 0 \quad \text{και} \quad p_k(+\infty) := \lim_{t \rightarrow +\infty} p_k(t) = +\infty.$$

Άρα, η (2.68) διαθέτει, τουλάχιστον, μία ρίζα στο $(0,1)$ και τουλάχιστον, μία ρίζα στο $(1, +\infty)$. Τώρα, η πρώτη παράγωγος της συνάρτησης $p_k(t)$ είναι

$$p'_k(t) = kt^{k-1} - k(k-1)t^{k-2} - k = k(t^{k-1} - (k-1)t^{k-2} - 1) = kq_k(t),$$

όπου για την $q_k(t) = t^{k-1} - (k-1)t^{k-2} - 1$ πραγματοποιούμε το Θεώρημα του Bolzano στο διάστημα $[0, +\infty]$. Συγκεκριμένα, η συνάρτηση q_k είναι ορισμένη στο κλειστό διάστημα $[0, +\infty]$. Επιπλέον,

- η q_k είναι συνεχής στο $[0, +\infty]$ (ως πολυωνυμική)
- $q_k(0) \cdot q_k(+\infty) < 0$ (διότι $q_k(0) = -1 < 0$ και $q_k(+\infty) = +\infty > 0$)

τότε, για την εξίσωση $q_k(t) = 0$, υπάρχει μία, τουλάχιστον, ρίζα $\rho \in (0, +\infty)$ τέτοια ώστε $q_k(\rho) = 0$. Επιπλέον, η πρώτη παράγωγος της $q_k(t)$, ήτοι, η

$$q'_k(t) = (k-1)t^{k-2} - (k-1)(k-2)t^{k-3} = (k-1)t^{k-3}[t - (k-2)]$$

είναι αρνητική για κάθε $t \in (0, k-2)$ και θετική για κάθε $t \in (k-2, +\infty)$. Αυτό σημαίνει ότι η $q_k(t)$ παρουσιάζει μείωση στο $(0, k-2)$ και αύξηση στο $(k-2, +\infty)$. Οπότε, η $q_k(t)$ είναι γνησίως αύξουσα και γνησίως φθίνουσα αντίστοιχα, στα ανοικτά διαστήματα $(0, k-2)$ και $(k-2, +\infty)$. Ως εκ τούτου, υπάρχει μοναδική λύση $\rho > k-2 \geq 1$ έτσι ώστε $q_k(t) < 0$ για κάθε $t \in (0, \rho)$ και $q_k(t) > 0$ για κάθε $t \in (\rho, +\infty)$. Από τη σχέση $q_k(t) = \frac{p'_k(t)}{k}$ προκύπτει ότι η $p_k(t)$ μειώνεται στο $(0, \rho)$ και αυξάνεται στο $(\rho, +\infty)$. Επειδή, όμως, $p_k(0) > 0$ και $p_k(+\infty) > 0$ αποφαινόμεστε ότι υπάρχουν μοναδικά $t_1 \in (0, 1)$ και $t_2 \in (\rho, +\infty)$ τέτοια ώστε $p_k(t_1) = 0 = p_k(t_2)$. Λόγω του ότι, $t_1 = 1/t_2$, το σύνολο όλων των λύσεων της (41) είναι το $\{1 - p_k, p_k\}$, όπου $p_k = \frac{t_2}{1+t_2}$. Εν τέλει, η σχέση $t_2 > \rho_2 > k-2$ αποδεικνύει ότι $p_k = \frac{t_2}{1+t_2} > \frac{k-2}{1+(k-2)} = \frac{k-2}{k-1}$. Από την άλλη, η ρίζα p_k δεν δύναται να ανήκει στο $\left(\frac{k-2}{k-1}, \frac{k}{k+1}\right]$ διότι για όλα τα p , σε αυτό το διάστημα, το αριστερό μέλος της εξίσωσης (41) είναι μη-θετικό ενώ το δεξιό μέλος είναι αυστηρά θετικό (εφόσον $p > \frac{k-2}{k-1}$, ισχύει ότι $(k+1)p - 1 > (k+1)\frac{k-2}{k-1} - 1 = \frac{(k-1)(k+2)+4}{k-1} > 0$, αφού $k \geq 3$). Τούτο επιβεβαιώνει ότι $p_k > \frac{k}{k+1}$, και έτσι, αν $X \sim b(p_k)$ ή $X \sim b(1-p_k)$ τότε οι σχέσεις (2.60) και (2.67) δείχνουν ότι οι εξισώσεις (2.61) και (2.58) πληρούνται, δηλαδή, $X \in \mathcal{F}_k$.

■

Πόρισμα 2.3: Εάν $k \geq 3$ είναι περιττός τότε οι μοναδικές δίτιμες τυχαίες μεταβλητές, που περιέχονται στο σύνολο όλων των τυποποιημένων ιδιαζόντων τυχαίων μεταβλητών τάξης k , $\mathcal{F}_k^0$, είναι οι ακόλουθες:

$$P\left(X = -\sqrt{\frac{p_k}{1-p_k}}\right) = 1 - p_k, \quad P\left(X = \sqrt{\frac{1-p_k}{p_k}}\right) = p_k$$

και

$$P\left(X = -\sqrt{\frac{1-p_k}{p_k}}\right) = p_k, \quad P\left(X = \sqrt{\frac{p_k}{1-p_k}}\right) = 1 - p_k,$$

όπου $p_k \in \left(\frac{k}{k+1}, 1\right)$ δίνεται από την σχέση (2.66).

Το Πόρισμα 2.3 περιγράφει όλες τις δίτιμες τυχαίες μεταβλητές του $\mathcal{F}_k^0$ όταν $k \geq 3$ είναι περιττός, εντούτοις, έχουμε ήδη υπόψη μας ότι το $\mathcal{F}_k^0$ περιλαμβάνει επίσης και κάποιες τρίτιμες τυχαίες μεταβλητές. Μεταξύ αυτών, ακριβώς μία είναι συμμετρική.

Λήμμα 2.6: Έστω $k \geq 3$ είναι περιττός τότε η μοναδική συμμετρική τυχαία μεταβλητή του $\mathcal{F}_k^0$ είναι εκείνη που έχει συνάρτηση κατανομής

$$P(X = \pm\sqrt{k}) = \frac{1}{2k}, \quad P(X = 0) = 1 - \frac{1}{k}.$$

Πιο γενικά, αυτή είναι η μοναδική τυχαία μεταβλητή του $\mathcal{F}_k^0$ με $\mu_k = 0$.

Απόδειξη: Εφόσον, $X \in \mathcal{F}_k^0$ έχουμε $\mu = E(X) = 0$ και $E(X^2) = Var(X) + E(X)^2 = 1 + 0^2 = 1$. Οπότε, δεδομένου ότι $\mu_k = 0$, η (2.58) γίνεται

$$X^k = k\mu_{k-1}X \Rightarrow X(X^{k-1} - k\mu_{k-1}) = 0 \quad \text{με πιθανότητα ένα,}$$

όπου, οι λύσεις της τελευταίας εξίσωσης είναι

$$X = 0 \quad \text{ή} \quad X = \pm(k\mu_{k-1})^{\frac{1}{k-1}}.$$

Επομένως, το στήριγμα της X είναι το υποσύνολο

$$A = \left\{ -(k\mu_{k-1})^{\frac{1}{k-1}}, 0, (k\mu_{k-1})^{\frac{1}{k-1}} \right\},$$

όπου $\mu_{k-1} = E(X^{k-1}) > 0$ διότι η X είναι μη-εκφυλισμένη και $k-1$ είναι άρτιος. Έστω τώρα,

$$p = P(X = 0), \quad p_1 = P\left(X = -(k\mu_{k-1})^{\frac{1}{k-1}}\right) \quad \text{και} \quad p_2 = P\left(X = (k\mu_{k-1})^{\frac{1}{k-1}}\right).$$

Οι πιθανότητες αυτές είναι προφανώς, μη-αρνητικές και φυσικά, αθροίζουν στη μονάδα, δηλαδή, $P(X \in A) = 1 \Rightarrow p + p_1 + p_2 = 1$. Όμως, από υπόθεση ισχύει, $E(X) = 0$, ήτοι,

$$\begin{aligned} \sum_x xP(X = x) &= 0 \\ \Rightarrow -(k\mu_{k-1})^{\frac{1}{k-1}} \cdot P\left(X = -(k\mu_{k-1})^{\frac{1}{k-1}}\right) &+ 0 \cdot P(X = 0) \\ &+ (k\mu_{k-1})^{\frac{1}{k-1}} \cdot P\left(X = (k\mu_{k-1})^{\frac{1}{k-1}}\right) = 0 \end{aligned}$$

$$\Rightarrow -(k\mu_{k-1})^{\frac{1}{k-1}} \cdot p_1 + (k\mu_{k-1})^{\frac{1}{k-1}} \cdot p_2 = 0 \Rightarrow p_1 = p_2$$

έτσι, προκύπτει, $p + p_1 + p_2 = 1 \Rightarrow p + 2p_1 = 1 \Rightarrow p_1 = \frac{1-p}{2} = p_2$ και τελικά,
 $P\left(X = \pm(k\mu_{k-1})^{\frac{1}{k-1}}\right) = \frac{1-p}{2}$.

Επιπρόσθετα,

$$\begin{aligned} \mu_{k-1} &= E(X - \mu)^{k-1} = E(X)^{k-1} = \sum_x x^{k-1} P(X = x) = \\ &= \left[-(k\mu_{k-1})^{\frac{1}{k-1}} \right]^{k-1} \cdot P\left(X = -(k\mu_{k-1})^{\frac{1}{k-1}}\right) + 0^{k-1} \cdot P(X = 0) \\ &\quad + \left[(k\mu_{k-1})^{\frac{1}{k-1}} \right]^{k-1} \cdot P\left(X = (k\mu_{k-1})^{\frac{1}{k-1}}\right) \end{aligned}$$

όμως, το k είναι περιττός, έτσι, το $k - 1$ θα είναι άρτιος, άρα,

$$\begin{aligned} \mu_{k-1} &= k\mu_{k-1}p_1 + k\mu_{k-1}p_2 \Rightarrow \mu_{k-1} = 2k\mu_{k-1}p_1 \\ \Rightarrow \mu_{k-1} &= 2k\mu_{k-1} \frac{1-p}{2} = (1-p)k\mu_{k-1} \Rightarrow p = 1 - \frac{1}{k}. \end{aligned}$$

Οπότε,

$$P\left(X = \pm(k\mu_{k-1})^{\frac{1}{k-1}}\right) = \frac{1-p}{2} = \frac{1 - \left(1 - \frac{1}{k}\right)}{2} = \frac{1}{2k}.$$

Τέλος,

$$\begin{aligned} 1 &= E(X^2) = \sum_x x^2 P(X = x) = \left[-(k\mu_{k-1})^{\frac{1}{k-1}} \right]^2 \cdot p_1 + 0 \cdot p + \left[(k\mu_{k-1})^{\frac{1}{k-1}} \right]^2 \cdot p_2 \\ &= (k\mu_{k-1})^{\frac{2}{k-1}} \cdot p_1 + (k\mu_{k-1})^{\frac{2}{k-1}} \cdot p_2 = (k\mu_{k-1})^{\frac{2}{k-1}} \cdot (p_1 + p_2) \\ &= (k\mu_{k-1})^{\frac{2}{k-1}} \cdot (1-p) = (k\mu_{k-1})^{2/(k-1)} \cdot \frac{1}{k} = \frac{\left[(k\mu_{k-1})^{\frac{1}{k-1}} \right]^2}{k}. \end{aligned}$$

Ορίζοντας $\alpha = (k\mu_{k-1})^{\frac{1}{k-1}} > 0$ θα λάβουμε $1 = \frac{\alpha^2}{k} \Rightarrow \alpha = \sqrt{k}$. Αφετέρου, είναι εύκολο να δούμε ότι για αυτή την τιμή του $\alpha = \sqrt{k}$ ισχύει ότι $\mu_k = 0$. Επίσης, $\alpha = (k\mu_{k-1})^{\frac{1}{k-1}} \Rightarrow \alpha^{k-1} = k\mu_{k-1} \Rightarrow \mu_{k-1} = k^{-1}\alpha^{k-1} = k^{-1}(k^{1/2})^{k-1}k^{\frac{k-1}{2}-1} \Rightarrow \mu_{k-1} = k^{(k-3)/2}$. Οπότε, $k\mu_{k-1} = k^{(k-1)/2} = (\pm\sqrt{k})^{k-1}$, έτσι, το στήριγμα παίρνει τη μορφή $A = \{-\sqrt{k}, 0, \sqrt{k}\}$ και $x(x^{k-1} - k\mu_{k-1}) = x(x^{k-1} - (\pm\sqrt{k})^{k-1}) \equiv 0$ για όλα τα $x \in A$. ■

Το ακόλουθο Λήμμα παρουσιάζει μία πλήρη περιγραφή όλων των τρίτιμων κατανομών του $\mathcal{F}_3^0$ και παρέχει μία εικόνα της φύσης του συνόλου $\mathcal{F}_k^0$ όταν $k \geq 3$ είναι περιττός.

Λήμμα 2.7: Για κάθε $\mu_3 \in [-\sqrt{2}, \sqrt{2}]$ υπάρχει μία μοναδική τυχαία μεταβλητή $X \in \mathcal{F}_3^0$ έτσι ώστε $E(X^3) = \mu_3$. Οι περιπτώσεις $\mu_3 = \pm\sqrt{2}$ αντιστοιχούν στις δίτιμες κατανομές που παρατέθηκαν στο Πόρισμα 2.3 για $k = 3$. Οποιαδήποτε άλλη τιμή του $\mu_3 \in (-\sqrt{2}, \sqrt{2})$ προσδιορίζει μοναδικά μία τρίτιμη κατανομή, και ειδικά, η $\mu_3 = 0$ αντιστοιχεί στην συμμετρική κατανομή του Λήμματος 2.6 για $k = 3$. Εξάλλου, στο σύνολο $\mathcal{F}_3^0$ δεν ανήκουν άλλες τυχαίες μεταβλητές. Ως εκ τούτου, το $\mathcal{F}_3^0$ επιδέχεται την παραμετροποίηση

$$\mathcal{F}_3^0 = \{X_\theta, -\sqrt{2} \leq \theta \leq \sqrt{2}\}, \quad (2.69)$$

όπου η τυχαία μεταβλητή X_θ χαρακτηρίζεται από τις σχέσεις

$$E(X_\theta) = 0, E(X_\theta^2) = 1, E(X_\theta^3) = \theta \text{ και } P[X_\theta(X_\theta^2 - 3) = \theta] = 1. \quad (2.70)$$

Απόδειξη: Έστω $X \in \mathcal{F}_3^0$ και υποθέτουμε ότι $\mu_3 = E(X^3) = \theta$ τότε $\mu = E(X) = 0$, $\sigma^2 = \mu_2 = E(X^2) = \text{Var}(X) + E(X)^2 = 1$ και σύμφωνα με την εξίσωση (2.58) συνάγουμε ότι $X(X^2 - 3) = \theta$ με πιθανότητα ένα. Συνεπώς, εφόσον η X είναι μη-εκφυλισμένη, το στήριγμα της X πρέπει να περιέχει τουλάχιστον δύο στοιχεία τα οποία περιλαμβάνονται στο σύνολο των ριζών της $y(y^2 - 3) = \theta$. Τούτο υποδεικνύει ότι $|\theta| \leq 2$ αφού, διαφορετικά, το σύνολο $\{y \in \mathbb{R}: y(y^2 - 3) = \theta\}$ είναι μονοσύνολο. Συνειδητοποιούμε ότι το $\theta = 0$ οδηγεί, μονοσήμαντα, στην συμμετρική τυχαία μεταβλητή του Λήμματος 2.6 με $k = 3$. Έτσι, τούδε και στο εξής υποθέτουμε ότι $\theta \neq 0$. Οι τιμές $\theta = \pm 2$ είναι αδύνατες επειδή οι εξισώσεις $y(y^2 - 3) = \pm 2$ διαθέτουν ακριβώς δύο πραγματικές λύσεις, ως είναι α, β , με $|\alpha| = 1$ και $|\beta| = 2$, ούτως ώστε να καθίστανται ανέφικτες οι σχέσεις $E(X) = 0$ και $E(X^2) = 1$. Πράγματι, για τις δύο αυτές εξισώσεις παίρνουμε διαδοχικά

$$\begin{cases} y(y^2 - 3) = +2 \\ y(y^2 - 3) = -2 \end{cases} \Rightarrow \begin{cases} y^3 - 3y - 2 = 0 \\ y^3 - 3y + 2 = 0 \end{cases} \Rightarrow \begin{cases} (y + 1)^2(y - 2) = 0 \\ (y - 1)^2(y + 2) = 0 \end{cases}$$

που έχουν ρίζες τις ± 1 και ± 2 .

Κατόπιν, διακρίνουμε δύο περιπτώσεις, την $-2 < \theta < 0$ και την $0 < \theta < 2$. Ας εξετάσουμε την πρώτη. Στην περίπτωση αυτή έχουμε $\{y: y(y^2 - 3) = \theta\} = \{-\alpha, \beta, \gamma\}$ όπου $0 < \beta < 1 < \gamma < \sqrt{3} < \alpha$ και, εξ' ορισμού, οι αριθμοί α, β, γ πληρούν την σχέση

$$-\alpha(\alpha^2 - 3) = \beta(\beta^2 - 3) = \gamma(\gamma^2 - 3) = \theta. \quad (2.71)$$

Από την τελευταία εξίσωση αντιλαμβανόμαστε ότι η $\theta = \theta(\beta) = \beta(\beta^2 - 3)$ είναι μία γνησίως φθίνουσα και συνεχής συνάρτηση του β η οποία απεικονίζει το $\beta \in (0, 1)$ σε $\theta \in (-2, 0)$, άρα η αντίστροφη συνάρτηση, $\beta(\theta): (-2, 0) \rightarrow (0, 1)$, είναι καλώς-ορισμένη, συνεχής και γνησίως φθίνουσα ως προς θ με $\beta(-2_+) = 1$ και $\beta(0_-) = 0$. Εν συνεχεία, από την (2.71) προκύπτει η δευτεροβάθμια εξίσωση

$$\begin{aligned} 3(\alpha + \beta) &= \alpha^3 + \beta^3 = (\alpha + \beta)(\alpha^2 - \alpha\beta + \beta^2) \\ \Rightarrow \alpha^2 - \alpha\beta + \beta^2 &= 3 \Rightarrow \alpha^2 - \alpha\beta + \beta^2 - 3 = 0 \end{aligned}$$

με λύσεις

$$\alpha_{1,2} = \frac{\beta \pm \sqrt{3(4 - \beta^2)}}{2},$$

όπου επειδή $\alpha > \beta > \frac{\beta}{2} > 0$, η $\frac{\beta + \sqrt{3(4 - \beta^2)}}{2} > 0$ γίνεται δεκτή, ενώ, η $\frac{\beta - \sqrt{3(4 - \beta^2)}}{2}$ απορρίπτεται, αφού, αν υποθέσουμε ότι $\frac{\beta - \sqrt{3(4 - \beta^2)}}{2} > 0$ τότε $\beta > \sqrt{3(4 - \beta^2)} \Rightarrow \beta^2 > 12 - 3\beta^2 \Rightarrow \beta^2 > 3 \Rightarrow \beta > \sqrt{3}$ ή $\beta < -\sqrt{3}$, δηλαδή, οδηγούμαστε σε άτοπο, διότι, $0 < \beta < 1$. Επομένως, $\frac{\beta - \sqrt{3(4 - \beta^2)}}{2} < 0$ και έτσι

$$\alpha = \alpha(\beta) = \frac{\beta + \sqrt{3(4 - \beta^2)}}{2}. \quad (2.72)$$

Παρομοίως, από την (2.71) προκύπτει η δευτεροβάθμια εξίσωση

$$\begin{aligned} 3(\gamma - \beta) &= \gamma^3 - \beta^3 = (\gamma - \beta)(\gamma^2 + \gamma\beta + \beta^2) \\ \Rightarrow \gamma^2 + \gamma\beta + \beta^2 &= 3 \Rightarrow \gamma^2 + \gamma\beta + \beta^2 - 3 = 0 \end{aligned}$$

με ρίζες

$$\gamma_{1,2} = \frac{-\beta \pm \sqrt{3(4 - \beta^2)}}{2},$$

όπου η $\frac{-\beta - \sqrt{3(4 - \beta^2)}}{2} < 0$ απορρίπτεται, επειδή είναι αρνητική, ενώ, η $\frac{-\beta + \sqrt{3(4 - \beta^2)}}{2}$ γίνεται αποδεκτή, αφού, αν υποθεθεί ότι $\frac{-\beta + \sqrt{3(4 - \beta^2)}}{2} > 0$ τότε $\beta < \sqrt{3(4 - \beta^2)} \Rightarrow \beta^2 < 12 - 3\beta^2 \Rightarrow \beta^2 < 3 \Rightarrow -\sqrt{3} < \beta < \sqrt{3}$, γεγονός που υφίσταται διότι, $-\sqrt{3} < 0 < \beta < 1 < \sqrt{3}$. Επομένως,

$$\gamma = \gamma(\beta) = \frac{-\beta + \sqrt{3(4 - \beta^2)}}{2}. \quad (2.73)$$

Αφαιρώντας κατά μέλη τις σχέσεις (2.72) και (2.73) καταλήγουμε ότι $\alpha = \beta + \gamma$. Κατόπιν, ορίζουμε $p_1 = P(X = -\alpha)$, $p_2 = P(X = \beta)$ και $p_3 = P(X = \gamma)$ και δοθέντος ότι $P(X \in \{-\alpha, \beta, \gamma\}) = 1$ και $E(X) = 0$, $E(X^2) = 1$, παίρνουμε, αντίστοιχα, το σύστημα των τριών εξισώσεων

$$\begin{cases} p_1 + p_2 + p_3 = 1, \\ -\alpha p_1 + \beta p_2 + \gamma p_3 = 0, \\ \alpha^2 p_1 + \beta^2 p_2 + \gamma^2 p_3 = 1, \end{cases}$$

το οποίο, λόγω του ότι $\alpha = \beta + \gamma$, έχει τη μοναδική λύση

$$\begin{cases} p_1 = \frac{1 + \beta\gamma}{(\beta + 2\gamma)(2\beta + \gamma)}, \\ p_2 = \frac{\gamma(\beta + \gamma) - 1}{(\gamma - \beta)(2\beta + \gamma)}, \\ p_3 = \frac{1 - \beta(\beta + \gamma)}{(\gamma - \beta)(\beta + 2\gamma)}. \end{cases}$$

Τώρα, επειδή $\gamma^2 + \gamma\beta + \beta^2 = 3$, ισχύει $\gamma(\beta + \gamma) = 3 - \beta^2$ και $\beta(\beta + \gamma) = 3 - \gamma^2$. Στη συνέχεια, αντικαθιστώντας αυτές τις τιμές στους αριθμητές των πιθανοτήτων p_2 και p_3 παίρνουμε τελικά

$$\begin{cases} p_1 = \frac{1 + \beta\gamma}{(\beta + 2\gamma)(2\beta + \gamma)}, \\ p_2 = \frac{2 - \beta^2}{(\gamma - \beta)(2\beta + \gamma)}, \\ p_3 = \frac{\gamma^2 - 2}{(\gamma - \beta)(\beta + 2\gamma)}. \end{cases} \quad (2.74)$$

Είναι φανερό ότι $p_1 > 0$ και $p_2 > 0$ για όλες, συλλήβδην, τις τιμές β και γ με

$$0 < \beta < 1 < \gamma = \frac{-\beta + \sqrt{3(4 - \beta^2)}}{2} < \sqrt{3}.$$

Εντούτοις, αυτό δεν ισχύει για την πιθανότητα p_3 διότι είναι μη-αρνητική και άρα, απαιτεί $\gamma^2 \geq 2 \Rightarrow \gamma \geq \sqrt{2}$ ή $\gamma \leq -\sqrt{2}$. Όμως, εφόσον $\gamma > 0$ θα έχουμε μόνο $\gamma \geq \sqrt{2}$. Τώρα, λόγω του ότι $\mu_3 = \theta = \gamma(\gamma^2 - 3)$ και $\gamma \in [\sqrt{2}, \sqrt{3})$ συμπεραίνουμε ότι όλες οι εφικτές τιμές του θ ($\theta < 0$) περιλαμβάνονται στο διάστημα $[-\sqrt{2}, 0)$. Επικαλούμενοι την (2.73) και το γεγονός ότι $\gamma \geq \sqrt{2}$ έπεται ότι

$$\begin{aligned} \frac{-\beta + \sqrt{3(4 - \beta^2)}}{2} &\geq \sqrt{2} \Rightarrow \sqrt{3(4 - \beta^2)} \geq 2\sqrt{2} + \beta \\ \Rightarrow 3(4 - \beta^2) &\geq 8 + \beta^2 + 4\sqrt{2}\beta \Rightarrow 4\beta^2 + 4\sqrt{2}\beta - 4 \leq 0 \end{aligned}$$

$$\Rightarrow \beta^2 + \sqrt{2}\beta - 1 \leq 0 \Rightarrow \beta^2 + \sqrt{2}\beta + \frac{1}{2} \leq \frac{3}{2} \Rightarrow \left(\beta + \frac{\sqrt{2}}{2}\right)^2 \leq \frac{3}{2}$$

$$\Rightarrow -\sqrt{\frac{3}{2}} \leq \beta + \frac{\sqrt{2}}{2} \leq \sqrt{\frac{3}{2}} \Rightarrow -\sqrt{\frac{3}{2}} - \frac{\sqrt{2}}{2} \leq \beta \leq \sqrt{\frac{3}{2}} - \frac{\sqrt{2}}{2},$$

ωστόσο, $\beta > 0$, επομένως, $0 < \beta \leq \sqrt{\frac{3}{2}} - \frac{\sqrt{2}}{2} = \frac{\sqrt{6}-\sqrt{2}}{2} = \sqrt{2-\sqrt{3}}$. Επιπλέον, βλέπουμε ότι η τιμή $\gamma = \sqrt{2}$ συνάδει σε μία τυποποιημένη δίτιμη τυχαία μεταβλητή με $\mu_3 = \theta = \gamma(\gamma^2 - 3) = -\sqrt{2}$. Τώρα, λαμβάνοντας τις τιμές $-\alpha = -\beta - \gamma = -\frac{\sqrt{6}+\sqrt{2}}{2} = -\sqrt{2+\sqrt{3}}$ και $\beta = \sqrt{2-\sqrt{3}} = \frac{\sqrt{6}-\sqrt{2}}{2}$ με αντίστοιχες πιθανότητες $1-p$ και p , όπου

$$\begin{aligned} p &= \frac{2 - \beta^2}{(\gamma - \beta)(2\beta + \gamma)} = \frac{2 - (2 - \sqrt{3})}{\left(\sqrt{2} - \frac{\sqrt{6}-\sqrt{2}}{2}\right)\left(2\frac{\sqrt{6}-\sqrt{2}}{2} + \sqrt{2}\right)} \\ &= \frac{\sqrt{3}}{\frac{3\sqrt{2}-\sqrt{6}}{2}\sqrt{6}} = \frac{1}{2} + \frac{\sqrt{3}}{6}, \end{aligned}$$

αποφαινόμαστε ότι αυτή είναι η πρώτη δίτιμη τυχαία μεταβλητή όπως επισημαίνεται στο Πόρισμα 2.3 για $k = 3$. Από την άλλη, κάθε γ που ανήκει στο ανοικτό διάστημα $(\sqrt{2}, \sqrt{3})$, αντιστοιχεί σε μία μοναδική τιμή $\mu_3 = \theta = \gamma(\gamma^2 - 3) \in (\sqrt{2}, 0)$, η οποία με τη σειρά της, προσδιορίζει μονοσήμαντα ότι $\beta = \beta(\theta) \in \left(0, \frac{\sqrt{6}-\sqrt{2}}{2}\right)$, δια μέσου της σχέσης $\beta = \frac{-\gamma + \sqrt{3(4-\gamma^2)}}{2}$ (βλ. την (2.73)), και $\alpha = \alpha(\theta)$, μέσω της $\alpha = \beta + \gamma$. Εν κατακλείδι, αυτές οι μοναδικά ορισμένες τιμές α , β και γ καθορίζουν τις (αυστηρά θετικές) πιθανότητες p_1 , p_2 και p_3 , όπως καταγράφονται στην (2.74), η οποία αποδεικνύει ότι κάθε τυχαία μεταβλητή $X_\theta \in \mathcal{F}_3^0$ καθίσταται μονοσήμαντα προσδιορισμένη από $E(X_\theta^3) = \theta$, $-\sqrt{2} < \theta < 0$.

Υπολείπεται να εξετάσουμε και την περίπτωση κατά την οποία $0 < \theta < 2$. Ως εκ τούτου, αν $X_\theta \in \mathcal{F}_3^0$ και $E(X_\theta^3) = \theta > 0$ τότε είναι εύκολο να επαληθεύσουμε ότι $-X_\theta \in \mathcal{F}_3^0$ και $E(-X_\theta^3) = -E(X_\theta^3) = -\theta < 0$. Χρησιμοποιώντας τα προαναφερθέντα επιχειρήματα προκύπτει, αναγκαστικά, ότι $-\sqrt{2} \leq -\theta < 0$, ότι η $-X_\theta$ καθορίζεται από την τιμή του $-\theta$, ότι η $-X_\theta$ είναι δίτιμη τυχαία μεταβλητή, εάν $-\theta = -\sqrt{2}$, και ότι η $-X_\theta$ είναι τρίτιμη τυχαία μεταβλητή αλλιώς, δηλαδή, όταν $-\theta \in (-\sqrt{2}, 0)$. Με ανάλογο τρόπο, το ίδιο αληθεύει και για την X_θ , και έτσι, η απόδειξη έχει ολοκληρωθεί. ■

Μολονότι, δεν είμαστε σε θέση να περιγράψουμε πλήρως το σύνολο $\mathcal{F}_k^0$ για περιπτώ $k \geq 5$, η κατάσταση φαίνεται να είναι σε μεγάλο βαθμό παρόμοια συγκριτικά με την περίπτωση κατά την οποία $k = 3$, ήτοι, κάθε $X_\theta \in \mathcal{F}_k^0$ χαρακτηρίζεται από την κεντρική ροπή, $\theta = E(X^k) = \mu_k$, και οι δυνατές τιμές του θ σχηματίζουν ένα συμμετρικό διάστημα της μορφής $[-\alpha_k, \alpha_k]$, όπου οι ακραίες τιμές $\theta = \pm \alpha_k$ αντιστοιχούν στις δίτιμες κατανομές του Πορίσματος 2.3, εν αντιθέσει με κάθε $\theta \in (-\alpha_k, \alpha_k)$ που καθορίζει κατά μοναδικό τρόπο μία τρίτιμη τυχαία μεταβλητή.

2.5 Μέθοδος Δέλτα πρώτης τάξης

Τώρα, είμαστε έτοιμοι, πλέον, να διατυπώσουμε τα Θεωρήματα της μεθόδου Δέλτα πρώτης τάξης, αφού, προηγουμένως καλέσουμε έναν χρήσιμο Ορισμό από το πρώτο κεφάλαιο.

Ορισμός: Μία ακολουθία εκτιμητριών $\{T_n\}_{n \geq 1}$ μίας παραμετρικής συνάρτησης $g(\theta)$ λέγεται ασυμπτωτικά κανονική εκτιμήτρια της $g(\theta)$ εάν και μόνο εάν

$$\exists r(\theta) \in [0, +\infty): \sqrt{n}(T_n - g(\theta)) \xrightarrow{d} \mathcal{N}(0, r^2(\theta)), \quad \forall \theta \in \theta, \quad \text{καθώς } n \rightarrow +\infty.$$

Θεώρημα 2.5: (μέθοδος Δέλτα πρώτης τάξης από τον $\mathcal{R}$ στον $\mathcal{R}$)

Έστω $\{T_n\}_{n \geq 1}$ ασυμπτωτικά κανονική εκτιμήτρια μίας παραμέτρου θ , δηλαδή,

$$\exists r(\theta) \in [0, +\infty): \sqrt{n}(T_n - \theta) \xrightarrow{d} \mathcal{N}(0, r^2(\theta)), \quad \forall \theta \in \theta, \quad \text{καθώς } n \rightarrow +\infty.$$

Τότε

$$\sqrt{n}(g(T_n) - g(\theta)) \xrightarrow{d} \mathcal{N}\left(0, (g'(\theta) \cdot r(\theta))^2\right), \quad \text{καθώς } n \rightarrow +\infty,$$

όπου $g: \mathcal{R} \rightarrow \mathcal{R}$ είναι Borel-μετρήσιμη συνάρτηση για την οποία $\exists g'(\theta)$ και $g'(\theta) \neq 0$.

Απόδειξη: Αναπτύσσουμε την ακολουθία T_n σε σειρά *Taylor* πρώτης τάξης γύρω από το σημείο θ και λαμβάνουμε

$$g(T_n) = g(\theta) + \frac{g'(\theta)}{1!} (T_n - \theta)^1 + R_{1,n} = g(\theta) + g'(\theta)(T_n - \theta) + R_{1,n},$$

όπου $R_{1,n} = o_p(T_n - \theta)$, δηλαδή, $\frac{R_{1,n}}{T_n - \theta} \xrightarrow{p} 0$,

έτσι,

$$g(T_n) - g(\theta) = g'(\theta)(T_n - \theta) + R_{1,n}.$$

Τώρα, πολλαπλασιάζοντας και τα δύο μέλη με $\sqrt{n}$ θα προκύψει

$$\sqrt{n}(g(T_n) - g(\theta)) = g'(\theta)\sqrt{n}(T_n - \theta) + \sqrt{n}R_{1,n}. \quad (2.75)$$

Κατόπιν, διαιρούμε και τα δύο μέλη με $\sqrt{n}(T_n - \theta)$ και έχουμε

$$\frac{\sqrt{n}(g(T_n) - g(\theta))}{\sqrt{n}(T_n - \theta)} = \frac{g'(\theta)\sqrt{n}(T_n - \theta)}{\sqrt{n}(T_n - \theta)} + \frac{\sqrt{n}R_{1,n}}{\sqrt{n}(T_n - \theta)}$$

όμως,

$$\frac{\sqrt{n}R_{1,n}}{\sqrt{n}(T_n - \theta)} = \frac{R_{1,n}}{T_n - \theta} \xrightarrow{p} 0 \text{ και } \sqrt{n}(T_n - \theta) = O_p(1) \text{ [αφού, } \sqrt{n}(T_n - \theta) \xrightarrow{d} \mathcal{N}(0, r^2(\theta))]$$

επομένως,

$$\sqrt{n}R_{1,n} \xrightarrow{p} 0 \Rightarrow \sqrt{n}R_{1,n} \xrightarrow{d} 0. \quad (2.76)$$

Τελικά, σύμφωνα με τις σχέσεις (2.75), (2.76) και το Θεώρημα *Slutsky* απορρέει ότι

$$\begin{aligned} \sqrt{n}(g(T_n) - g(\theta)) &= g'(\theta)\sqrt{n}(T_n - \theta) + \sqrt{n}R_{1,n} \xrightarrow{d} g'(\theta) \cdot \mathcal{N}(0, r^2(\theta)) + 0 \Rightarrow \\ &\Rightarrow \sqrt{n}(g(T_n) - g(\theta)) \xrightarrow{d} \mathcal{N}\left(0, (g'(\theta) \cdot r(\theta))^2\right), \text{ καθώς } n \rightarrow +\infty. \blacksquare \end{aligned}$$

Θεώρημα 2.6: (πολυδιάστατη μέθοδος Δέλτα πρώτης τάξης από τον $\mathcal{R}^k$ στον $\mathcal{R}^s$)

Θεωρούμε ότι $\bar{X}_n$ είναι μια ακολουθία k –διάστατων τυχαίων διανυσμάτων τέτοια ώστε

$$\sqrt{n}(\bar{X}_n - \mu) \xrightarrow{d} \begin{pmatrix} W_1 \\ \vdots \\ W_k \end{pmatrix} = \mathbf{W}, \text{ καθώς } n \rightarrow +\infty,$$

όπου $\mu \in \mathcal{R}^k$ και $\mathbf{W}$ είναι k –διάστατο τυχαίο διάνυσμα. Έστω επίσης, το διάνυσμα των *Borel*–μετρησίμων συναρτήσεων $\mathbf{g} = (g_1, g_2, \dots, g_s)'$: $\mathcal{R}^k \rightarrow \mathcal{R}^s$ για το οποίο υποθέτουμε ότι οι συναρτήσεις $g_1, g_2, \dots, g_s$ έχουν συνεχείς

μερικές παραγώγους πρώτης τάξης γύρω από τον k –διάστατο μέσο μ , που ορίζονται ως εξής:

$$g_i^{(j)}(\mathbf{x}) = \frac{\partial g_i(\mathbf{x})}{\partial x_j}, \quad j = 1, 2, \dots, k, \quad i = 1, 2, \dots, s.$$

Τότε

$$\sqrt{n}(\mathbf{g}(\bar{\mathbf{X}}_n) - \mathbf{g}(\mu)) \xrightarrow{d} H(\mu) \cdot \mathbf{W}, \quad \text{καθώς } n \rightarrow +\infty,$$

όπου

$$H(\mu) = (g_i^{(j)}(\mu)), \quad j = 1, 2, \dots, k, \quad i = 1, 2, \dots, s,$$

δηλαδή,

$$\begin{aligned} \sqrt{n} \begin{pmatrix} g_1(\bar{\mathbf{X}}_n) - g_1(\mu) \\ g_2(\bar{\mathbf{X}}_n) - g_2(\mu) \\ \vdots \\ g_s(\bar{\mathbf{X}}_n) - g_s(\mu) \end{pmatrix} &\xrightarrow{d} \begin{pmatrix} g_1^{(1)}(\mu) & \dots & \dots & g_1^{(k)}(\mu) \\ g_2^{(1)}(\mu) & \dots & \dots & g_2^{(k)}(\mu) \\ \vdots & \vdots & \vdots & \vdots \\ g_s^{(1)}(\mu) & \dots & \dots & g_s^{(k)}(\mu) \end{pmatrix} \begin{pmatrix} W_1 \\ \vdots \\ W_k \end{pmatrix} \\ &= \begin{pmatrix} g_1^{(1)}(\mu)W_1 + \dots + g_1^{(k)}(\mu)W_k \\ g_2^{(1)}(\mu)W_1 + \dots + g_2^{(k)}(\mu)W_k \\ \vdots \\ g_s^{(1)}(\mu)W_1 + \dots + g_s^{(k)}(\mu)W_k \end{pmatrix} = \begin{pmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_s \end{pmatrix} = \mathbf{Z}. \end{aligned}$$

Απόδειξη: Επικαλούμενοι το τέχνασμα³⁶ των *Cramér–Wald*, αρκεί να δείξουμε ότι ο γραμμικός συνδυασμός

$$\begin{aligned} c_1 [\sqrt{n}(g_1(\bar{\mathbf{X}}_n) - g_1(\mu))] + c_2 [\sqrt{n}(g_2(\bar{\mathbf{X}}_n) - g_2(\mu))] + \dots \\ + c_s [\sqrt{n}(g_s(\bar{\mathbf{X}}_n) - g_s(\mu))] \end{aligned}$$

συγκλίνει ασθενώς προς τον γραμμικό συνδυασμό

³⁶Σύμφωνα με αυτό, αν $\mathbf{X}, \{\mathbf{X}_n\}_{n \geq 1}$ είναι ακολουθίες τυχαίων διανυσμάτων στον $\mathcal{R}^k$ τότε, καθώς $n \rightarrow \infty, \mathbf{X}_n \xrightarrow{d} \mathbf{X} \Leftrightarrow \mathbf{c}^T \mathbf{X}_n \xrightarrow{d} \mathbf{c}^T \mathbf{X}, \quad \forall \mathbf{c} \in \mathcal{R}^k$.

$$c_1 Z_1 + c_2 Z_2 + \cdots + c_s Z_s \text{ για κάθε } \mathbf{c} \in \mathcal{R}^s.$$

Θεωρούμε λοιπόν, την συνάρτηση

$$g(\mathbf{x}) = c_1 g_1(\mathbf{x}) + c_2 g_2(\mathbf{x}) + \cdots + c_s g_s(\mathbf{x}).$$

Αυτή έχει συνεχείς μερικές παραγώγους πρώτης τάξης ως προς $x_1, x_2, \dots, x_k$ που ορίζονται ως εξής

$$g^{(j)}(\mathbf{x}) = \frac{\partial g(\mathbf{x})}{\partial x_j} = \sum_{i=1}^s c_i g_i^{(j)}(\mathbf{x}), \quad j = 1, 2, \dots, k.$$

Από το Θεώρημα της πολυδιάστατης μεθόδου Δέλτα m τάξης για $m = 1$ (βλ. παράγραφο 3.3 του επόμενου κεφαλαίου) προκύπτει ότι

$$\begin{aligned} \sqrt{n} \left(g(\bar{\mathbf{X}}_n) - g(\boldsymbol{\mu}) \right) &\xrightarrow{d} [D^1 g(\boldsymbol{\mu})] \cdot \mathbf{W}^1 = [\nabla g(\boldsymbol{\mu})]' \cdot \mathbf{W} = \sum_{j=1}^k g^{(j)}(\boldsymbol{\mu}) W_j \\ &= \sum_{j=1}^k W_j \cdot \sum_{i=1}^s c_i g_i^{(j)}(\boldsymbol{\mu}) = \sum_{i=1}^s c_i \cdot \sum_{j=1}^k g_i^{(j)}(\boldsymbol{\mu}) W_j = \sum_{i=1}^s c_i Z_i. \end{aligned}$$

Όμως,

$$\begin{aligned} \sqrt{n} \left(g(\bar{\mathbf{X}}_n) - g(\boldsymbol{\mu}) \right) &= \sqrt{n} \left[\sum_{i=1}^s c_i g_i(\bar{\mathbf{X}}_n) - \sum_{i=1}^s c_i g_i(\boldsymbol{\mu}) \right] \\ &= \sum_{i=1}^s c_i \left[\sqrt{n} \left(g_i(\bar{\mathbf{X}}_n) - g_i(\boldsymbol{\mu}) \right) \right] \xrightarrow{d} \sum_{i=1}^s c_i Z_i, \quad \forall \mathbf{c} \in \mathcal{R}^s, \end{aligned}$$

Επομένως,

$$\sqrt{n} \left(g(\bar{\mathbf{X}}_n) - g(\boldsymbol{\mu}) \right) \xrightarrow{d} \mathbf{Z}. \quad \blacksquare$$

Κεφάλαιο 3

Μέθοδος Δέλτα ανώτερης τάξης

3.1 Μέθοδος Δέλτα δεύτερης τάξης

Εξίσου σημαντική είναι και η μέθοδος Δέλτα πέραν της πρώτης τάξεως η οποία, αν και ολοένα και δυσκολότερα εφαρμόζεται όσο μεγαλώνει η τάξη της μεθόδου, ωστόσο, αρκετές φορές δύναται να συναντήσουμε στην ασυμπτωτική στατιστική.

Θεώρημα 3.1: (μέθοδος Δέλτα δεύτερης τάξης από τον $\mathcal{R}$ στον $\mathcal{R}$)

Έστω $\{T_n\}_{n \geq 1}$ ασυμπτωτικά κανονική εκτιμήτρια μίας παραμέτρου θ , δηλαδή,

$$\exists r(\theta) \in [0, +\infty): \sqrt{n}(T_n - \theta) \xrightarrow{d} \mathcal{N}(0, r^2(\theta)), \quad \forall \theta \in \Theta, \text{ καθώς } n \rightarrow +\infty.$$

Τότε,

$$n(g(T_n) - g(\theta)) \xrightarrow{d} \frac{r^2(\theta)g''(\theta)}{2} \chi_1^2, \text{ καθώς } n \rightarrow +\infty,$$

όπου $g: \mathcal{R} \rightarrow \mathcal{R}$ είναι *Borel*-μετρήσιμη συνάρτηση για την οποία $\exists g'(\theta)$, $g''(\theta)$ έτσι ώστε $g'(\theta) = 0$ και $g''(\theta) \neq 0$, και όπου χ_1^2 είναι η κατανομή *χι-τετράγωνο* με ένα βαθμό ελευθερίας.

Απόδειξη: Αναπτύσσουμε την ακολουθία $g(T_n)$ σε σειρά *Taylor* δεύτερης τάξης περί το $g(\theta)$ και έτσι προκύπτει

$$\begin{aligned} g(T_n) &= g(\theta) + \frac{g'(\theta)}{1!}(T_n - \theta)^1 + \frac{g''(\theta)}{2!}(T_n - \theta)^2 + R_{2,n} \\ &= g(\theta) + \frac{g''(\theta)}{2}(T_n - \theta)^2 + R_{2,n}, \end{aligned}$$

όπου $R_{2,n} = o_p((T_n - \theta)^2)$, ήτοι, $\frac{R_{2,n}}{(T_n - \theta)^2} \xrightarrow{p} 0$,

έτσι,

$$g(T_n) - g(\theta) = \frac{g''(\theta)}{2}(T_n - \theta)^2 + R_{2,n}.$$

Τώρα, πολλαπλασιάζοντας και τα δύο μέλη με n θα προκύψει

$$n(g(T_n) - g(\theta)) = \frac{g''(\theta)}{2}n(T_n - \theta)^2 + nR_{2,n}. \quad (3.1)$$

Κατόπιν, διαιρούμε και τα δύο μέλη με $n(T_n - \theta)^2$ και έχουμε

$$\frac{n(g(T_n) - g(\theta))}{n(T_n - \theta)^2} = \frac{\frac{g''(\theta)}{2}n(T_n - \theta)^2}{n(T_n - \theta)^2} + \frac{nR_{2,n}}{n(T_n - \theta)^2},$$

όμως,

$$\frac{nR_{2,n}}{n(T_n - \theta)^2} = \frac{R_{2,n}}{(T_n - \theta)^2} \xrightarrow{p} 0$$

και

$$n(T_n - \theta)^2 = O_p(1),$$

επομένως,

$$nR_{2,n} \xrightarrow{p} 0 \Rightarrow nR_{2,n} \xrightarrow{d} 0. \quad (3.2)$$

Επίσης,

$$\begin{aligned} \sqrt{n}(T_n - \theta) &\xrightarrow{d} \mathcal{N}(0, r^2(\theta)) \Rightarrow \frac{\sqrt{n}(T_n - \theta)}{r(\theta)} \xrightarrow{d} \mathcal{N}(0, 1) \\ &\Rightarrow \frac{n(T_n - \theta)^2}{r^2(\theta)} \xrightarrow{d} \mathcal{N}^2(0, 1) \equiv \chi_1^2. \end{aligned} \quad (3.3)$$

Τελικά, σύμφωνα με τις σχέσεις (3.1), (3.2), (3.3) και το Θεώρημα *Slutsky* απορρέει ότι

$$\begin{aligned} n(g(T_n) - g(\theta)) &= \frac{g''(\theta)}{2}n(T_n - \theta)^2 + nR_{2,n} \xrightarrow{d} \frac{g''(\theta)}{2}r^2(\theta)\chi_1^2 + 0 \Rightarrow \\ &\Rightarrow n(g(T_n) - g(\theta)) \xrightarrow{d} \frac{g''(\theta)}{2}r^2(\theta)\chi_1^2, \text{ καθώς } n \rightarrow +\infty. \blacksquare \end{aligned}$$

Θεώρημα 3.2: (μέθοδος Δέλτα δεύτερης τάξης από τον $\mathcal{R}^k$ στον $\mathcal{R}$)

Υποθέτουμε ότι $\bar{X}_n$ είναι μια ακολουθία k -διάστατων τυχαίων διανυσμάτων τέτοια ώστε

$$\sqrt{n}(\bar{X}_n - \mu) \xrightarrow{d} W, \text{ καθώς } n \rightarrow +\infty, \quad (3.4)$$

όπου $\mu \in \mathcal{R}^k$ και W είναι k -διάστατο τυχαίο διάνυσμα. Θεωρούμε ότι η *Borel*-μετρήσιμη συνάρτηση $g: \mathcal{R}^k \rightarrow \mathcal{R}$ είναι δύο φορές συνεχώς διαφορίσιμη γύρω από τον k -διάστατο μέσο μ και ορίζουμε

$$\nabla g(\mu) = \left(\frac{\partial g(x)}{\partial x_i} \right) \Big|_{x=\mu} \in \mathcal{R}^k \text{ και } H_k(\mu) := \left(\frac{\partial^2 g(x)}{\partial x_i \partial x_j} \right) \Big|_{x=\mu} \in \mathcal{R}^{k \times k}.$$

Αν

$$n[\nabla g(\mu)]'(\bar{X}_n - \mu) \xrightarrow{p} 0 \quad (3.5)$$

τότε

$$n(g(\bar{X}_n) - g(\mu)) \xrightarrow{d} \frac{1}{2} W^T H_k(\mu) W, \text{ καθώς } n \rightarrow +\infty. \quad (3.6)$$

Απόδειξη: Από την (3.4) έπεται ότι $\bar{X}_n \xrightarrow{p} \mu$. Εν συνεχεία, αναπτύσσοντας την ακολουθία $g(\bar{X}_n)$ σε σειρά *Taylor* 2^{ης} τάξης γύρω από το σημείο $g(\mu)$ θα λάβουμε

$$g(\bar{X}_n) = g(\mu) + \frac{[\nabla g(\mu)]^T}{1!} (\bar{X}_n - \mu)^1 + \frac{(\bar{X}_n - \mu)^T H_k(\mu) (\bar{X}_n - \mu)}{2!} + R_{2,n},$$

όπου $R_{2,n} = o_p \left((\bar{X}_n - \mu)^T (\bar{X}_n - \mu) \right)$, δηλαδή, $\frac{R_{2,n}}{(\bar{X}_n - \mu)^T (\bar{X}_n - \mu)} \xrightarrow{p} 0$,

έτσι,

$$g(\bar{X}_n) - g(\mu) = [\nabla g(\mu)]^T (\bar{X}_n - \mu) + \frac{1}{2} (\bar{X}_n - \mu)^T H_k(\mu) (\bar{X}_n - \mu) + R_{2,n}.$$

Τώρα, πολλαπλασιάζοντας και τα δύο μέλη με n θα προκύψει

$$\begin{aligned} n(g(\bar{X}_n) - g(\mu)) &= n[\nabla g(\mu)]^T (\bar{X}_n - \mu) \\ &\quad + \frac{1}{2} n(\bar{X}_n - \mu)^T H_k(\mu) (\bar{X}_n - \mu) + nR_{2,n}. \end{aligned}$$

Από υπόθεση ισχύει

$$n[\nabla g(\boldsymbol{\mu})]^T (\bar{\mathbf{X}}_n - \boldsymbol{\mu}) \xrightarrow{p} 0,$$

επομένως,

$$\begin{aligned} n(g(\bar{\mathbf{X}}_n) - g(\boldsymbol{\mu})) &= \frac{1}{2} n(\bar{\mathbf{X}}_n - \boldsymbol{\mu})^T H_k(\boldsymbol{\mu}) (\bar{\mathbf{X}}_n - \boldsymbol{\mu}) + nR_{2,n} \\ &= \frac{1}{2} [\sqrt{n}(\bar{\mathbf{X}}_n - \boldsymbol{\mu})]^T H_k(\boldsymbol{\mu}) [\sqrt{n}(\bar{\mathbf{X}}_n - \boldsymbol{\mu})] + nR_{2,n}. \end{aligned}$$

Έπειτα, διαιρούμε και τα δύο μέλη με $n(\bar{\mathbf{X}}_n - \boldsymbol{\mu})^T (\bar{\mathbf{X}}_n - \boldsymbol{\mu})$ και έχουμε

$$\begin{aligned} \frac{n(g(\bar{\mathbf{X}}_n) - g(\boldsymbol{\mu}))}{n(\bar{\mathbf{X}}_n - \boldsymbol{\mu})^T (\bar{\mathbf{X}}_n - \boldsymbol{\mu})} &= \frac{\frac{1}{2} [\sqrt{n}(\bar{\mathbf{X}}_n - \boldsymbol{\mu})]^T H_k(\boldsymbol{\mu}) [\sqrt{n}(\bar{\mathbf{X}}_n - \boldsymbol{\mu})]}{n(\bar{\mathbf{X}}_n - \boldsymbol{\mu})^T (\bar{\mathbf{X}}_n - \boldsymbol{\mu})} \\ &\quad + \frac{nR_{2,n}}{n(\bar{\mathbf{X}}_n - \boldsymbol{\mu})^T (\bar{\mathbf{X}}_n - \boldsymbol{\mu})}, \end{aligned}$$

όμως,

$$\frac{nR_{2,n}}{n(\bar{\mathbf{X}}_n - \boldsymbol{\mu})^T (\bar{\mathbf{X}}_n - \boldsymbol{\mu})} = \frac{R_{2,n}}{(\bar{\mathbf{X}}_n - \boldsymbol{\mu})^T (\bar{\mathbf{X}}_n - \boldsymbol{\mu})} = O_p(1),$$

άρα,

$$nR_{2,n} \xrightarrow{p} 0 \Rightarrow nR_{2,n} \xrightarrow{d} 0.$$

Εν τέλει, σύμφωνα με τη σχέση (3.4) και το Θεώρημα *Slutsky* απορρέει ότι

$$\begin{aligned} n(g(\bar{\mathbf{X}}_n) - g(\boldsymbol{\mu})) &= \frac{1}{2} [\sqrt{n}(\bar{\mathbf{X}}_n - \boldsymbol{\mu})]^T H_k(\boldsymbol{\mu}) [\sqrt{n}(\bar{\mathbf{X}}_n - \boldsymbol{\mu})] \\ &\quad + nR_{2,n} \xrightarrow{d} \frac{1}{2} \mathbf{W}^T H_k(\boldsymbol{\mu}) \mathbf{W} + 0 \Rightarrow \\ &\Rightarrow n(g(\bar{\mathbf{X}}_n) - g(\boldsymbol{\mu})) \xrightarrow{d} \frac{1}{2} \mathbf{W}^T H_k(\boldsymbol{\mu}) \mathbf{W}, \text{ καθώς } n \rightarrow +\infty. \blacksquare \end{aligned}$$

3.2 Η οριακή συμπεριφορά της ιδιάζουσας κατανομής

Εάν το τυχαίο δείγμα προέρχεται από μια ιδιάζουσα κατανομή τάξης $k \geq 2$, τότε η ασυμπτωτική κανονικότητα της σχέσης (2.7) ανάγεται στο όριο $\sqrt{n}(M_{k,n} - \mu_k) \xrightarrow{p} 0$ (βλ. Ορισμό 2.2). Αυτό καταδεικνύει ότι η τάξη σύγκλισης της $M_{k,n}$ προς τη μ_k είναι ταχύτερη από $o(1/\sqrt{n})$, και έτσι, μια δεύτερης τάξης προσέγγιση μπορεί να εφαρμοστεί, σύμφωνα με τα ανωτέρω Θεωρήματα.

Το τελευταίο Θεώρημα εφαρμόζεται, πάραυτα, για τη δειγματική κεντρική ροπή k -τάξης, $M_{k,n}$, κάθε φορά που το τυχαίο δείγμα προέρχεται από μια ιδιάζουσα κατανομή.

Θεώρημα 3.3: Έστω ότι $M_{k,n}$ είναι η δειγματική κεντρική ροπή μίας ιδιάζουσας κατανομής τάξης $k \geq 2$ τότε, καθώς $n \rightarrow +\infty$,

$$n(M_{k,n} - \mu_k) \xrightarrow{d} \frac{1}{2}k(k-1)\mu_{k-2}W_1^2 - kW_1W_{k-1}, \quad (3.7)$$

όπου

$$\begin{pmatrix} W_1 \\ W_{k-1} \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2 & \mu_k \\ \mu_k & \mu_{2k-2} - \mu_{k-1}^2 \end{pmatrix}\right). \quad (3.8)$$

Απόδειξη: Καταρχάς, υπενθυμίζουμε ότι $M_{k,n} - \mu_k = g(\mathbf{m}_n) - g(\boldsymbol{\mu})$, όπου $g: \mathcal{R}^k \rightarrow \mathcal{R}$ είναι *Borel*-μετρήσιμη συνάρτηση όπως αυτή δίνεται στη σχέση (2.3), δηλαδή,

$$g(x_1, x_2, \dots, x_k) = (-1)^{k-1}(k-1)x_1^k + \sum_{j=2}^{k-1} \binom{k}{j} (-1)^{k-j} x_1^{k-j} x_j + x_k$$

και $\boldsymbol{\mu} = (0, \sigma^2, \mu_3, \dots, \mu_k)'$. Επίσης,

$$\sqrt{n}(\mathbf{m}_n - \boldsymbol{\mu}) \xrightarrow{d} \mathbf{W} = (W_1, W_2, \dots, W_k)' \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma})$$

με $\boldsymbol{\Sigma} = (\mu_{i+j} - \mu_i\mu_j)$ (βλ. την (2.2)). Επομένως, το Θεώρημα 3.2 ισχύει για $\bar{\mathbf{X}}_n = \mathbf{m}_n$, και έτσι αποδεικνύεται ότι η σχέση (3.5) πληρούται για $\mathbf{m}_n$, δηλαδή, καθίσταται σαφές ότι

$$n[\nabla g(\boldsymbol{\mu})]^T(\mathbf{m}_n - \boldsymbol{\mu}) \xrightarrow{p} 0.$$

Χρησιμοποιώντας την (2.6) παρατηρούμε ότι

$$[\nabla g(\boldsymbol{\mu})]^T(\mathbf{m}_n - \boldsymbol{\mu}) = -k\mu_{k-1}m_{1,n} + (m_{k,n} - \mu_k).$$

Τώρα, εφόσον $E(m_{j,n}) = \mu_j$ για όλα τα n και j θα έχουμε

$$E[-k\mu_{k-1}m_{1,n} + (m_{k,n} - \mu_k)] = 0$$

και επιπλέον, για την διασπορά προκύπτει διαδοχικά ότι

$$\text{Var}[-k\mu_{k-1}m_{1,n} + (m_{k,n} - \mu_k)]$$

$$\begin{aligned}
&= (-k\mu_{k-1})^2 \text{Var}(m_{1,n}) + \text{Var}(m_{k,n} - \mu_k) + 2\text{Cov}[-k\mu_{k-1}m_{1,n}, (m_{k,n} - \mu_k)] \\
&= k^2\mu_{k-1}^2 \text{Var}(m_{1,n}) + \text{Var}(m_{k,n}) - 2k\mu_{k-1}\text{Cov}(m_{1,n}, m_{k,n}). \tag{3.9}
\end{aligned}$$

Όμως,

$$m_{j,n} := \frac{1}{n} \sum_{i=1}^n (X_i - \mu)^j, \quad \forall j = 1, 2, \dots, k$$

έτσι,

$$m_{1,n} := \frac{1}{n} \sum_{i=1}^n (X_i - \mu)^1 = \frac{1}{n} \sum_{i=1}^n X_i - \frac{1}{n} n\mu = \bar{X}_n - \mu,$$

ήτοι,

$$\text{Var}(m_{1,n}) = \frac{\sigma^2}{n}.$$

Επιπροσθέτως,

$$\begin{aligned}
\text{Var}(m_{k,n}) &= \text{Var}\left(\frac{1}{n} \sum_{i=1}^n (X_i - \mu)^k\right) = \frac{1}{n^2} \sum_{i=1}^n \text{Var}(X_i - \mu)^k = \frac{1}{n^2} n \text{Var}(X_i - \mu)^k \\
&= \frac{1}{n} \{E(X_i - \mu)^{2k} - [E(X_i - \mu)^k]^2\} = \frac{1}{n} (\mu_{2k} - \mu_k^2)
\end{aligned}$$

και

$$\text{Cov}(m_{1,n}, m_{k,n}) = E(m_{1,n}m_{k,n}) - E(m_{1,n})E(m_{k,n}).$$

Ωστόσο,

$$E(m_{1,n}) = \mu_1 = 0,$$

οπότε,

$$\begin{aligned}
\text{Cov}(m_{1,n}, m_{k,n}) &= E(m_{1,n}m_{k,n}) = E\left(\frac{1}{n} \sum_{i=1}^n (X_i - \mu) \cdot \frac{1}{n} \sum_{i=1}^n (X_i - \mu)^k\right) \\
&= E\left(\frac{1}{n^2} \sum_{i=1}^n (X_i - \mu)^{k+1}\right) = \frac{1}{n^2} \sum_{i=1}^n E(X_i - \mu)^{k+1} = \frac{1}{n^2} n E(X_i - \mu)^{k+1} = \frac{\mu_{k+1}}{n}.
\end{aligned}$$

Τελικά, η εξίσωση (3.9) θα λάβει τη μορφή

$$\text{Var}[-k\mu_{k-1}m_{1,n} + (m_{k,n} - \mu_k)]$$

$$\begin{aligned}
&= k^2 \mu_{k-1}^2 \frac{\sigma^2}{n} + \frac{\mu_{2k} - \mu_k^2}{n} - 2k \mu_{k-1} \frac{\mu_{k+1}}{n} \\
&= \frac{1}{n} (k^2 \mu_{k-1}^2 \sigma^2 + \mu_{2k} - \mu_k^2 - 2k \mu_{k-1} \mu_{k+1}) \\
&= \frac{1}{n} [\mu_{2k} - \mu_k^2 + k \mu_{k-1} (k \sigma^2 \mu_{k-1} - 2 \mu_{k+1})] = \frac{v_k^2}{n} = 0,
\end{aligned}$$

διότι, υποθέσαμε ότι η κατανομή είναι ιδιάζουσα πράγμα που σημαίνει ότι $v_k^2 = 0$. Ως εκ τούτου, $[\nabla g(\boldsymbol{\mu})]^T (\mathbf{m}_n - \boldsymbol{\mu}) = 0$ με πιθανότητα ένα και, έτσι, η σχέση $n[\nabla g(\boldsymbol{\mu})]^T (\mathbf{m}_n - \boldsymbol{\mu}) \xrightarrow{p} 0$ είναι τετριμμένη. Κατόπιν, με έναν απλό υπολογισμό, χρησιμοποιώντας την (2.5), αποδεικνύεται ότι

$$H_k(\boldsymbol{\mu}) = \begin{pmatrix} k(k-1)\mu_{k-2} & 0 & \cdots & \cdots & \cdots & 0 & -k & 0 \\ 0 & 0 & \cdots & \cdots & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \cdots & \cdots & 0 & 0 & 0 \\ -k & 0 & \cdots & \cdots & \cdots & 0 & 0 & 0 \\ 0 & 0 & \cdots & \cdots & \cdots & 0 & 0 & 0 \end{pmatrix},$$

για παράδειγμα,

$$H_2(\boldsymbol{\mu}) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}, H_3(\boldsymbol{\mu}) = \begin{pmatrix} 0 & -3 & 0 \\ -3 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, H_4(\boldsymbol{\mu}) = \begin{pmatrix} 12\sigma^2 & 0 & -4 & 0 \\ 0 & 0 & 0 & 0 \\ -4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \text{ κ.ο.κ.}$$

Εφαρμόζοντας την (3.6) βλέπουμε ότι η ποσότητα $n(M_{k,n} - \mu_k)$ συγκλίνει ασθενώς στην κατανομή $\frac{1}{2} \mathbf{W}^T H_k(\boldsymbol{\mu}) \mathbf{W} = \frac{1}{2} k(k-1) \mu_{k-2} W_1^2 - k W_1 W_{k-1}$, ενώ, από τη σχέση (2.2), η κατανομή του $(W_1, W_{k-1})'$ δίνεται από την (3.8). ■

Παρατήρηση: Η οριακή κατανομή της (3.7) δύναται να εκφρασθεί σε όρους δύο ανεξάρτητων και ισόνομων τυπικών κανονικών τυχαίων μεταβλητών Z_1, Z_2 .

Απόδειξη: Πράγματι, παρατηρώντας ότι

$$\sigma^2(\mu_{2k-2} - \mu_{k-1}^2) - \mu_k^2 = \text{Var} \left[\sigma(X - \mu)^{k-1} - \mu_k \frac{X - \mu}{\sigma} \right] \geq 0$$

είναι εύκολο να δούμε ότι

$$\begin{pmatrix} W_1 \\ W_{k-1} \end{pmatrix} = \begin{pmatrix} \sigma Z_1 \\ \frac{\mu_k}{\sigma} Z_1 + \frac{\gamma_k}{\sigma} Z_2 \end{pmatrix}, \quad \text{όπου } \gamma_k := \sqrt{\sigma^2(\mu_{2k-2} - \mu_{k-1}^2)}. \quad (3.10)$$

Συνεπώς, η (3.7) μπορεί να ξαναγραφτεί ως

$$n(M_{k,n} - \mu_k) \xrightarrow{d} \left(\frac{1}{2} k(k-1) \sigma^2 \mu_{k-2} - k \mu_k \right) Z_1^2 - k \gamma_k Z_1 Z_2. \quad \blacksquare \quad (3.11)$$

Σε περίπτωση που επιθυμούμε μία πιο απλούστερη μορφή για την τελευταία σχέση σύγκλισης μπορούμε να χρησιμοποιήσουμε την επόμενη πρόταση.

Πρόταση 3.1: Ας είναι Z_1, Z_2 ανεξάρτητες και ισόνομες τυπικές κανονικές τυχαίες μεταβλητές τότε για αυθαίρετες πραγματικές σταθερές α, β ισχύει,

$$\alpha Z_1^2 + \beta Z_1 Z_2 =_d \frac{1}{2} (\sqrt{\alpha^2 + \beta^2} + \alpha) Z_1^2 - \frac{1}{2} (\sqrt{\alpha^2 + \beta^2} - \alpha) Z_2^2. \quad (3.12)$$

Απόδειξη: Ο ισχυρισμός είναι προφανής εάν $\beta = 0$. Τώρα, θεωρούμε ότι $\beta \neq 0$ και θέτουμε $\rho = \sqrt{\alpha^2 + \beta^2} > 0$. Είναι εύκολο να αντιληφθούμε ότι η ροπογεννήτρια συνάρτηση του δεξιού μέλους της τελευταίας σχέσης δίνεται από την εξίσωση

$$M_2(t) = \frac{1}{\sqrt{1 - 2at - \beta^2 t^2}}$$

και αυτή είναι πεπερασμένη μέσα στο διάστημα $\{t \in \mathcal{R}: 1 - 2at - \beta^2 t^2 > 0\} = \left(-\frac{1}{\rho - \alpha}, \frac{1}{\rho + \alpha}\right)$ το οποίο περιέχει την αρχή επειδή $\rho - \alpha > 0$ και $\rho + \alpha > 0$. Επίσης, η ροπογεννήτρια συνάρτηση του αριστερού μέλους της (3.12) είναι

$$M_1(t) = E(e^{atZ_1^2 + btZ_1Z_2}) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}\gamma(x,y)} dx dy,$$

όπου

$$\gamma(x, y) = x^2 + y^2 - 2atx^2 - 2\beta txy = (1 - 2at - \beta^2 t^2)x^2 + (y - \beta tx)^2.$$

Ως εκ τούτου, για κάθε $t \in \left(-\frac{1}{\rho-\alpha}, \frac{1}{\rho+\alpha}\right)$ έχουμε

$$\begin{aligned} M_1(t) &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}[(1-2at-\beta^2 t^2)x^2 + (y-\beta tx)^2]} dx dy \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}(1-2at-\beta^2 t^2)} \left(\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}(y-\beta tx)^2} dy \right) dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}(1-2at-\beta^2 t^2)} dx = M_1(t), \end{aligned}$$

και η απόδειξη είναι πλήρης. ■

Πόρισμα 3.1: Έστω ότι το διάνυσμα $(X_1, X_2)'$ ακολουθεί μία διδιάστατη κανονική κατανομή με $E(X_1) = \mu_1$, $E(X_2) = \mu_2$, $Var(X_1) = \sigma_1^2$, $Var(X_2) = \sigma_2^2$ και $Cov(X_1, X_2) = \rho\sigma_1\sigma_2$, όπου $\mu_1, \mu_2 \in \mathcal{R}$, $\sigma_1 \geq 0$, $\sigma_2 \geq 0$ και $-1 \leq \rho \leq 1$ είναι αυθαίρετες σταθερές. Τότε

$$(X_1 - \mu_1)(X_2 - \mu_2) =_d \sigma_1\sigma_2 \left\{ \frac{1}{2}(1+\rho)Z_1^2 - \frac{1}{2}(1-\rho)Z_2^2 \right\}.$$

Απόδειξη: Εφόσον,

$$\begin{pmatrix} X_1 - \mu_1 \\ X_2 - \mu_2 \end{pmatrix} =_d \begin{pmatrix} \sigma_1 Z_1 \\ \sigma_2 (\rho Z_1 + \sqrt{1-\rho^2} Z_2) \end{pmatrix},$$

θα ισχύει

$$(X_1 - \mu_1)(X_2 - \mu_2) =_d \sigma_1\sigma_2 \{ \rho Z_1^2 + \sqrt{1-\rho^2} Z_1 Z_2 \},$$

και εν κατακλείδι, ο ισχυρισμός έπεται από την (3.12) για $\alpha = \rho$ και $\beta = \sqrt{1-\rho^2}$. ■

Το κυριότερο αποτέλεσμα περιέχεται στο ακόλουθο Θεώρημα, η απόδειξη του οποίου, είναι άμεση συνέπεια της (3.11) και της Πρότασης 3.1.

Θεώρημα 3.4: Αν $M_{k,n}$ είναι η δειγματική κεντρική ροπή μίας ιδιάζουσας κατανομής τάξης $k \geq 2$ τότε, καθώς $n \rightarrow +\infty$,

$$n(\mu_k - M_{k,n}) \xrightarrow{d} \frac{k}{2} \{ \sigma \sqrt{\theta_k} + \alpha_k \} Z_1^2 - \frac{k}{2} \{ \sigma \sqrt{\theta_k} - \alpha_k \} Z_2^2, \quad (3.13)$$

όπου Z_1, Z_2 είναι ανεξάρτητες και ισόνομες τυπικές κανονικές τυχαίες μεταβλητές και

$$\alpha_k = \mu_k - \frac{1}{2}(k-1)\sigma^2\mu_{k-2}, \quad (3.14)$$

$$\theta_k = \mu_{2k-2} - \mu_{k-1}^2 - (k-1)\mu_{k-2} \left[\mu_k - \frac{1}{4}(k-1)\sigma^2\mu_{k-2} \right]. \quad (3.15)$$

Πόρισμα 3.2: Εάν $M_{k,n}$ είναι η δειγματική κεντρική ροπή μίας ιδιάζουσας κατανομής τάξης $k \geq 2$ τότε υπάρχει μία σταθερά $\lambda_k \in \mathcal{R}$ τέτοια ώστε

$$n(\mu_k - M_{k,n}) \xrightarrow{d} \lambda_k \chi_1^2$$

όταν και μόνο όταν

$$\mu_k^2 = \sigma^2(\mu_{2k-2} - \mu_{k-1}^2). \quad (3.16)$$

Αν η προηγούμενη εξίσωση ισχύει τότε $\lambda_k = k \left[\mu_k - \frac{1}{2}(k-1)\sigma^2\mu_{k-2} \right]$

και έτσι,

$$n(\mu_k - M_{k,n}) \xrightarrow{d} k \left(\mu_k - \frac{1}{2}(k-1)\sigma^2\mu_{k-2} \right) \chi_1^2. \quad (3.17)$$

Αν όμως, η σχέση (3.16) δεν υφίσταται τότε

$$n(\mu_k - M_{k,n}) \xrightarrow{d} \lambda_k \chi_1^2 - \tilde{\lambda}_k \tilde{\chi}_1^2 \quad (3.18)$$

με $\lambda_k = \{ \sigma \sqrt{\theta_k} + \alpha_k \} > 0$, $\tilde{\lambda}_k = \{ \sigma \sqrt{\theta_k} - \alpha_k \} > 0$ και α_k, θ_k όπως δίνονται από τις σχέσεις (3.14) και (3.15) αντίστοιχα, και όπου χ_1^2 και $\tilde{\chi}_1^2$ είναι ανεξάρτητες και ισόνομες τυχαίες μεταβλητές από την κατανομή χ -τετράγωνο με ένα βαθμό ελευθερίας.

Ύστερα από πολλούς υπολογισμούς καταλήγουμε ότι η (3.16) ικανοποιείται για όλες τις δίτιμες κατανομές που επισημαίνονται στα Πορίσματα 2.2 και 2.3. Συγκεκριμένα, από την (3.17) είμαστε σε θέση να δείξουμε ότι

$$n(\mu_k - M_{k,n}) \xrightarrow{d} \frac{k(k-1)}{2} p_k^{k-1} \frac{(k+1)p_k^2 - (k+1)p_k + 1}{(k+1)p_k - 1} \chi_1^2, \quad k = 3, 4, 5, \dots$$

Για παράδειγμα, η δίτιμη τυποποιημένη κατανομή του Πορίσματος 2.2 με πιθανότητα $p_6 = \frac{1}{2} + \frac{\sqrt{15(4\sqrt{10}-5)}}{30}$ έχει έκτης τάξης κεντρική ροπή ίση με $\mu_6 = \frac{4\sqrt{10}-5}{135}$ και ισχύει

$$n(\mu_6 - M_{6,n}) \xrightarrow{d} \lambda_6 \chi_1^2$$

ή

$$n\left(\frac{4\sqrt{10}-5}{135} - M_{6,n}\right) \xrightarrow{d} 15p_6^5 \cdot \frac{7p_6^2 - 7p_6 + 1}{7p_6 - 1} \chi_1^2$$

ή

$$n\left(\frac{4\sqrt{10}-5}{135} - M_{6,n}\right) \xrightarrow{d} \frac{50 - 13\sqrt{10}}{45} \chi_1^2.$$

Για τις συμμετρικές κατανομές του Λήμματος 2.6 αποδεικνύεται ότι η (3.16) δεν πληρούται και ότι $\lambda_k = \tilde{\lambda}_k = \frac{\sqrt{k-1}}{2} k^{\frac{k-1}{2}}$. Έτσι, εφόσον $\mu_k = 0$, από τη σχέση (3.18) συμπεραίνουμε το όριο

$$nM_{k,n} \xrightarrow{d} \frac{\sqrt{k-1}}{2} k^{\frac{k-1}{2}} (\chi_1^2 - \tilde{\chi}_1^2), \quad k = 3, 5, 7, \dots$$

3.3 Πολυδιάστατη μέθοδος Δέλτα ανώτερης τάξης

Τελειώνοντας, παραθέτουμε την πολυδιάστατη μέθοδο Δέλτα ανωτέρας τάξεως, η οποία γενικεύει όλες τις περιπτώσεις.

Θεώρημα 3.5: (πολυδιάστατη μέθοδος Δέλτα m τάξης από τον $\mathcal{R}^k$ στον $\mathcal{R}$)

Έστω ότι $\bar{X}_n$ είναι μια ακολουθία k -διάστατων τυχαίων διανυσμάτων τέτοια ώστε

$$\sqrt{n}(\bar{X}_n - \mu) \xrightarrow{d} W, \quad \text{καθώς } n \rightarrow +\infty,$$

όπου $\mu \in \mathcal{R}^k$ και W είναι k -διάστατο τυχαίο διάνυσμα. Θεωρούμε ότι η Borel-μετρήσιμη συνάρτηση $g: \mathcal{R}^k \rightarrow \mathcal{R}$ είναι m φορές συνεχώς διαφορίσιμη γύρω από τον k -διάστατο μέσο μ και ας είναι j_i το πλήθος των φορών που

παραγωγίζουμε ως προς την μεταβλητή x_i για κάθε $i = 1, 2, \dots, k$. Υποθέτουμε επιπλέον, ότι

$$g_{j_1, \dots, j_k}(\boldsymbol{\mu}) = 0, \quad \forall j_1, j_2, \dots, j_k \in \{0, 1, \dots\} \text{ με } 1 \leq j_1 + j_2 + \dots + j_k \leq m - 1.$$

Τότε,

$$n^{m/2} \left(g(\bar{\mathbf{X}}_n) - g(\boldsymbol{\mu}) \right) \xrightarrow{d} \frac{1}{m!} [D^m g(\boldsymbol{\mu})] \mathbf{W}^m, \quad \text{καθώς } n \rightarrow +\infty,$$

όπου

$$[D^m g(\boldsymbol{\mu})] \mathbf{W}^m = \sum_{j_1 + \dots + j_k = m} g_{j_1, \dots, j_k}(\boldsymbol{\mu}) W_1^{j_1} \dots W_k^{j_k}.$$

Δηλαδή, καθώς $n \rightarrow +\infty$,

$$n^{m/2} \left(g(\bar{\mathbf{X}}_n) - g(\boldsymbol{\mu}) \right) \xrightarrow{d} \frac{1}{m!} \sum_{j_1 + \dots + j_k = m} g_{j_1, \dots, j_k}(\boldsymbol{\mu}) W_1^{j_1} \dots W_k^{j_k}.$$

Απόδειξη: Εφόσον, $g_{j_1, \dots, j_k}(\boldsymbol{\mu}) = 0$ έπεται ότι

$$D^j g(\boldsymbol{\mu})(\mathbf{x} - \boldsymbol{\mu})^j = 0, \quad \forall j = 1, 2, \dots, m - 1.$$

Συνεπώς, από το συμπέρασμα του πολυδιάστατου Θεωρήματος *Taylor*, απορρέει ότι

$$\lim_{\mathbf{x} \rightarrow \boldsymbol{\mu}} \frac{g(\mathbf{x}) - g(\boldsymbol{\mu}) - \frac{D^m g(\boldsymbol{\mu})}{m!} (\mathbf{x} - \boldsymbol{\mu})^m}{\sum_{i=1}^k |x_i - \mu_i|^m} = 0. \quad (3.19)$$

Ακολούθως, έχουμε

$$\begin{aligned} n^{m/2} \left(g(\bar{\mathbf{X}}_n) - g(\boldsymbol{\mu}) \right) &= n^{m/2} \left(g(\bar{\mathbf{X}}_n) - g(\boldsymbol{\mu}) - \frac{1}{m!} [D^m g(\boldsymbol{\mu})] (\bar{\mathbf{X}}_n - \boldsymbol{\mu})^m \right) \\ &\quad + \frac{1}{m!} n^{m/2} [D^m g(\boldsymbol{\mu})] (\bar{\mathbf{X}}_n - \boldsymbol{\mu})^m = T_n + Q_n. \end{aligned}$$

Όμως,

$$\begin{aligned} Q_n &= \frac{1}{m!} n^{m/2} [D^m g(\boldsymbol{\mu})] (\bar{\mathbf{X}}_n - \boldsymbol{\mu})^m = \frac{1}{m!} [D^m g(\boldsymbol{\mu})] \left(\sqrt{n} (\bar{\mathbf{X}}_n - \boldsymbol{\mu}) \right)^m \\ &= \frac{1}{m!} \sum_{j_1 + \dots + j_k = m} g_{j_1, \dots, j_k}(\boldsymbol{\mu}) \left(\sqrt{n} (\bar{X}_{n,1} - \mu_1) \right)^{j_1} \dots \left(\sqrt{n} (\bar{X}_{n,k} - \mu_k) \right)^{j_k} \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{m!} \sum_{j_1 + \dots + j_k = m} g_{j_1, \dots, j_k}(\boldsymbol{\mu}) W_{n,1}^{j_1} \dots W_{n,k}^{j_k} \\
&= h(W_{n,1}, \dots, W_{n,k}) \xrightarrow{d} h(W_1, \dots, W_k) \\
&= \frac{1}{m!} \sum_{j_1 + \dots + j_k = m} g_{j_1, \dots, j_k}(\boldsymbol{\mu}) W_1^{j_1} \dots W_k^{j_k},
\end{aligned}$$

αφού η h είναι συνεχής ως πολυωνυμική.

Μένει να δειχθεί ότι $T_n \xrightarrow{p} 0$. Πράγματι, έχουμε

$$\begin{aligned}
|T_n| &= \left| n^{m/2} \left(g(\bar{X}_n) - g(\boldsymbol{\mu}) - \frac{1}{m!} [D^m g(\boldsymbol{\mu})] (\bar{X}_n - \boldsymbol{\mu})^m \right) \right| \\
&= n^{m/2} \left| g(\bar{X}_n) - g(\boldsymbol{\mu}) - \frac{1}{m!} [D^m g(\boldsymbol{\mu})] (\bar{X}_n - \boldsymbol{\mu})^m \right| \\
&= \frac{n^{m/2} \left| g(\bar{X}_n) - g(\boldsymbol{\mu}) - \frac{1}{m!} [D^m g(\boldsymbol{\mu})] (\bar{X}_n - \boldsymbol{\mu})^m \right|}{n^{m/2} \sum_{i=1}^k |\bar{X}_{n,i} - \mu_i|^m} \cdot n^{m/2} \sum_{i=1}^k |\bar{X}_{n,i} - \mu_i|^m \\
&= \frac{\left| g(\bar{X}_n) - g(\boldsymbol{\mu}) - \frac{1}{m!} [D^m g(\boldsymbol{\mu})] (\bar{X}_n - \boldsymbol{\mu})^m \right|}{\sum_{i=1}^k |\bar{X}_{n,i} - \mu_i|^m} \cdot \sum_{i=1}^k (\sqrt{n} |\bar{X}_{n,i} - \mu_i|)^m = A_n \cdot B_n
\end{aligned}$$

όπου $B_n = \sum_{i=1}^k (\sqrt{n} |\bar{X}_{n,i} - \mu_i|)^m \xrightarrow{d} \sum_{i=1}^k |W_i|^m$, από υπόθεση.

Εν συνεχεία,

$$A_n = \frac{\left| g(\bar{X}_n) - g(\boldsymbol{\mu}) - \frac{1}{m!} [D^m g(\boldsymbol{\mu})] (\bar{X}_n - \boldsymbol{\mu})^m \right|}{\sum_{i=1}^k |\bar{X}_{n,i} - \mu_i|^m} = w(\bar{X}_n),$$

όπου $|w(x)| \rightarrow 0$ για $x \rightarrow \boldsymbol{\mu}$, από (3.19).

Εφόσον, $\bar{X}_n \xrightarrow{p} \boldsymbol{\mu}$ έπεται από το Θεώρημα *Skorokhod* ότι μπορούμε να βρούμε ακολουθία $\bar{Y}_n$ τέτοια ώστε, $\bar{Y}_n(\omega) =_d \bar{X}_n(\omega)$ και $\bar{Y}_n(\omega) \xrightarrow{a.s.} \boldsymbol{\mu}$. Άρα, $|w(\bar{Y}_n(\omega))| \xrightarrow{a.s.} 0$ και συνεπώς, $w(\bar{Y}_n(\omega)) \xrightarrow{p} 0$. Όμως, $w(\bar{Y}_n) =_d w(\bar{X}_n)$, και έτσι, $A_n \xrightarrow{p} 0$.

Επομένως, από το Θεώρημα *Slutsky* προκύπτει ότι

$$A_n \cdot B_n \xrightarrow{p} 0,$$

δηλαδή, τελικά, θα ισχύει

$$T_n \xrightarrow{p} 0,$$

και έτσι, καταλήγουμε ότι

$$n^{m/2} \left(g(\bar{\mathbf{X}}_n) - g(\boldsymbol{\mu}) \right) \xrightarrow{d} \frac{1}{m!} \sum_{j_1 + \dots + j_k = m} g_{j_1, \dots, j_k}(\boldsymbol{\mu}) W_1^{j_1} \dots W_k^{j_k}.$$

■

Βιβλιογραφία

Afendras G. (2013). Moment–based inference for Pearson's quadratic q subfamily of distributions. *Comm. Statist. – Theory Methods* (to appear), arXiv:1104.0040v2.

Βαγγελάτου Ε. (2012). *Μεταπτυχιακές Σημειώσεις Ασυμπτωτικής Στατιστικής*, Εθνικό & Καποδιστριακό Πανεπιστήμιο Αθηνών.

Billingsley P. (1986). *Probability and Measure*, Second Edition, John Wiley, New York.

Chung C.L. (1974). *A Course in Probability Theory*, Second Edition, Academic Press, New York.

Cramér H. (1946). *Mathematical Methods of Statistics*, Princeton University Press, Princeton.

Feller W. (1968). *An Introduction to Probability Theory and its Applications*. Vol. 1, Third Edition, John Wiley, New York.

Fisz M. (1963). *Probability Theory and Mathematical Statistics*, Third Edition, John Wiley, New York.

Freund J., Walpole R. (1980). *Mathematical Statistics*, Third Edition, Prentice–Hall, New Jersey.

Geary R.C. (1936). The distribution of "Student's" ratio for non–normal samples. *J. Roy. Statist. Soc. Ser. B*, **3**, 178–184.

Gut A. (1988). *Stopped Random Walks: Limit Theorems and Applications*. Springer–Verlag, New York.

Kagan A.M., Linnik Y.V. and Rao C.R. (1973). *Characterization Problems in Mathematical Statistics*, John Wiley, New York.

Laha R.G., Lukacs E. and Newman M. (1960). On the independence of a sample central moment and the sample mean. *Ann. Math. Statist.*, **31**, 1028–1033.

Lehmann E.L. (1999). *Elements of Large–Sample Theory*, Springer, New York.

Παπαδάτος Ν. (2006). *Θεωρία Πιθανοτήτων*, Εθνικό & Καποδιστριακό Πανεπιστήμιο Αθηνών.

Παπαδάτος Ν. (2013). "*Ομοιόμορφη Ολοκληρωσιμότητα τυποποιημένων αθροισμάτων ανεξαρτήτων και ισονόμων τυχαίων μεταβλητών με πεπερασμένη διασπορά*".

Roussas G. (1997). *A Course of Mathematical Statistics*, Second Edition, Academic Press, New York.

Serfling R.J. (1980). *Approximation Theorems of Mathematical Statistics*, John Wiley, New York.

Φερεντίνος Κ. & Παπαϊωάννου Τ. (2000). *Μαθηματική Στατιστική*, Πανεπιστήμιο Ιωαννίνων.

Van der Vaart A.W. (1998). *Asymptotic Statistics*, Cambridge University Press, New York.

Wikipedia (2013). Βιογραφικά Σημειώματα, *Harald Cramér*, *Brook Taylor* & άλλων επιστημόνων.

Zinger A.A. (1958). Independence of quasi polynomial statistics and analytical properties of distributions. *Teor. Veroy. i Primenen*, **3**, 265–284 (in Russian).